Physical aspects of origins of life scenarios

Alex Blokhuis

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'ESPCI Paris

Aspects physiques des scénarios d'origines de la vie

Soutenue par
Alexander Willem Peter Blokhuis
Le 18 11 2019

École doctorale n°564
Physique en île de France

Spécialité
Physique

Composition du jury :

Annie Lemarchand Directeur de recherche, Sorbonne Université - CNRS	*Présidente*
Ulrich Gerland Full Professor, Technical University of Munich	*Rapporteur*
Sandeep Krishna Associate professor, National Centre for Biological Sciences - TIFR	*Rapporteur*
Massimiliano Esposito Full Professor, University of Luxembourg	*Examinateur*
Ludovic Jullien Professeur des universités, ENS - SU	*Examinateur*
Pierre Gaspard Full Professor, Université Libre de Bruxelles	*Examinateur*
David Lacoste Directeur de recherche, ESPCI Paris - CNRS	*Directeur de thèse*
Philippe Nghe Maître de conférences, ESPCI Paris	*CoDirecteur de thèse*

ESPCI PARIS | PSL

Acknowledgements

David, you provided me with the freedom to explore a lot of different things, but never failed to remind me of an essential point: we cannot do everything, and we need to finish a project before we burden ourselves with more. I value this as the most important lesson of this thesis, and it cannot be repeated enough. Our endless discussions have been a journey of discovery. Most importantly, you showed me the qualities a scientist needs: the patience to ponder on the most appropriate notation, a capacity to be concise, and choosing one's words with care. It has been a privilege learning from you. I hope you will have many more PhD students that will share this privilege with me.

Working on the origins of life has something absurd to it, one can easily feel lost. I was fortunate enough to have an excellent guide. Philippe, your rationality is matched by a great sense of humor. In a field where we do not even know what the right questions are, our habit of questioning everything, even ourselves (notably through Cyrille's anthropological work) felt like a breath of fresh air. I fondly look back at our theory discussions, notably those that grew into Ch.5, and look forward to extending our work.

Andrew, thank you for your time and trust. After our inspiring conversations, I would leave the office with plenty of new ideas, suggestions and tablet notes on things to test and try. Now that the thesis is done, I can give them full attention and address them at the rate they pop up.

Luca, thank you for teaching me, for the articles you sent me to read and engaging in long discussions on thermodynamics, information and evolution. It has been a pleasure working with you on transient compartments. Without your input, encouragement and information, chapter 4 would not have seen the light of day. I hope we may one day build that macroscopic information engine.

I wholeheartedly thank Pierre Gaspard for his careful reading and insightful feedback on our work, as well as for our fruitful collaborations. Working together on CSTRs and Serial Transfer has been a delight. Our discussions, such as those on diffusiophoresis, have provided lasting food for

thought. I hope I can return the favor.

I wish to thank the OCAV program for their support (ANR-10-IDEX-0001-02) and providing a platform for the exchange of ideas. Through François Guyot I got a deep appreciation for the key role of geology in the origins of life and to this day I am trying to fill this hiatus in my knowledge. I also wish to acknowledge Christophe Malaterre, who taught me about the intellectual history behind the notion of 'chemical evolution', leading to useful insights and discussions.

Sandeep Ameta and Simon, thank you for teaching me the ins and outs of experimental RNA biochemistry.

I humbly thank my jury for accepting to read the manuscript, which has become bigger than anticipated, even without the experiments. I sincerely hope that a fascination for the topic may somewhat counterbalance the burden of reading.

Daan van de Weem and Anatoliy Babic enlightened me on the joy of branching processes, renewal theorems and other branches of mathematics. Chapter 6 and 8 are a testament to your useful suggestions. Heel erg bedankt.

I wish to thank Sandeep Krishna for interesting discussions and useful suggestions, and his swift rescue during my stay in Bangalore.

I thank our collaborators at ENS Ludovic Jullien, Jean Pierre-Fitoussi and Fabien Ferrage, as well as everyone else in the formose project: Heng Lu, Cyrille Jeancolas, Éstanislau Guilherme, Gabrielle Woronoff, Rebecca Turk Macleod, Eörs Szathmary, Andrew Griffths and Philippe Nghe.

Joachim H. Rosenberger and Tobias Göppel reignited my passion for polymer length distributions, after a lengthy discussion about their experimental results. I warmly acknowledge their input, which led me to brush up calculations which have been included in Ch.7.

I wish to thank Yannick for his scientific advice, interesting discussions and important suggestions.

I wish to thank Anton for our interesting discussions on compartments, evolution, the Gibbs paradox, dynamical systems, the Leontief matrix and many more. Your exposé on the fossil record and natural history has dramatically increased the rate at which I go to the natural history museum.

I wish to thank Zorana for her support, feedback and keen critical sense. Your comments have been well taken and they formed the basis of the reflections that led to Ch. 5.

My long discussions with Jean-François Derivaux and Basile Nguyen have been thought-provoking and productive.

I wish to thank N. Lehman for the critical reading of our work, along with R. Mizuuchi, L. Vincent, N. and D. Baum for a fruitful collaboration.
I wish to thank Reinaldo for insightful discussions and critical reading of our work.

I acknowledge stimulating exchanges with Bahram Houchmandzadeh.

I wish to thank my colleagues in the lab, old and new, and everyone else that made it double the fun to come to the office.

I extend my gratitude to all my friends and family for their moral support, especially the last 6 months.

I am forever indebted to Oriane. You helped me out in every way imaginable. Whether it was teaching me all the tricks Illustrator has to offer to draw my figures, or helping me out with my home experiments, or simply brightening up my day. You have come to refer to this thesis as my 'mistress', after having to endure my numerous long nights of redacting without end. I have a hard time telling who's more relieved now that this chapter is closed.

Publication list

Published

[1] A. Blokhuis, D. Lacoste. Length and sequence relaxation of copolymers under recombination reactions. *J. Chem. Phys.* 147: 94905, 2017

[2] A. Blokhuis, D. Lacoste, P. Nghe, L. Peliti. Selection dynamics in transient compartmentalization. *Phys. Rev. Lett.* 120: 158101, 2018

[3] A. Blokhuis, D. Lacoste, P. Gaspard. Reaction kinetics in open reactors and serial transfers between closed reactors. *J. Chem. Phys* 148: 144902, 2018

[4] R. Mizuuchi, A. Blokhuis, L. Vincent, P. Nghe, N. Lehman, D. Baum. Mineral surfaces select for longer RNA molecules. *Chem. Commun.* 55: 2090, 2019

[5] A. Blokhuis, P. Nghe, L. Peliti, D. Lacoste. The generality of transient compartmentalization and its associated error thresholds. *J. Theor. Biol.*, 487, 2020, 110110

Other

[6] A. Blokhuis, D. Lacoste, P. Nghe. Autocatalysis in chemical networks: unifications and extensions. (in preparation)

[7] A. Blokhuis, D. Lacoste. The kinetic and thermodynamic role of activation in prebiotic polymer scenarios. (in preparation)

[8] H. Lu, A. Blokhuis, C. Jeancolas, P. Pelupessy, G. Éstanislau, G. Woronoff, R.T. Macleod, F. Ferrage, L. Jullien, E. Szathmary, A. Griffiths, P. Nghe. Natural selection of compartmentalized autocatalytic chemical reactions. (in preparation, order of authors to be determined)

[9] A. Blokhuis, D. Lacoste, L. Peliti. Macroscopic Information Engines. (in preparation)

[10] A. Blokhuis. Thermodynamically allowed current propensities and the 0th law of thermodynamics. (in preparation)

Sommaire en Français

La science qui s'intéresse aux scénarios des origines de la vie (Origins of Life, ou tout court OOL), cherche à expliquer l'émergence d'une biosphère à partir de matière abiotique. Ce processus est appelé «abiogenèse», mais son fonctionnement n'est pas clair. Pour mieux comprendre l'abiogenèse, les chercheurs en OOL combinent des connaissances de divers domaines, tels que la biologie, la chimie, la géologie, la physique, l'astronomie, l'histoire des sciences et bien d'autres. Certaines de ces idées ont conduit à la formulation de scénarios prébiotiques : spéculation sur le lieu, la chimie et les mécanismes physiques de l'abiogenèse.

Dans cette thèse, nous introduisons des cadres rigoureux, pour l'étude systématique des aspects physiques de l'abiogenèse. Ces cadres s'appuient sur des connaissances récentes en thermodynamique hors équilibre, en réseaux de réactions chimiques et sélection sur plusieurs niveaux. Ils soulignent la cohérence thermodynamique et la structure fondamentale de la chimie. Nous identifions des contraintes importantes sur les scénarios ainsi que de nouveaux mécanismes d'émergence et d'évolution en chimie et au-delà.

Le plan de la thèse est le suivant :

dans le chapitre 1, une introduction critique au domaine des origines de la vie est donnée, soulignant i) ce que nous pouvons raisonnablement considérer comme connu, ii) ce que supposent les scénarios populaires et iii) les développements historiques qui ont façonné la pensée actuelle.

i) Si la vie est originaire de la terre, notre géologie contraint les éléments chimique qu'elle pourrait utiliser et sa fenêtre temporelle. La terre est considérée comme ayant 4,5 milliards d'années, tandis que des fossiles attribués à la vie microbienne (microfossiles) ont été trouvés remontant à environ 3,8 milliards ans. Une théorie scientifique sur ce qui se passe entre les deux ne devrait pas être en contradiction avec la thermodynamique, les lois de conservation, la cinétique et la réactivité chimiques ou d'autres aspects établis de la chimie et de la physique modernes. Ces contraintes constituent la base des résultats qui seront dérivés dans nos cadres.

ii) Au-delà de ce qui est considéré comme «connu», les scénarios actuels nécessitent des hypothèses supplémentaires, notamment sur «ce qui peut et ne peut pas se produire dans l'évolution chimique», «quels produits chimiques sont arrivés en premier» et «où la vie a commencé».

La famille des scénarios du monde de l'ARN considère les polymères génétiques comme

indispensables pour établir une progression semblable à l'évolution. Comme pour le scénario Oparin-Haldane pour les protéines, il est souvent postulé que des événements très rares ont fourni les espèces clé pour déclencher l'abiogenèse, comme une réplicase d'ARN à base d'ARN, et beaucoup considèrent que l'abiogenèse commence par l'ARN. Au sein de la communauté mondiale de l'ARN, toutes ces hypothèses sont remises en question : certains considèrent l'AXN précédant l'ARN thermodynamiquement plus favorable, d'autres qu'une réplicase à base d'ARN est invraisemblable et voient plutôt l'ARN évoluer à travers des réseaux catalytiques.

Le scénario du monde fer-soufre est une description détaillée d'une transition d'une chimie adsorbée sur des minéraux (par exemple la pyrite), vers des bicouches superficielles lipophiles, qui se détachent finalement pour former des protocoles. Le scénario considère une chaîne déterministe de changements chimiques maintenue par l'autocatalyse, avec une arrivée tardive de la génétique. La chimie utilise des espèces rencontrées dans la biochimie moderne, mais les utilise de différentes manières pour la chimie de surface. Le scénario est souvent placé dans des évents hydrothermaux acides. D'autres scénarios «métaboliques» changent dans certains de ces détails : certains démarrent directement avec des protocoles. D'autres considèrent qu'un scénario plus parcimonieux peut être formulé dans des évents hydrothermaux alcalins. Tous décrivent l'abiogenèse comme un processus de réorientation des biomolécules existantes vers de nouvelles utilisations.

Les scénarios du monde lipidique, suggérés par des modèles informatiques comme GARD de Lancet et al, soulignent que des états de composition distincts (composomes), maintenus par autocatalyse, peuvent servir de précurseur d'un génome. Il est supposé qu'une abondance d'amphiphiles distincts ont été créés prébiotiquement, et que l'évolution chimique était initialement due à l'assemblage non-covalent et à la division des micelles ou des vésicules.

Une hypothèse commune qui est partagée entre les scénarios est que certains éléments de la vie existante devaient être là dès le départ (X-first). Cela peut par exemple être un génome (genes-first), un métabolisme (metabolism-first) analogue au métabolisme moderne ou un compartiment amphiphile (compartments-first, lipids-first). Plus généralement, la vie est expliquée en termes de biomolécules existantes, mais les approches diffèrent dans leur hypothèse quant à celles qui devaient venir en premier. Ces hypothèses sont clairement contradictoires : elles ne peuvent pas être vraies en même temps. Tout au long de l'histoire de la recherche OOL, l'hypothèse de la biochimie d'abord a été remise en question à plusieurs reprises. Étant donné le grand espace de structures possibles occupé par la matière non biochimique, il est possible que ce qui est arrivé en premier n'ait même pas encore été proposé.

Plus généralement, les suggestions sur le début de la vie conduisent à des dilemmes ou à des multilemmes. Les dilemmes bien connus sont l'ARN d'abord contre les peptides d'abord et le métabolisme d'abord contre la génétique d'abord et la vie a commencé dans les étangs contre les évents hydrothermaux. Bien que de tels dilemmes facilitent la communication, il convient de garder à l'esprit qu'ils sont tous de fausses dichotomies : la littérature suggère beaucoup plus de deux réponses pour chaque problème, et il est fort possible que les bonnes réponses n'aient pas encore été proposées. Cela est illustré par le fait que de nouvelles réponses à ces questions sont proposées à ce jour.

Les visions modernes de la recherche Origines de la vie sont fortement marquées par l'histoire du domaine, dont nous rappellerons une petite partie. L'idée que les formes de vie modernes se forment spontanément (génération spontanée) a été réfutée de manière convaincante par L. Pasteur. Oparin a estimé que l'abiogenèse aurait dû se produire au moins une fois, mais de manière non triviale. Sur la base d'idées sur la terre et l'atmosphère primitives, Oparin (et plus tard, Haldane) considérait les océans comme un «bouillon prébiotique», dans lequel les oligopeptides s'accumulaient pour finalement former des protocoles et répliquer des peptides. Certaines autres idées influentes se sont concentrées sur l'imitation de différentes propriétés réalistes par la chimie physique, telles que les jardins chimiques de Leduc et le protoplasme d'Herrera.

Lorsque les acides aminés ont été trouvés dans la simulation Urey et Millers d'un bouillon prébiotique, alimenté par une atmosphère sous une décharge électrique, la théorie du bouillon prébiotique est devenue l'image dominante et est rapidement devenue une partie des programmes scolaires. Avec l'avènement de la génétique, l'idée de l'ARN agissant à la fois comme catalyseur et génome a commencé à être envisagée, afin de résoudre une énigme de «la génétique d'abord contre les enzymes d'abord». Dans les années 80, une telle activité catalytique a été trouvée. Un éditorial très influent de Gilbert en 1986 a suggéré un scénario axé sur l'ARN, et le terme monde d'ARN a été inventé, qui est rapidement devenu une idée dominante.

Les jardins chimiques ont connu une résurgence avec la découverte de grands champs de monts hydrothermales. En tant que source généreuse d'énergie libre et de produits chimiques, les sources hydrothermaux ont conduit à l'épanouissement de scénarios métaboliques, ce qui conduirait au dilemme apparent du métabolisme d'abord contre la génétique d'abord. Une autre influence majeure fut le livre de Gold, paru en 1992 : «The Deep Hot Biosphere», suggérant que la vie aurait émergé sur une terre Hadéenne, à plusieurs kilomètres sous la surface. Après, la vie a été retrouvée à de grandes profondeurs, et de nouvelles explorations de ses revendications sont activement en cours.

Une publication de 1996 d'un fossile biogénique potentiel de Mars dans la météorite ALH84001 a conduit à une augmentation de la recherche en exobiologie, un nouvel institut d'astrobiologie et la réintégration des programmes d'exploration sur Mars. Des recherches ultérieures suggèrent que les signatures réalistes d'ALH84001 pourraient également être générées par des processus abiotiques. Cependant, il a fourni un tremplin pour que l'exobiologie devienne un domaine majeur en OOL. L'échantillon d'idées présenté ici illustre le fait que la recherche OOL est dispersée autour de différentes idées. En effet, des études bibliométriques exhaustives trouvent de grandes sous-communautés en OOL, dont certaines se connaissent ou se citent à peine, les «microfossiles et témoignages de la vie sur la terre primitive» étant les plus déconnectés. Les disciplines scientifiques qui composent ces communautés se révèlent très différentes. La pluridisciplinarité inhérente à la recherche en OOL pose des défis, car les chercheurs maîtrisent rarement toutes les disciplines qui contribuent à la recherche en OOL. Les programmes de recherche en OOL lancent aujourd'hui des initiatives pour combler les principales lacunes dans les connaissances et les attitudes disciplinaires. Une majorité d'auteurs en OOL ont écrit et continuent d'écrire que la compréhension de la thermodynamique hors équilibre est essentielle. Paradoxalement, la plupart de ces auteurs pensent qu'il n'y a actuellement aucun cadre théorique pour cela, malgré son développement actif depuis plus de 40 ans. Un objectif important de ce manuscrit est d'utiliser et de vulgariser la thermodynamique hors équilibre et la thermodynamique stochastique comme un outil pour obtenir un aperçu fondamental des problèmes en OOL.

Dans le chapitre 2, nous passons en revue le formalisme de la matrice stoechiométrique pour les réseaux chimiques, qui permet d'analyser les aspects topologiques et thermodynamiques généraux des systèmes chimiques avec un nombre arbitraire de réactions et de réactifs. La théorie est illustrée en pensant au chimiste expérimental, en se concentrant sur des exemples et des situations chimiques. Pour notre approche, nous introduisons quelques conventions à motivation chimique, qui seront fructueuses dans la dérivation des résultats dans les chapitres suivants.

Une matrice stœchiométrique contient les changements stœchiométriques du nombre de réactifs dans toutes les réactions d'un réseau de réactions. Son contenu peut également être représenté plus visuellement, par exemple via des hypergraphes ou des graphes bipartites. De telles représentations deviennent rapidement impraticables avec l'augmentation de la taille du réseau. Une représentation plus pratique (graphiques simples) peut être faite avec une matrice d'incidence, qui relie des «complexes», des collections d'espèces chimiques qui réagissent ensemble dans une réaction.

En traitant de la thermodynamique, il est souhaitable qu'il existe une inversion bien définie des transitions microscopiques pertinentes. Par conséquent, une première convention sera que toute

réaction chimique décrite est définie dans les deux sens. Une deuxième convention sera qu'une étape de réaction implique au plus deux espèces distinctes (par exemple des complexes, des molécules) à la fois. Une réaction d'ordre supérieur peut toujours être décomposée en étapes impliquant au plus deux molécules. Enfin, nous introduisons une convention appelée «non-ambiguïté», ce qui signifie qu'au sein d'une étape de réaction, une espèce n'est pas à la fois un réactif et un produit. Ces conventions ne limitent en rien la chimie que nous pouvons décrire, mais elles nous obligent à inclure suffisamment d'étapes pour le faire. Une conséquence est que la matrice stœchiométrique contient désormais des coefficients stœchiométriques au lieu de changements dans le nombre de réactifs, ce qui signifie qu'il ne peut correspondre qu'à un seul réseau. Ce sera déterminant pour révéler certains motifs.

Les propriétés topologiques d'un réseau chimique sont contenus dans ses sous-espaces fondamentaux. L'espace nul droit (noyau) correspond à des combinaisons de réactions qui laissent le système inchangé, appelées cycles. L'espace nul gauche (cokernel) de la matrice stœchiométrique décrit les quantités conservées : combinaisons linéaires d'espèces inchangées par les réactions chimiques dans le système. Nous distinguons deux types : i) lois de conservation de type masse, dans lesquelles tous les coefficients sont positifs (par exemple nombre d'atomes de carbone), ii) lois de conservation mixtes, dans lesquelles une différence est conservée (par exemple charge totale). Le nombre de cycles indépendants et de lois de conservation est lié au nombre de réactions et d'espèces par le théorème de nullité de rang.

Pour inspecter certains aspects chimiques d'un réseau, il s'avérera indispensable d'étudier les sous-réseaux. Nous décrivons une approche systématique pour ce faire en utilisant des sous-matrices, obtenues en supprimant des lignes et/ou des colonnes de la matrice d'origine. Une telle opération peut par exemple être utilisée pour étudier l'effet des chimiostats, qui fixent la concentration d'un composé par échange avec un réservoir. Cela permet de «retirer» le composé de la description efficace et un seul chimiostat enfreint nécessairement une loi de conservation au sein du système.

Une simple réaction en chaîne ou à un mécanisme de réaction de Michaelis-Menten, donne une bonne impression sur ce que font ces concepts. Les catalyseurs de ces schémas ont leur propre loi de conservation de type masse. Lorsqu'elles sont étudiées en isolément, dans une sous-matrice, un nouveau cycle émerge, correspondant exactement au cycle catalytique. Cette idée est précisée en Ch.5 et Ch.6, où le cadre est appliqué pour étudier la catalyse et l'autocatalyse en général.

Dans le chapitre 3, quelques aspects thermodynamiques des réseaux chimiques ouverts sont examinés. Un point clé que nous souhaitons illustrer est qu'un réseau chimique peut être ouvert de différentes manières. Les détails de l'ouverture d'un système à un environnement sont d'une importance cruciale pour son comportement, ce qui est illustré par le traitement des chimiostats simples et des chimiostats composites, un réacteur CSTR, un transfert en série et des compartiments à couplage osmotique. Ouvrir un système chimique par une de ces manières le soumet à une dynamique et à des lois de conservation très distinctes.

Le chimiostat thermodynamique idéal décrit l'échange d'une espèce chimique avec un grand bain qui, en raison d'un échange rapide, fixe son potentiel chimique et les fluctuations de concentration. Dans ce cas, il peut être vu comme un réservoir infini séparé du système par une membrane parfaitement spécifique échangeant des molécules uniques. Dans la pratique expérimentale, le rôle d'un «bain» peut être rempli par des composés au sein du système qui servent de tampon (chimiostats homogènes), dont la capacité est intrinsèquement limitée par la taille du système. Un bain peut également être dû à des composés dans d'autres phases (chimiostat externe), comme une phase fluide, des précipités, des gaz, etc. Souvent, le potentiel chimique des espèces réservoirs est couplé à d'autres espèces, conduisant à un chimiostat composite (par exemple ions dissous en équilibre avec leur sel). De tels chimiostats fixent les produits de concentrations au lieu de concentrations individuelles, ce qui signifie que les espèces chimiques considérées comme des

«aliments» ne sont plus indépendantes. Nous montrons que dans l'annexe que les chimiostats thermodynamiques nécessitent un traitement plus détaillé en thermodynamique stochastique et impliquent une loi de zéros modifiée pour l'échange de quantités entières conservées (par exemple des atomes).

Dans un réacteur à cuve à agitation continue (Continuously stirred tank reactor, CSTR), un mélange chimique entre et sort d'un réacteur bien agité. L'écoulement de sortie sert de voie de dégradation inhérente, et il peut être inclus dans la matrice stoechiométrique en tant que tel (Cependant, il s'agit d'un écoulement de liquide macroscopique. Une interprétation microscopique comme pour les réactions réversibles ne s'applique donc pas). Sous l'action de cette dégradation, toutes les lois de conservation du réseau sont brisées.

Cependant, de manière asymptotique (c'est-à-dire pour $t \to \infty$), l'équilibre entre les flux entrants et sortants conduit à des contraintes de déséquilibre similaires à celles des lois de conservation pour le système fermé : la masse qui entre en équilibre la masse qui sort, et un écoulement stationnaire obéit aux lois de conservation d'un réacteur fermé. Lorsque l'écoulement est rapide par rapport à la chimie, la composition chimique du CSTR peut rester loin de l'équilibre. Lorsque la chimie dépasse fortement le débit, le CSTR se rapproche du comportement d'un réacteur fermé. Une démonstration claire de ce fait peut être fournie en termes de chimie des polymères en étudiant les distributions de longueur.

Une manière similaire d'ouvrir le système consiste à transférer à plusieurs reprises une fraction f du mélange réactionnel vers un approvisionnement en produits chimiques frais, après un temps Δt s'est écoulé. Ce transfert en série est souvent utilisé comme simulation expérimentale d'un CSTR qui est plus facile à configurer et est parfois décrit comme équivalent. Nous montrons que le temps de séjour effectif $\tau_{eff} = \Delta t/(1-f)$ devient équivalent au temps de séjour τ dans un CSTR, dans la limite où $\Delta t \to 0$, qui est quand $f \to 1$.

Il existe donc une analogie étroite entre le CSTR et le transfert série. Cependant, le transfert sériel expérimental n'est généralement pas effectué dans les limites $\Delta t \to 0$, $f \to 1$, ce qui est un régime très peu pratique pour travailler. La modélisation détaillée et les arguments de l'échelle de temps montrent que, même en dehors de cette limite, un bon accord entre la composition du CSTR et la composition du transfert en série peut souvent encore être atteint.

Enfin, nous examinons les compartiments couplés par osmose (gouttelettes), qui peuvent échanger des petites molécules (dont leur solvant) par diffusion. Ces systèmes sont ouverts et peuvent croître et se diviser. Cependant, pour croître et se diviser de manière persistante, les gradients appropriés doivent être maintenus. Pour deux situations simples (échange avec un réservoir, échange avec une gouttelette voisine régulièrement rafraîchie), nous considérons un certain nombre de situations pour maintenir un tel gradient, pour croître et se diviser de façon persistante.

Lorsque les gouttelettes ont une chimie identique, il est facilement démontré que les gradients s'annulent rapidement, ce qui signifie qu'un cycle de division de croissance conduira finalement à l'élimination totale de la gouttelette. Lorsqu'une gouttelette a une chimie distincte, elle doit avoir des composés distincts. Nous pouvons alors imaginer qu'une chimie distincte se produise parce que ces composés : I. sont consommés dans une réaction distincte. II. sont des catalyseurs qui ne se forment pas (allocatalyse). III. Sont des catalyseurs qui se forment (autocatalyse). Les considérations de droit de la conservation montrent que I et II sont exclus. L'autocatalyse, cependant, peut maintenir un cycle de croissance-division avec un seul voisin ou réservoir. Cette conclusion est illustrée par un modèle de jouet pour la réaction de Formose dans un système de gouttelettes couplées, ce qui conduit à une récurrence de Poincaré stable.

Dans le chapitre 4, le concept de l'information en thermodynamique est discuté d'une manière chimiquement explicite. Le concept d'information a semé la confusion dans la littérature OOL, notamment parce qu'il est souvent confondu avec des codes (génétiques). Pour démontrer le

caractère thermodynamique explicite des informations, nous commençons par une discussion sur le «paradoxe de Gibbs» et le niveau de description. Par la suite, nous montrons comment l'information peut être rendue très concrète en chimie : nous montrons comment un moteur d'information macroscopique peut extraire le travail de la racémisation d'énantiomères purs. De la même manière, des protocoles peuvent être conçus pour convertir de manière réversible ($\Delta S = 0$) des énantiomères purs de leur image miroir. L'annexe développe une discussion sur la thermodynamique hors équilibre et les réseaux de réactions qui distinguent deux produits (tels que deux énantiomères).

Le paradoxe de Gibbs est une expérience de pensée, dans laquelle deux gaz sont autorisés à mélanger. Si les gaz sont distincts, défaire ce mélange requiert une certain quantité de travail. Si les gaz sont les mêmes, alors aucune différence significative n'a été introduite par le mélange, et il suffit de réintroduire une partition au milieu. Car Gibbs a commencé avec une expression qui exigeait également du travail pour séparer deux gaz équivalents, une paradoxe est apparu. Le paradoxe a été résolu en ajoutant une contribution manquante à l'expression que nous identifierions aujourd'hui avec une entropie de mélange ou l'information.

En mécanique statistique, le mélange d'entropie découle immédiatement de considérations combinatoires, en considérant la dégénérescence des états due aux permutations. Cette dégénérescence est dans l'œil du spectateur, c'est-à-dire que le détail du niveau de description détermine si les états sont traités de manière équivalente. Les colloïdes en dispersion ne sont pratiquement pas atomiquement identiques, et les macromolécules n'ont pas la même composition isotopique. Dans la plupart des cas, notre description est trop grossière pour tenir compte de ces aspects : en mécanique statistique, l'entropie de mélange découle immédiatement de considérations combinatoires, en considérant la dégénérescence des états due aux permutations. Cette dégénérescence est dans l'œil du spectateur, c'est-à-dire que le détail du niveau de description détermine si les états sont traités de manière équivalente. Les colloïdes en dispersion ne sont pratiquement pas atomiquement identiques, et les macromolécules n'ont pas la même composition isotopique. Dans la plupart des cas, notre description est trop grossière pour tenir compte de ces aspects. Notre fonction d'entropie ne tient alors tout simplement pas compte de ces détails : alors notre fonction d'entropie ne les contient pas.

Cadré comme tel, le paradoxe de Gibbs n'est pas un paradoxe. Il s'agit d'un exemple macroscopique de perte d'information, en faisant passer le système vers un état mixte qui est plus probable. Si nous avions choisi une description moins détaillée, ce changement entropique serait imperceptible. Il est cependant très réel et nous pouvons l'utiliser pour extraire le travail et vice versa. À cette fin, nous décrivons un moteur d'information macroscopique, qui exploite la configuration d'une molécule chirale.

Nous commençons par une première chambre (I) remplie de molécules de gaz pur d'une configuration (S). Une membrane laisse passer les molécules S jusqu'à une deuxième chambre (II), qui est initialement contractée. Un catalyseur de racémisation y interconvertit S et son image miroir R. A l'équilibre, cela conduit à doubler la pression dans la chambre II. Les deux chambres ont des pistons mobiles qui sont couplés et via un protocole réversible, l'entropie de mélange complète due à la racémisation peut être extraite sous forme de travail (par contact avec un bain de chaleur) de ces pistons.

Lors du couplage avec un second bain de chaleur, un moteur peut être construit, analogue à un moteur Carnot, mais où l'expansion du volume isotherme a été remplacée par la racémisation isotherme. Par conséquent, l'efficacité Carnot est récupérée. Alternativement, l'espèce chirale peut être utilisée comme 'carburant non-combustible' et le racémate comme gaz d'échappement. Nous montrons qu'un moteur autonome à quatre chambres peut être construit de cette façon, couplé à des réservoirs d'énantiomères.

Enfin, nous décrivons une configuration à trois chambres, avec une membrane spécifique S, un catalyseur de racémisation et une membrane spécifique R, avec des pistons mobiles à chaque

extrémité. Grâce à un fonctionnement quasi-statique, ces pistons permettent la conversion réversible de R pur en S. pur. Ceci est strictement interdit dans un seul compartiment, où le mélange d'entropie conduit à un seul minimum d'énergie libre. Ici, chaque compartiment conserve une composition fixe, seule la taille est modifiée.

En annexe, nous examinons les structures de réseaux chimiques qui améliorent la synthèse d'un composé, en ajoutant des réactions irréversibles ou des échanges dynamiques (réversibles). Plusieurs réseaux importants sont discutés en détail. La dichotomie classique entre un produit cinétique ou un produit thermodynamique ne vaut que pour le plus simple des réseaux, et la chimie des systèmes découvre rapidement de nouvelles stratégies élégantes pour pousser les réactions vers de nouveaux extrêmes en termes d'efficacité. La simplicité de ces réseaux chimiques correcteurs d'erreurs (certains n'impliquent qu'une seule réaction supplémentaire) est provocatrice : les réseaux fonctionnels abiotiques peuvent être considérablement plus répandus que ce qui est actuellement considéré et il sera instructif d'en tenir compte dans la conceptualisation de l'évolution chimique.

Dans le chapitre 5, l'élucidation stoechiométrique de l'allocatalyse et de l'autocatalyse est développée. Le point de départ est la définition officielle IUPAC de la catalyse et de l'autocatalyse. Ensuite, différentes formes de catalyse sont illustrées à titre d'illustration. Par la suite, nous montrons comment les propriétés définies au chapitre 2 comme la non-ambiguïté permettent une caractérisation générale. Cela permet, pour la première fois, d'identifier ces caractéristiques (c'est-à-dire les motifs de réseau catalytique) dans les réseaux de réaction en général, grâce à l'utilisation de techniques sous-matricielles. Nous montrons également que l'échange entre compartiments peut conduire à l'émergence de nouveaux types de cycles autocatalytiques impossibles pour des compartiments uniques : l'autocatalyse par multiples compartiments.

Il existe différentes définitions contradictoires de la catalyse dans la littérature OOL et dans les ontologies numériques. Pour obtenir des résultats qui se rapportent à la chimie, nos définitions doivent correspondre à leur utilisation correcte en chimie. Cela implique d'utiliser la définition de catalyseur recommandée par l'IUPAC : «Une substance qui augmente la vitesse d'une réaction (*nette*) sans modifier le changement d'énergie standard de Gibbs dans la réaction; le processus est appelé catalyse. Le catalyseur est à la fois un réactif et un produit de la réaction (*catalysé*). Les mots catalyseur et catalyse ne doivent pas être utilisés lorsque la substance ajoutée réduit la vitesse de réaction (voir inhibiteur). La catalyse peut être classée comme catalyse homogène, dans laquelle une seule phase est impliquée, et catalyse hétérogène, dans laquelle la réaction se produit à ou près d'une interface entre les phases. La catalyse provoquée par l'un des produits d'une réaction (*nette*) est appelée autocatalyse. La catalyse provoquée par un groupe sur une molécule de réactif elle-même est appelée catalyse intramoléculaire. Le terme catalyse est également souvent utilisé lorsque la substance est consommée dans la réaction (par exemple : hydrolyse catalysée par une base d'esters). Strictement, une telle substance devrait être appelée un activateur. » Nous voyons que l'autocatalyse est une forme particulière de catalyse, dans laquelle le catalyseur se produit lui-même. Dans ce qui suit, il sera important de faire la distinction entre les catalyseurs qui se forment eux-mêmes et ceux qui aident uniquement à former d'autres espèces. La cohérence lexicologique nous oblige à désigner ce processus sous le nom d'allocatalyse.

Un exemple classique d'allocatalyse est une réaction catalytique en une seule étape comme $S + E \rightleftharpoons E + P$. Puisque nous exigeons la non-ambiguïté par convention, des étapes supplémentaires doivent être introduites, ce qui conduit à un mécanisme de type Michaelis-Menten. Cependant, fournir la stœchiométrie ne suffit pas : la définition de l'IUPAC spécifie qu'une accélération de la vitesse doit se produire. Un composé peut accélérer les réactifs à une température (et donc être un catalyseur) et les piéger à une autre (et donc être un inhibiteur); un composé n'est pas un catalyseur en soi, mais seulement dans un contexte donné.

Contrairement à certaines autres définitions, la définition de l'IUPAC ne nécessite pas le retour du «catalyseur d'origine», une nouvelle copie est également valide. Cela signifie que les étapes

de propagation dans les réactions en chaîne sont également des exemples de catalyse. Nous pouvons également construire un cycle catalytique en couplant une réaction autocatalytique avant et arrière. Un tel schéma obéit à la définition de l'IUPAC de la catalyse, mais est clairement d'un caractère différent de l'allocatalyse sans autocatalyse : il ne vient pas avec une voie inhérente à l'accumulation ou à la dégradation. Comme nous le verrons plus loin, de tels cas peuvent être supprimés en exigeant que l'allocatalyseur ait une loi de conservation de type massique.

Une propriété clé pour identifier les motifs catalytiques dans les sous-matrices se révélera être l'absence de réactions vides ($\emptyset$), ou 'autonomie'. Formulé positivement, nous exigeons que chaque réaction ait au moins un réactif et au moins un produit. Cette propriété capture qu'un cycle catalytique (dans les deux sens) ne peut se produire que sous réserve de la présence d'un catalyseur, et qu'un tel catalyseur est consommé pour former d'autres catalyseurs et est finalement formé à nouveau à partir d'un autre catalyseur. Le concept d'autonomie est lié au concept de siphon, dans un cadre différent, où il s'est avéré être une caractéristique clé pour une définition particulière (non-IUPAC) de la catalyse.

Nous pouvons formaliser les exigences stoechiométriques pour l'allocatalyse, en définissant l'allocatalyse stoechiométrique :

Un ensemble de substances $\{\mathrm{X_k}\}$ qui permettent une transformation d'autres espèces, sans consommation ou production nette de ces substances. La transformation se produit à travers un cycle avec le vecteur de réaction $\boldsymbol{c}^*$, qui conduit à une conversion nette des espèces externes Y ($\forall\, i,j\ \ \mathrm{X_i} \neq \mathrm{Y_j}$). Cela conduit à une réaction globale

$$\sum_k \bar{n}_k^{(+)}\mathrm{X_k} + \sum_l n_l^{(+)}\mathrm{Y_l} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k \bar{n}_k^{(-)}\mathrm{X_k} + \sum_l n_l^{(-)}\mathrm{Y_l}, \tag{1}$$

Contraint par

$$\sum_k n_k^{(+)}\mathrm{Y_k} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k n_k^{(-)}\mathrm{Y_k}, \quad n_j^{(+)} \neq n_j^{(-)}, \tag{2}$$

$$\sum_k \bar{n}_k^{(+)}\mathrm{X_k} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k \bar{n}_k^{(-)}\mathrm{X_k}, \quad \bar{n}_j^{(+)} = \bar{n}_j^{(-)}. \tag{3}$$

Les réactions utilisées dans *pmbc** respectent une loi de conservation de masse L^* pour la population d'allocatalyseurs :

$$L^* = \sum_k a_k \mathrm{X_k}, \ \ \forall k\ a_k \geq 1. \tag{4}$$

Où, par la conservation de masse, nous choisissons d'exclure les compositions de réactions autocatalytiques comme forme d'allocatalyse.

Nous pouvons les regrouper dans une caractérisation d'une «sous-matrice allocatalytique» : Une sous-matrice $\boldsymbol{\nu}^*$ qui est i) autonome, ii) admet un cycle émergent $\boldsymbol{c}^*$ impliquant toutes les espèces et réactions dans $\boldsymbol{\nu}^*$, et iii) admet une loi de conservation de masse L^* contenant tous les éléments internes est une sous-matrice allocatalytique. $\boldsymbol{c}^*$ est un cycle allocatalytique, et toutes les espèces $x_{\boldsymbol{c}^*}$ sont des catalyseurs.

Un bilan de réaction pour une réaction autocatalytique peut contenir des réactifs, des déchets et des allocatalyseurs supplémentaires (par exemple des cofacteurs). Une caractéristique distinctive est que les autocatalyseurs, par définition, peuvent augmenter en nombre. Il existe donc un combinaison de réactions qui conduit à la production nette de tous les autocatalyseurs (vecteur de droite de la matrice qui donne une vecteur strictement positive). Nous appelons la reproduction stoechiométriquement réalisable (SFR), la propriété qu'une sous-matrice est i) autonome, ii) monobloc, iii) il existe un combinaison de réactions qui produit toutes les espèces décrit par la matrice.

Une propriété plus forte, une autocatalyse stœchiométriquement réalisable (SFA), se produit lorsqu'en plus du SFR, la production nette de chaque espèce implique également la consommation de chaque espèce. Pour SFR, il peut y avoir d'espèces qui ne contribuent pas à l'autocatalyse. SFA nécessite la participation (par consommation) de chaque espèce. À la fin du chapitre, nous démontrons qu'un réseau SFR qui n'est pas SFA peut toujours être réduit davantage à un réseau SFR plus petit. Le plus petit SFR doit au moins être carré de taille 2 et doit également être SFA. Un SFA avec des lois de conservation mixtes admet toujours un sous-réseau SFR plus petit. Comme les cycles peuvent être supprimés librement et que toutes les lois de conservation disparaissent lors de la réduction du réseau, il résulte du théorème de la nullité de rang que les matrices SFA irréductibles sont inversibles. Cependant, être inversible et SFA n'implique pas d'irréductibilité.

Cette approche décrit l'autocatalyse et l'allocatalyse comme on le trouve dans la littérature chimique dans un seul cadre théorique. En approfondissant l'analogie entre les processus d'échange (évaporation, diffusion, etc.) et les réactions, les réseaux autocatalytiques devraient également émerger dans un contexte multicompartimental. Un exemple simple est fourni par un réseau qui envoie une molécule B à la phase II, pour aller chercher une espèce BC, formant BCB. À son retour dans la phase I, BC est à nouveau libéré et converti en un autre B, ce qui donne deux B pour récupérer plus de BC.

Cette chimie ne devient autocatalytique qu'en présence de plusieurs compartiments, dans lequel BC peut être abondant dans l'un (comme nourriture / chimiostat) et rare dans l'autre compartiment (en assumant le rôle d'autocatalyseur). Ces rôles s'excluent mutuellement dans un seul compartiment. Plus généralement, de nouveaux cycles autocatalytiques sont à prévoir du fait du couplage entre compartiments chimiquement distincts, car ils admettent le couplage de réactions chimiques avec des exigences incompatibles (pH, potentiel redox, état, etc.). Ces conditions sont courantes en écologie, par ex. pour les organismes pratiquant l'alimentation croisée (syntrophie).

Le nombre de motifs autocatalytiques est strictement supérieur lorsqu'un système est couplé à l'environnement. Que cela se manifeste d'avantage est dépendant du contexte.'

Nous pouvons faire la distinction entre l'autocatalyse solitaire, impliquant des réactions directes qui ne sont que de premier ordre en termes d'autocatalyseur, et l'autocatalyse jointe, où les autocatalyseurs doivent rencontrer d'autres autocatalyseurs. Les deux sont communs, mais ils diffèrent sensiblement dans leur comportement. Les autocatalyseurs solitaires peuvent se multiplier et atteindre des nombres macroscopiques, à partir d'une seule espèce. Les autocatalyseurs jointes doivent surmonter une concentration seuil telle que les rencontres d'autocatalyseurs soient suffisamment courantes, ce qui peut conduire à des bistabilités, comme en témoigne la réaction termoléculaire de Schlögl. Thermodynamiquement, l'autocatalyse solitaire devrait être favorisée dans un premier temps, en raison de sa contribution favorable à l'entropie de mélange.

Dans le chapitre 6, le concept d'évolution chimique (parfois appelé évolution «pré-darwinienne») est discuté, un processus hypothétique qui a servi pendant plus d'un siècle de deus ex machina de scénarios d'origines de vie pour expliquer comment les molécules simples progressent progressivement. est devenu plus complexe. Nous passons en revue certains aspects des ensembles RAF, GARD, et un modèle pour un métabolisme évolutif, qui reposent tous sur l'autocatalyse. Ensuite, nous dérivons lorsque les perturbations d'une seule molécule entraînent des changements durables dans un réacteur, ce qui donne des réseaux autocatalytiques de premier ordre. En utilisant la théorie des processus de branchement, nous dérivons des expressions analytiques pour la probabilité qu'un nouvel autocatalyseur forme avec succès une grande population stable, la «fixation». Nous examinons ensuite les aspects de cette forme d '«évolution» autocatalytique et avançons les obstacles les plus courants rencontrés dans l'évolution chimique et ses modèles et esquissons une évolution vers une synthèse de l'évolution chimique. Une première extension du corpus d'idées en évolution chimique est l'autocatalyse multicompartimentale, pour laquelle nous étudions la probabilité de fixation dans une variété d'environnements et de contextes. Nous montrons notamment que des

mécanismes bien connus de coopération en écologie peuvent être en jeu.

L'évolution chimique a été définie, redéfinie et laissée indéfinie pendant plus d'un siècle. C. Malaterre distingue deux traditions intellectuelles : une approche descriptive «historique» et des tentatives théoriques pour définir son fonctionnement. Un tel concept est considéré comme nécessaire par certains, car l'émergence de mécanismes complexes de la vie (comme le génome) peut alors s'expliquer par un processus évolutif. En pratique, ceci est traité par des modèles avec une certaine chimie autocatalytique.

Les ensembles autocatalytiques (ou ensembles RAF) décrivent des types particuliers de systèmes autocatalytiques, dans différentes espèces de catalyseurs qui se catalysent mutuellement la formation. La théorie fait une distinction entre la catalyse et les réactions, et un ensemble RAF nécessite que chaque réaction soit catalysée. On constate aisément que cela implique un niveau de détail particulier pour la description : on ne peut pas décrire les étapes de réaction en catalyse, car une telle description de mécanisme introduirait des réactions non catalysées. Une prédiction de scénario faite dans le contexte de la théorie de la RAF, est que lorsqu'un mélange chimique devient suffisamment diversifié (par exemple, 10^6 sortes de molécules alimentaires différents), il doit y avoir des motifs autocatalytiques. Un tel situation est généralement considéré en termes de copolymères.

Le modèle GARD prend en compte les collectifs d'espèces réciproquement allocatalytiques. En règle générale, le terme GARD est utilisé comme un toto pro pars pour désigner l'amphiphile GARD, dans lequel les amphiphiles catalysent leur incorporation mutuelle dans une micelle ou une vésicule. Si l'autocatalyse ainsi induite est suffisamment forte et spécifique, et que le nombre de types d'amphiphiles est important, il peut y avoir un nombre considérable d'attracteurs distincts pour la composition du système (composome). Dans le même temps, une augmentation du nombre d'espèces entraîne un bruit intrinsèque plus élevé lors de l'incorporation. Une diversité d'amphiphiles trop importante peut alors empêcher la transmission stable d'un composome à travers plusieurs cycles de croissance et de division. Une idée fausse persistante au sujet des composomes est que leur contenu d'information est caractérisé par le nombre de compositions moléculaires possibles. La nature fluctuante de la composition rend cette mesure thermodynamique problématique. En utilisant la théorie du chapitre 4, il est montré que le nombre d'attracteurs peut avoir une interprétation thermodynamique de l'information rigoureux.

Une autre idée, lancée par King, était que la chimie autocatalytique conduit à plus de chimie autocatalytique, et a discuté de la notion d'une fidélité de réaction seuil pour l'autocatalyse pour survivre, aujourd'hui connu comme le «seuil de décroissance». Un travail ultérieur de Bagley, Farmer et Fontana a modélisé le déclenchement de réactions autocatalytiques, qui à leur tour modifieraient le mélange de réactifs et conduiraient à de nouveaux motifs autocatalytiques à déclencher, décorant ainsi de manière autocatalytique un " métabolisme " par évolution autocatalytique. Ce que toutes ces idées d'évolution chimique ont en commun, c'est l'autocatalyse.

Nous proposons l'expérience de pensée suivante : Supposons un grand (disons, $N = O(10^{23})$ CSTR dans un état d'équilibre hors d'équilibre. La composition est légèrement perturbée par l'arrivée d'une seule espèce, et nous nous demandons si cela peut avoir des conséquences macroscopiques. Nous constatons que seule l'autocatalyse peut entraîner des changements macroscopiques dans la composition. Nous en déduisons ensuite une description en termes d'espèces rares et abondantes. Le nouvel autocatalyseur est rare, tout comme les autres dans son cycle autocatalytique, ce qui signifie que leurs rencontres mutuelles peuvent (initialement) être négligées. Il s'ensuit que les réactions directes dans le cycle autocatalytique doivent toutes être de premier ordre dans un autocatalyseur, et qu'il existe au moins une étape de fragmentation, qui est nécessairement irréversible.

La probabilité de fixation est maintenant obtenu à partir des probabilités de terminer avec succès les cycles, qui pour un seul chemin cyclique peuvent être refondus dans un processus de

naissance et de mort, avec une probabilité de naissance correspondant à la probabilité de terminer avec succès le cycle autocatalytique. Lorsque plusieurs chemins cycliques sont impliqués, nous utilisons des probabilités de chemin pour construire les «statistiques de naissance» pour un seul autocatalyseur : le nombre de copies de lui-même qu'il générera efficacement. Ensuite, trouver la probabilité de fixation est trouvé en mappant le problème sur un processus de branchement. Cela permet d'étudier l'effet de la structure du réseau sur la survie. Par exemple, une branche interne, représentant un cycle allocatalytique, peut fortement stabiliser la survie. Puisque la survie diminue avec l'intensité des réactions secondaires, qui sont irréversibles, la sélection autocatalytique a une tendance à choisir contre une dissipation 'inutile'.

Le modèle d'évolution autocatalytique à réacteur unique est fortement analogue à l'évolution du métabolisme de Farmer et al. Cependant, ces processus ne sont pas considérée comme une évolution chimique suffisamment convaincant pour l'abiogenèse. Nous considérons quatre principaux obstacles rencontrés dans les théories de l'évolution chimique : I. Variation et rareté de l'autocatalyse : il n'est pas clair qu'il y aura suffisamment d'autocatalyse dans un réacteur unique pour avoir des trajectoires d'évolution contingentes.

II. Réactions secondaires, états de piégeage, complexes inactifs : la solution populaire consistant à mettre de nombreux composants différents dans la même solution est un moyen évident d'obtenir plus de motifs autocatalytiques. Orgel et Szathmàry ont soutenu que c'est une recette pour un désastre : cela conduit à une augmentation spectaculaire des réactions secondaires, ce qui empêcherait les motifs de se développer réellement, ce qui découle également de notre calcul de la probabilité de fixation. Les exigences contradictoires d'existence (grand nombre d'espèces) et de survie (petit nombre d'espèces) ont mené au concept de la «paradoxe de la spécificité». III. Graphes de réseaux chimiques et modélisation : de nombreuses approches reposent sur une chimie artificielle formulés en termes de graphes aléatoires. En pratique, la chimie est loin d'aléatoire, elle est très structurée : les acides réagissent avec les bases, les agents oxydants réagissent avec les agents réducteurs, etc. Souvent, une réaction donnera des espèces moins réactives et plus faible énergétiquement, comme les acides et les bases conjugués faibles. IV. Restrictions de modèle et restrictions de scénario : les scénarios prébiotiques ont tendance à se concentrer sur un seul type de molécule et un seul type de mécanisme (par exemple GARD). Cependant, la plupart des molécules et des mécanismes n'impliquent pas que d'autres molécules et mécanismes soient exclus. La levée des restrictions inutiles est une étape nécessaire vers une synthèse de l'évolution chimique, qui considère tous les mécanismes qui peuvent se produire.

Le problème de fixation pour l'autocatalyse à plusieurs compartiments dépend fortement de la géométrie. Si un composé doit diffuser dans les deux sens entre un petit compartiment et un grand milieu environnant, il peut «se dégrader efficacement», simplement parce que les trajectoires de diffusion ne reviennent pas toujours à l'origine (en particulier en 3D), ou le font de manière trop lentement (comparé à d'autres processus de dégradation).

Pour augmenter la probabilité de fixation, nous considérons : i) le confinement spatial : les espèces restent déclenchées dans un volume fini, ce qui rend le retour considérablement plus probable ii) la coopération : lorsque plusieurs compartiments sont présents qui effectuent la même réaction, un autocatalyseur n'a pas besoin de revenir à son compartiment d'origine : il lui suffit d'atteindre l'un d'entre eux. iii) Structure du réseau : certains réseaux autocatalytiques sont plus résistants à la dégradation, tels que les réseaux qui effectuent une allocatalyse dans un compartiment, pour générer de nouvelles molécules qui récupèrent plus d'allocatalyseurs. Toutes ces solutions ont un équivalent en écologie microbienne, comme dans les biofilms pour confiner le matériel ou les communautés microbiennes qui libèrent des sidérophores pour trouver Fe^{3+}.

Dans le chapitre 7, nous considérons un certain nombre de mécanismes hors équilibre qui génèrent de longs polymères, tels que la recombinaison couplée à l'échange de réservoir, l'adsorption et la chimie de la ligature chimiquement dirigée et discutons de l'heuristique et des arguments

physiques pour comprendre certaines des distributions de longueur de polymère typiques que nous pouvons rencontrer dans la littérature. Nous déduisons des limites thermodynamiques pour des scénarios prébiotiques basés sur des copolymères et considérons les conséquences de ces restrictions. Notre discussion est largement tirée de deux publications publiées, une publication en préparation, une mémoire de M2 et des travaux en cours.

Les réactions de recombinaison sont intéressantes en OOL : elles fournissent un moyen dynamique d'explorer de nombreuses séquences de copolymères différentes et de générer de longues espèces, sans nécessiter d'activation. Nous étudions ces réactions en utilisant la thermodynamique hors équilibre et la thermodynamique stochastique. En suivant les espèces au niveau de la séquence, le taux de production d'entropie est divisé en une contribution pour l'énergie libre standard, le désordre dans la distribution de tailles et le désordre dans les distributions de séquences à l'intérieur de chaque taille donnée. En principe, ces contributions peuvent être couplées, ce qui permettrait un contrôle thermodynamique sur la séquence et/ou la longueur. Un tel couplage apparait notamment due à des effets de taille fini. La distribution de la longueur à l'équilibre est exponentielle, également lorsque des interactions plus proches voisins sont introduites. Enfin, nous dérivons des temps de relaxation pour la séquence et la longueur, constatant que la relaxation de séquence est soit aussi rapide que la relaxation de longueur, soit considérablement plus lente, selon le mécanisme de recombinaison.

En utilisant le formalisme de la matrice stoechiométrique, nous étudions les réactions de recombinaison et de ligature dans un système ouvert, qui échange certaines séquences de copolymères avec un environnement. Via l'introduction d'une notation de séquence, cela permet l'étude topologique de réseaux de polymères de dimension infinie, à partir desquels nous dérivons des lois de conservation et des cycles dépendants de la séquence. Celles-ci changent considérablement en fonction du mécanisme de réaction. Pour certaines concentrations dans les réservoirs, un état de croissance perpétuelle de polymère induit par recombinaison est atteint. Dans cet état, de gros oligomères pénètrent dans le système à partir d'un réservoir et se recombinent pour former un polymère plus gros et un petit oligomère, dont ce dernier se déplace à nouveau vers un réservoir. Lorsqu'une taille maximale est imposée et en l'absence de paysage énergétique, cela donne une distribution exponentielle croissante.

Les minéraux forment une partie considérable de la recherche OOL, et pour les ARN, il a été démontré que les minéraux peuvent protéger contre la dégradation et favoriser la synthèse des monomères et la ligation de l'ARN activé. Dans une étude conjointe avec les laboratoires de D. Baum et N. Lehman, on constate expérimentalement qu'à partir de populations d'ARN en solution, c'est préférentiellement de l'ARN de grande taille qui s'adsorbe sur une surface minérale. On obtient un modèle minimal pour expliquer cette observation, en commençant par étudier d'abord l'adsorption par multiples sites en 1D, puis en l'étendant ensuite en 2D. Cette extension est largement capturée par un changement de paramètres pour le modèle 1D. On constate que les oligomères plus petits sont favorisés par une forme d'entropie de mélange et les oligomères plus grands par l'énergie libre d'adsorption. Pour obtenir une adsorption appréciable, cette dernière contribution ne peut pas être trop faible. Si l'adsorption est importante, nous nous attendons donc à un enrichissement exponentiel d'oligomères plus gros.

De nombreux scénarios pour les origines de la vie partagent une étape clé : la ligature des premiers monomères par des moyens abiotiques. Nous revisitons une grande classe de ces scénarios, impliquant l'utilisation de modèles de Ligature-Fragmentation, en incluant explicitement l'étape d'activation chimique qui est normalement laissée implicite. La distribution de tailles de polymères à en état stationnaire de ce modèle de Activation -Ligature-Fragmentation (ALF) peut être exprimée en termes de deux quantités sans dimension : le rapport de ligature et le rapport d'activation, qui quantifient les taux relatifs d'activation et de ligature par rapport à l'hydrolyse. Dans les régimes limitatifs, une seule de ces quantités est caractéristique. En raison de l'absence d'équilibre, la

catalyse peut modifier la distribution de taille à l'état stationnaire. Cependant, cela doit concerner une étape de limitation de débit, par ex. lorsque l'activation est l'étape la plus lente, l'accélération de la ligature par l'ajout d'une ligature assistée par matrice n'entraîne pas une nouvelle augmentation de la longueur moyenne du polymère. D'un autre côté, le pliage et l'hybridation deviennent plus efficaces pour augmenter la taille des polymères dans ce régime.

L'inclusion de l'étape d'activation permet également d'utiliser la thermodynamique hors équilibre pour faire des déclarations générales, indépendantes du modèle, par ex. en ce qui concerne la dissipation. À partir de la dissipation pour effectuer un cycle ALF en état stationnaire, nous trouvons un coût thermodynamique minimale associé à l'exploration de séquences, indépendamment de la catalyse. Ce coût est une propriété de la chimie d'activation et est absent pour les schémas de réaction non dissipatifs qui effectuent une telle exploration, tels que les réactions de recombinaison. Nous illustrons comment ce coût fournit des limites quantitatives sur une grande famille de scénarios prébiotiques qui impliquent la génération dissipative de séquences aléatoires jusqu'à ce qu'une séquence rare soit trouvée.

Il existe un compromis entre l'énergie, le temps de recherche et la complexité des structures recherchées. En utilisant les distributions connues des structures secondaires et des limites supérieures pour les flux d'énergie planétaire, nous constatons que la thermodynamique empêche cette recherche aléatoire dissipative - souvent invoquée - de trouver les structures très complexes qu'elle est supposé trouver. Cependant, d'autres stratégies de recherche plus efficaces sont actuellement envisagées, telles que la réplication sans enzyme.

De nombreuses distributions de taille se retrouvent souvent dans la littérature, lorsque l'on considère les polymères et leur formation dans diverses circonstances. Pour obtenir une intuition pour les distributions de longueur en équilibre et hors équilibre, nous considérons quelques équations cinétiques typiques pour diverses situations invoquées en OOL et leurs solutions approximatives. Sans décoration (par exemple paysage énergétique), une seule exponentielle est la distribution de longueur d'équilibre. Hors équilibre, une deuxième échelle de temps (dégradation modifiée due au pliage, ligature assistée par modèle) pour les espèces plus longues conduit approximativement à une distribution à deux exponentielles. L'argument peut être répété pour obtenir des distributions multi-exponentielles. Des lois de puissance ont été obtenues, lorsque des processus sont ajoutés qui sont proportionnel au longueur de manière différente que les processus déjà présents. Par exemple, dans un CSTR, une dégradation uniforme se produit en raison de l'écoulement, tandis que la dégradation hydrolytique est proportionnelle à la longueur. Si la dégradation uniforme domine, la distribution atteint asymptotiquement une loi de puissance.

En évaluant une grande variété de situations, nous montrons qu'il existe de nombreuses voies vers des polymères longs, en exploitant diverses stratégies hors équilibre. Cependant, d'après l'analyse thermodynamique de l'exploration des séquences, nous constatons que l'image «longs polymères = polymérase = vie» n'est pas satisfaisante. En conséquence, nous pouvons souhaiter reconsidérer le rôle des polymères dans les scénarios OOL, et quelles exigences nous invoquons sur leur longueur et leur séquence. Une considération instructive est qu'avant les gènes, les polymères et les oligomères auraient déjà pu être utiles, mais avec des exigences moins strictes sur la longueur et la séquence. Par exemple, certaines familles de séquences tripeptidiques peuvent s'auto-organiser avec le fer pour former des catalyseurs fonctionnels de ferrédoxine. Par condensation de polyélectrolytes, les polymères peuvent former des coacervats, des compartiments liquides sans membrane avec des propriétés très distinctes de l'eau. Nous soutenons que des scénarios plus progressifs devraient être formulés, en considérant les fonctions initiales pour les polymères, qui apparaissent plus facilement, avant d'insister sur les fonctions nécessitant des séquences longues spécifiques.

Dans chapitre 8, nous étudions la dynamique de compartimentation transitoire pour les populations moléculaires. Dans la littérature OOL, les compartiments sont maintenant principalement

considérés comme intéressants pour deux raisons : i) un potentiel d'accumulation locale de produits chimiques prébiotiquement pertinents et ii) une sélection à plusieurs niveaux. Notre discussion se concentrera sur ce dernier, avec une synthèse basée sur deux articles publiés, suivie de travaux en cours (à savoir une sous-section sur la coopération induite par le bruit et une section sur la catastrophe de complexation).

Les destins des molécules compartimentés peuvent être alignés par des mécanismes de sélection agissant sur les compartiments et son contenu. Une population moléculaire qui favorise la survie de son compartiment peut ainsi devenir plus abondante. Cela peut se produire grâce à des mécanismes de coopération entre les molécules, favorisant des compositions qui ne survivraient en solution, par un processus de sélection sur multiples échelles.

Un mode similaire de sélection (group selection) se trouve dans la division cellulaire, comme illustré par le modèle de correction stochastique. Ces mécanismes ont été considérés comme un moyen important de sauver les populations de réplicases des parasites formés par les erreurs de réplication, et donc la maintenance des informations. Bien que ces réplicases appartiennent à des branches de scénarios très spécifiques, ce mode de sélection de groupes est un mécanisme bien plus général. Un mécanisme très similaire (mais utilisant l'incorporation catalytique des amphiphiles) agit dans le modèle GARD.

Un autre mécanisme récemment proposé et testé expérimentalement est la compartimentation transitoire, dans laquelle les (sous) populations sont encapsulées, cultivées, sélectionnées et libérées, et le cycle peut être répété. Car le compartiment n'est plus divisé en deux mais détruit après un cycle, les molécules enfermées ne sont plus contraintes à leur multiplication par un cycle de réplication. Ils peuvent ainsi se multiplier par des facteurs bien supérieurs à 2 (dans les expériences : 10^6).

Nous développons un formalisme pour la compartimentation transitoire qui prend en compte les statistiques de compartimentation et de croissance, et l'illustrons pour des espèces concurrentes à croissance indépendante. Les réplicateurs fonctionnels (dans l'expérience : ribozymes) peuvent être stabilisés par la sélection et un diagramme de phase pour le cas de ribozyme-parasite est obtenu en fonction de la taille de l'inoculum et de la croissance relative. En modifiant la fonction fitness 1D, nous pouvons traiter le cas de la coopération. En ajoutant des mutations déterministes au modèle, la dimension du modèle est augmentée d'une unité et il est démontré que les compartiments transitoires peuvent surmonter les catastrophes d'erreur.

En passant d'un petit nombre de réplicateurs à une grande population, le bruit dans le taux de réplication est amplifié de façon exponentielle, ce qui peut entraîner des fluctuations géantes dans la composition finale de la population. En utilisant la théorie des processus de branchement, nous dérivons le bruit dans la composition de la population pour les réplicateurs concurrents en fonction des taux de réplication et des longueurs de polymère. Étant donné la sélection rigoureuse de la composition que nous pouvons imposer dans la compartimentation transitoire, des gains de fitness considérables peuvent être réalisés en réduisant le bruit de composition. Les réseaux autocatalytiques à petites molécules, qui comportent généralement une étape ou peu d'étapes de limitation de débit, subiraient de grandes fluctuations. Les polymères, en revanche, peuvent réduire fortement leur bruit de composition : une incorporation successive de monomères (comme dans la réplication de modèle) peut conduire à une distribution du temps d'attente très étroite. En utilisant la théorie des processus de branchement, nous montrons que cela conduit à une réduction du bruit de composition. Comme la sélection agit sur la composition, le mode de réplication peut avoir de forts effets de fitness en raison de la réduction du bruit. Cela pourrait favoriser l'apparition de polymérases.

Lorsqu'une réplicase peut copier des réplicases, mais aussi des parasites, les choses deviennent plus délicates. Si la polymérisation limite la vitesse, la formation d'un complexe réplicase-parasite empêche la réplicase impliquée de copier d'autres réplicases. La théorie des processus de branche-

ment ne peut plus être utilisée ici : les populations interagissent et ne peuvent donc pas être traitées comme indépendantes dans le théorème de renouvellement. Cependant, en regardant les simulations et les trajectoires de réplication possibles, il devient clair que le traitement typique avec des modèles d'action de masse à 2 espèces devient très inapproprié : de tels modèles ne tiennent pas compte du fait que les molécules sont séquestrées dans des complexes. Dans un traitement plus détaillé, les parasites deviennent considérablement plus dangereux : ils se copient intrinsèquement plus rapidement (même s'ils ont la même longueur que la réplicase) et s'ils s'accumulent, ils peuvent rapidement occuper toutes les réplicases disponibles. Dans une telle situation, une réplicase fraîchement libérée rencontrera rapidement un autre parasite pour former un complexe avec (la réplication est à limitation de vitesse, donc la complexation est relativement rapide), et rencontrer une autre réplicase libre devient un événement de plus en plus rare. Nous appelons une telle prise de contrôle une «catastrophe de complexation». La prise en compte d'une telle catastrophe introduit de nouvelles contraintes importantes sur l'émergence et l'évolution des polymérases.

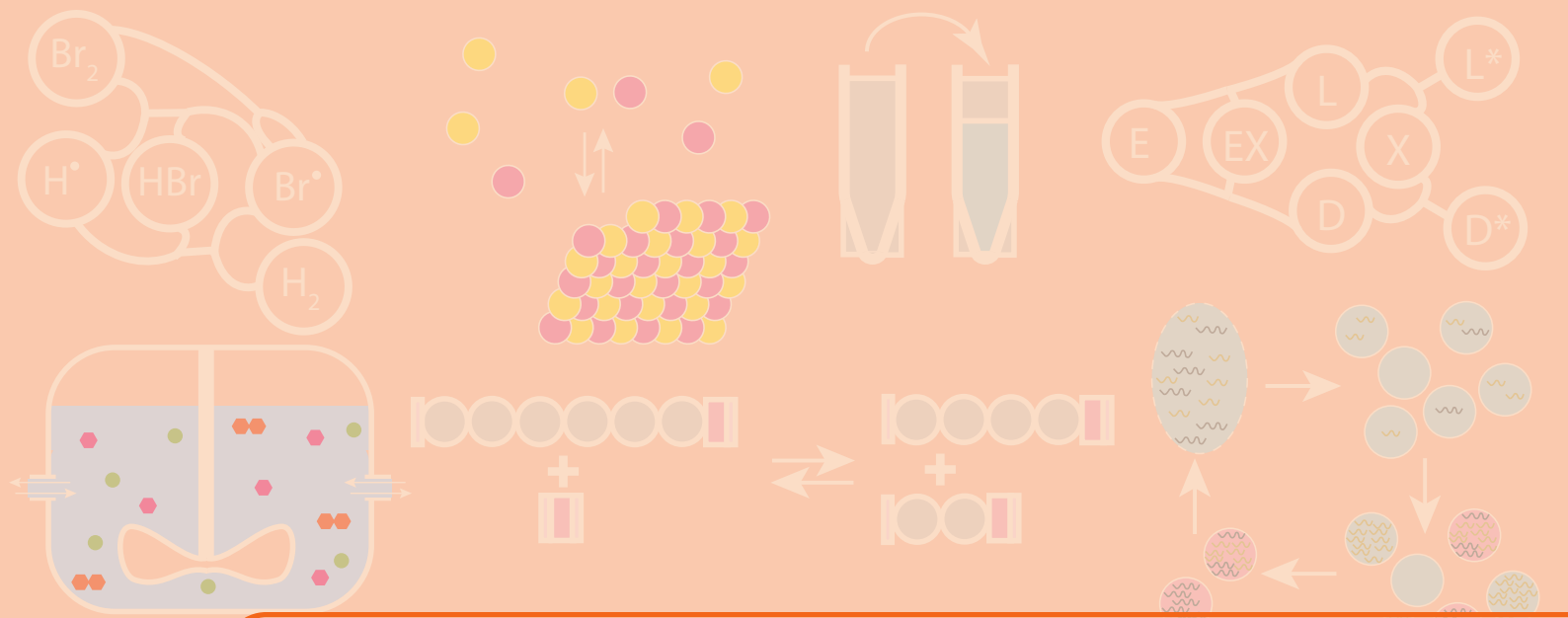
Dans Chapitre 9, un nouveau quasi-scénario est formulé, basé sur les résultats des chapitres précédents. Le scenario a pour but d'être provocateur, et souligne qu'il y a encore beaucoup de place pour de nouvelles idées et scénarios en OOL.

Il a été suggéré, le plus récemment par Krishnamurthy, que l'abiogenèse pourrait être un processus pour lequel notre biochimie n'est qu'un résultat particulier parmi d'autres. Notre scénario est adapté à cette philosophie et spécule sur les mécanismes qui entrent en jeu dans un tel processus.

Le scénario considère des réseaux chimiques hors d'équilibre à multiples compartiments, avec des barrières de transport. L'évolution autocatalytique déclenche une modification permanente dans les propriétés de transport et la chimie, modifiant l'auto-tri (self-sorting) et conduisant à de nouvelles innovations au niveau du réseau (comme la résolution cinétique dynamique ou diverses formes de relecture). De plus, de nouvelles phases (compartiments), interfaces et agrégats se forment et renforcent la complexification.

Par construction, le scénario devient de plus en plus tiré par les cheveux : à mesure que nous nous déplaçons au-delà des molécules uniques vers des structures d'ordre supérieur, un nombre croissant de nouveaux phénomènes émergents peuvent apparaître. Traiter ces étapes avancées avec rigueur nécessite de nouveaux cadres et connaissances, qui devraient probablement provenir de la chimie des systèmes dans le futur proche.

Le point principal reste inchangé : nous soutenons qu'une variété de mécanismes autocatalytiques, parmi lesquels l'autocatalyse à plusieurs compartiments, a fourni une évolution chimique à caractère écologique, qui a commencé à former de nouveaux compartiments et à modifier l'environnement. À son tour, cela a favorisé la sélection à plusieurs niveaux des populations moléculaires, des compartiments avec un contenu particulier (par exemple des réseaux de correction d'erreurs) et des collections d'ordre supérieur de ceux-ci. Ces compartiments dépendaient de voies catalytiques à plusieurs compartiments pour leur formation et des structures d'ordre supérieur ont émergé pour favoriser ces voies. Ainsi, de nouvelles couches de sélection ont été introduites et de nouvelles pressions de sélection ont commencé à agir. Cela devait de plus en plus être résolu par des chimies flexibles et évolutives, et a abouti à une course aux armements évolutive vers l'évolutivité, produisant des systèmes de plus en plus semblable du vivant.



Contents

1. Introduction to the Origins of Life

τούτου μὲν τοῦ ἀνθρώπου ἐγὼ σοφώτερός εἰμι· κινδυνεύει μὲν γὰρ ἡμῶν οὐδέτερος οὐδὲν καλὸν κἀγαθὸν εἰδέναι, ἀλλ᾽ οὗτος μὲν οἴεταί τι εἰδέναι οὐκ εἰδώς, ἐγὼ δέ, ὥσπερ οὖν οὐκ οἶδα, οὐδὲ οἴομαι· ἔοικα γοῦν τούτου γε σμικρῷ τινι αὐτῷ τούτῳ σοφώτερος εἶναι, ὅτι ἃ μὴ οἶδα οὐδὲ οἴομαι εἰδέναι.

— Σωκράτης

I am wiser than this man, for neither of us appears to know anything great and good; but he fancies he knows something, although he knows nothing; whereas I, as I do not know anything, so I do not fancy I do. In this trifling particular, then, I appear to be wiser than he, because I do not fancy I know what I do not know.

— Sokrates

This chapter is an informal introduction to the science that concerns itself with scenarios for the origins of life (OOL). The need for such an introduction comes from the rich heterogeneity of disciplines and perspectives in the field, too numerous to be covered here. In doing so, we provide some of the baggage to justify the work in following chapters. We hope to provide the reader with an appreciation and some context for the questions that are addressed in different branches of OOL. We also hope to provide handles in the critical interpretation of the literature and its corresponding press statements.

Specifically, we will discuss the folowing things:

i) *What we know absolutely and what we often assume*: One can learn and practice a branch of science because there are general principles and truths that pertain to that discipline. These thruths and principles provide an important sanity check in the elaboration of scenario and their mechanisms.

To make progress, scenarios need to make simplifying assumptions and disregard some details. Often this happens implicitly, due to thinking by analogy, and we only identify assumptions when we see how things can be different. Some typical assumptions in popular prebiotic scenarios will be identified and discussed.

ii) *Popular perspectives and scenarios in origins of life and how they shifted*: Origins of life (OOL) addresses a plethora of questions with some common themes. These questions bear the mark of history: scientific paradigms have shifted dramatically over the years, often due to milestone discoveries. Such developments may well repeat themselves in the foreseeable future.

iii) *Communities and multidisciplinarity*: OOL is composed of numerous large subcommunities that do not have full knowledge of each other's science and practices. These subcommunities have very different compositions of scientific disciplines, which is reflected in different assumptions and scenarios, but also methods and practices. Origins of Life Initiatives are now actively mapping out these differences through bibliometric studies, and organize conferences to bridge disciplinary gaps, to come to a stronger collective effort.

In Sec. 1.1, constraints on prebiotic scenarios are discussed. Constraints follow from 'established facts' (e.g. 2nd law of thermodynamics) that, to the best of our knowledge, are robust. In Sec. 1.2, common assumptions in the literature are discussed, in the context of their corresponding scenarios. Contrary to constraints, assumptions are not established facts and many of the assumptions made in OOL will, by necessity, conflict with each other.

Most of the current prebiotic scenarios and research questions are a product of major discoveries, leading to paradigm shifts and fragmentation. This is further discussed in Sec. 1.3. In Sec. 1.4, we will discuss some of the research that quantifies and addresses the multidisciplinary issues in the field. Finally, in Sec. 1.5, we lay out what the contribution is of the manuscript in the big picture of OOL research and summarize chapter by chapter what will be discussed.

1.1 Constraints on the Origins of Life

Research in the field of Origins of Life (OOL) assumes that life came forth from abiotic matter. What is largely unknown is where this happened, which abiotic matter was involved, what physical chemical mechanisms came into play and through which different stages prelife proceeded. While the exact system of interest is presently unknown, it is commonly considered that one starts with a soup of simple chemicals. What these chemicals subsequently do must initially lack the sophistication of biochemistry, whose emergence it tries to explain. Phenomena of prebiotic interest are thought to emerge from an interplay between chemistry, physical chemistry, and geology. This places us in a privileged position to study history: established facts from different fields of natural sciences strongly constrain the narrative.

Such constraints may not fully allow us to establish what happened, but they allow to establish what did not happen. Scenarios that break mass conservation, violate the second law, invoke nonexistent chemistry or rely on exploiting ultra-rare elements must be amended until such problems are addressed. Many results in origins of life come in the form of positve results, sometimes accompanied by new scenarios. This allows us to accumulate plenty of incriminating evidence pointing towards potential scenarios responsible for life. For a focused investigation, we also need to be able to narrow down the number of suspects: we wish to provide an alibi for the majority of alternative scenarios and mechanisms.

In the following, important constraints imposed by different fields are discussed. The list is by no means exhaustive and the distinction by field is somewhat arbitrary. Nevertheless, it gives an insight into arguments brought forth by corresponding scientific communities, and ones that are sometimes overlooked by other ones.

1.1.1 Chemical Constraints

- Conservation of atoms and isotopes: A chemical reaction conserves all atoms and isotopes involved in its process. It is a particular form of local mass conservation*.
- Concentrations govern reaction times: An elementary bimolecular reaction has a rate proportional to local reactant concentrations, following from the probability of random encounter of molecules. Concentrations can reach from moles to single molecules per unit volume, creating strong separations of relevant timescales. The rate of a reaction is bounded by the rate at which reactants can meet, which microscopically can be identified with diffusion rates.
- Chemical reaction steps are, a priori, unimolecular or bimolecular: 'Ternary reactions' can always be considered as a succession of fast and slow bimolecular reactions, sometimes proceeding through activated, noncovalent complexes.
- Stability and degradation: Molecules of interest are degraded on their particular timescale by an environment, e.g. due to hydrolysis, thermolysis and radiation. An important example is DNA, which is stable on a human timescale, but prone to enzymatic degradation on longer timescales. Willerslev et al observe a practical limit[1] for recovering bacterial DNA between 400ka and 1.5Ma.
- Reactivity and reaction networks: A considerable part of chemistry is dedicated to understanding which species react with each other, and why. From this, some clear patterns emerge: acids react with bases, yielding weaker conjugate acids and conjugate bases. Oxidizing agents react with reducing agents to yield weaker oxidizing and reducing agents. Nucleophiles react on electron poor sites, electrophiles react on electron-rich sites. Sterically hindered (bulky, impenetrable) reaction sites react slower than their more readily accessible counterpart. A much-used framework that aims to capture reactivity trends in simple rules is HSAB (hard-soft acid-base) theory. Chemical reaction rules have also been formalized in computer languages, to generate extensive reaction networks upon repeated application of known reaction pathways. Large chemical networks of real reactions are highly constrained by what actual chemistry does. Approximating such a network as a random graph typically leads to absurd chemistry, only very particular chemical systems seem well-suited for such an approach.

1.1.2 Physical Constraints

- Conservation laws: Mass-energy, charge, linear momentum, angular momentum and probability are some quantities that are often considered as conserved in the universe. For most systems, a number of approximate conserved quantities emerge. We often can treat mass conservation and energy conservation as separate conservation laws. The network structure in chemical networks conserves certain combinations of reactants. Such conservation laws place absolute bounds on what a prebiotic scenario can do: it can expend the energy there is, it can mobilize the matter that is provided, but not more.
- Thermodynamics: A physically realizable process must obey the laws of thermodynamics, as well as the extensions to nonequilibrium thermodynamics.† By being in accord with these laws, we get equilibrium constants and constraints on kinetic rate constants (detailed balance). A reaction vessel that relaxes to equilibrium reaches a unique macroscopic steady-state and subsequently produces no more entropy. A reaction vessel that stays out of equilibrium (e.g. by a replenishing influx of resources for the further production of entropy) may reach various nonequilibrium steady states and produce a host of other behaviors, which must again obey

*Implicitly, we are considering that nuclear reactions do not fall under the umbrella of chemical reactions. Whether this is a proper classification is an issue of semantics and not relevant for the overall point.

†Unfortunately, many scientists still think thermodynamics only has something to say about equilibrium, despite over 40 years of nonequilibrium thermodynamics.

clear constraints.

1.1.3 Geological Constraints

- Composition: The chemistry that takes place on a given site is dictated by the chemical composition of that environment. In geological records, much can be learned about local compositions from a long time ago, as well as more global trends, such as the atmospheric composition. The composition of the atmosphere has changed considerably over geological timescales. Scenarios for the origins of life normally start over 4 billion of years ago, in a time where the atmosphere was largely deprived of oxygen.
- (Bio)geochemical cycles: The chemical composition of the earth is essentially fixed. If important elements leave the (prebiotic) biosphere too quickly, it is of vital importance that such a loss is counterbalanced by a replenishment from elsewhere. Ultimately, this must lead to a cycling of elements through the lithosphere, atmosphere and hydrosphere in a roughly balanced fashion.
- Age of the earth: For a historical overview of the issue see [2]. Radiological dating of lead samples on earth compared to the isotopic composition of lead in iron meteorites (suggested by Houtermans [3]) has led to the current estimate of 4.54 ± 0.05 Ga as the age of the earth. For scenarios relying on abiogenesis on earth, this provides a constraint on the timescales that can be invoked.
- Microfossils: Microfossils are fossilized remains of microscopic size. The dating of fossilized remains of microorganisms fixes a time before which such life must have emerged. Microfossils are the object of some controversy[4]: there are abiotic processes that make very similar pseudofossils. To make convincing microfossil claims, it is therefore instructive to use several lines of evidence[‡].

1.1.4 Biological Constraints

- The endpoint of abiogenesis on earth: A scenario for the origins of life on earth must be consistent with modern biology as an endpoint. Chemically, this must be consistent with all presently known organisms using similar homochiral building blocks, genetic machinery and metabolic pathways. All currently known organisms are composed of one or more cells and use water as a solvent.

1.1.5 Other constraints

Many other constraints exist that can be considered to have the status of a fact or be close to it. For example, results from molecular phylogenetics are providing a more detailed picture of evolutionary history, providing clues of what aspects of life can be considered more 'ancient'.

The purpose of this section is not to be exhaustive, but rather to illustrate that there are several things we already know with great certainty. The insights in this highly incomplete list come from different scientific disciplines and underlines that OOL is a multidisciplinary effort. The list also provides a contrast with the upcoming section, where we consider some common assumptions in different branches of OOL.

1.2 Assumptions in the origins of life

Origins of life research is rife with domain-specific assumptions. Of course, making some assumptions is necessary if we want to make progress. Nevertheless, we should be aware that they are

[‡]An example of this is provided by Sugitani et al[5]. Their samples were covered with relatively uniform layers of carbon with a distinct C-13 isotope abundance, dated to be 3.43 Ga old. The isotopic composition of surrounding pyrite crystals (S-33, S-34) was also distinct and they were interpreted as metabolic byproducts.

being made. Many assumptions derive from current paradigms and from our limits of imagining things differently. The paradigms in origins of life have shifted and diverged many times in history (see Sec. 1.3). This has repeatedly been a consequence of new observations, experiments and occasionally theoretical considerations. Major paradigm shifts will likely happen again. It is therefore instructive to inspect some assumptions that are commonly made. In this section, some common assumptions that are explicit or implicit in the literature will be listed and marked in the following way:

- Assumption.

In doing so, we hope to provide some structural guidance in understanding the arguments of different schools of thought in OOL. We will also highlight what are the limits of some of these assumptions.

1.2.1 Chemical Evolution

Chemical evolution is a hypothetical form of evolution that is thought by many to precede modern Darwinian evolution involving genomes. It has been given many different definitions and mechanistic interpretations [6]. There is presently no consensus on what chemical evolution is or how it works, even less so a 'modern synthesis' of prebiotic chemical evolution. This vagueness and historical baggage leads some to avoid the term entirely.

Given the complexity of the genetic machinery, it seems that even a minimalistic protogenetic replication system is too complex to arise spontaneously. It is then speculated that a more rudimentary chemical process is at play, with some form of 'memory' and means of further progression. Concretely, such a memory is proposed to come in the form of self-replication, through various forms of autocatalysis.

Some commonly entertained assumptions concerning chemical evolution are the following:

- Chemical Evolution must be Darwinian Evolution
- Chemical Evolution implies polymers
- Chemical Evolution requires replicating compartments (e.g. protocells)
- Autocatalysis is exceedingly rare outside biochemistry

There are early models and speculations of chemical evolution that don't need these assumptions, and a number of upcoming chapters are devoted to this point. A model in Ch.8 strongly relaxes the first assumption. In Ch.5 and Ch.6 it is demonstrated that, to acquire the properties often invoked in chemical evolution, chemical systems can be constructed that do not require any of these assumptions.

1.2.2 RNA World Hypothesis

Speculations on the origins of the genetic code due to an early role of polynucleotides began at the end of the 60's, with a book by Woese [7] and papers by Crick [8] and Orgel [9]. Back then, mRNA, tRNA and rRNA had been discovered, highlighting the versatile use of RNA, but it was considered that it had no further catalytic role. An important element of these speculations was the consideration that RNA could have played such a role in an earlier stage. E.g. concerning the ribozyme, Crick notes *we cannot help feeling that the more significant reason for rRNA and tRNA is that they were part of the primitive machinery for protein synthesis.*

The discovery that RNA still performs catalytic tasks in extant biochemistry came highly unexpected. In a 1986 editorial article, W. Gilbert (known for Maxam-Gilbert sequencing [10]) coined the term 'RNA world', and speculated on the start of evolution as a soup of RNA molecules that would catalytically assemble themselves from a nucleotide soup, exploit recombination ('molecular sex' [11]) and mutations *to explore new functions and to adapt to new niches*. After the discovery of catalytic RNAs the RNA world hypothesis quickly rose to prominence. Modern formulations are numerous and diverse, insisting on particular details, mechanisms, or impossibilities [12]. A

well-known formulation of the scenario by Joyce and Orgel is known as " the Molecular Biologists' Dream", of which Orgel formulated a less radical version in 2004:

First we suppose that nucleoside bases and sugars were formed by prebiotic reactions on the primitive Earth and/or brought to the Earth in meteorites, comets, etc. Next, nucleotides were formed from prebiotic bases, sugars, and inorganic phosphates or polyphosphates, and they accumulated in an adequately pure state in some special little "pool." A mineral catalyst at the bottom of the pool—for example, montmorillonite—then catalyzed the formation of long single-stranded polynucleotides, some of which were then converted to complementary double strands by template-directed synthesis. In this way a library of double-stranded RNAs accumulated on the primitive Earth. We suppose that among the double-stranded RNAs there was at least one that on melting yielded a (single-stranded) ribozyme capable of copying itself and its complement. Copying the complement would then have produced a second ribozyme molecule, and then repeated copying of the ribozyme and its complement would have led to an exponentially growing population. In this scenario this is where natural selection takes over. Darwin suggested that all life is descended from one or a few simple organisms that evolved on the Earth long ago. According to the more radical scenario of the Molecular Biologists' Dream, the whole biosphere descends from one or a few replicating polynucleotides that formed on the primitive Earth about four billion years ago

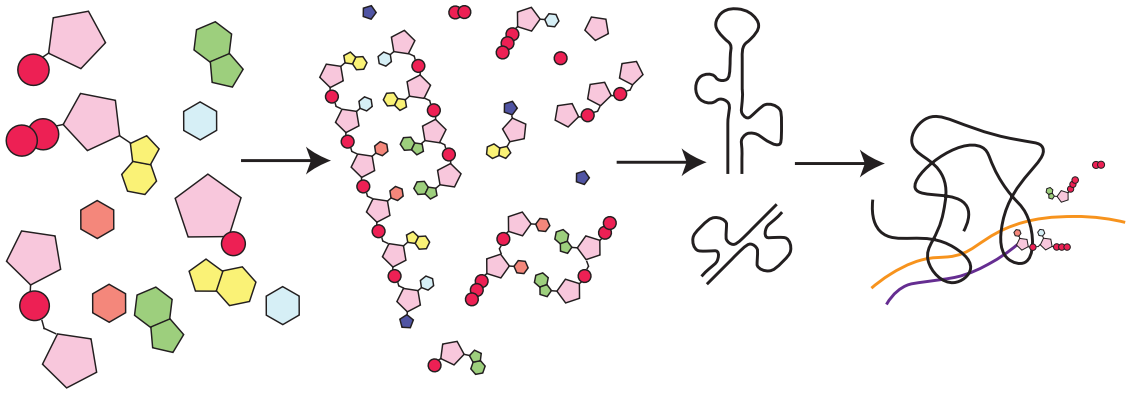


Figure 1.1: A typical scenario of the RNA world hypothesis. Drawing inspired by Refs. [13] and [14]. In Joyce's timeline, the first stage corresponds to prebiotic chemistry (4.2-4.0 Ga). A second stage (implicit in Joyce's timeline, which merges it with prebiotic chemistry, explicit in many other works RNA world research) invokes prebiotic activation chemistry, template-assisted ligation and other mechanisms to convert prebiotic monomers to increasingly polymerized and activated species[15, 13, 16]. Some of these species acquire functional folds (4.0 Ga for pre-RNA species) and support a transition to an RNA world (3.8 Ga, here represented by an RNA-based RNA polymerase, as in Joyce's timeline).

Various stages of such a world have been placed on a very influential origins of life timeline, proposed by G.F. Joyce [14]. On this timeline, Joyce reserves three periods for abiogenesis: i) prebiotic chemistry (4.2-4.0 Ga), a period that he argued would serve to form the activated biomonomers, which later on would be proficient enough to polymerize and form ii) a pre-RNA world (4.0 Ga) of alternative genetic polymeric species, which would later on be overtaken entirely by iii) the RNA world (3.8 Ga), where RNA becomes the genetic molecule. Some of these conceptual transitions have been drawn in Fig. 1.1. What is often lost in translation, is that the timeline is a 'best guess' for something that is itself a hypothesis. The details of a potential RNA world are still under heavy debate. If it turns out there was something reminiscent of Joyce's stage iii), then that does not tell us how we got there. Joyce's argument for stage ii), genetic pre-RNA polymers, is the exact same argument Orgel gave in 1968 [9], namely that only genetic polymers

have the capacity to perform effective enough evolution to get anywhere in early abiogenesis. A central debate in OOL is whether that is true or not.

It is hard to overstate the influence of the RNA world hypothesis on our thinking about the origins of life, having even entered some high school biology textbooks. The picture of prebiotic chemistry being the process that activates monomers to form genetic polymers as soon as possible is paradigmatic to this day within some OOL communities. A lot of experimental and theoretical RNA world research has been aimed at finding chemical pathways that fit this exact picture. The synthesis, activation and (sufficient) polymerization of RNA or XNA's is considered as one of the key challenges in the field and some to consider such a transition as the start of life.

Many proposals for where life started have coevolved with the RNA world hypothesis. Publications discuss in length how a particular environment helps RNA to form, such as for the fairly recent Mica-first scenario [17], how it can accumulate, such as in a pore with a temperature gradient [18], or be stabilized against degradation, as happens due to low temperature in the vicinity of ice [19], and so forth.

If an environment is inhospitable towards RNA, it is considered a less plausible candidate for life's origin. For example, Lepper et al. recently reported the hydrolysis rate of cytosine for 'Deep-Sea Black Smoker' conditions [20], noting *Our results on the stability of cytosine point away from a high-temperature/high-pressure origin-of-life scenario in favor of a low-temperature/low pressure environment for the emergence of the first RNA-based life forms*.

Joyce's timeline proposes an elegant answer to the question: "how can chemistry evolve to something complex, by the most direct means possible?". It does, however, need to overcome the initial problem of forming functional polymeric replicators, which is considered to be an extremely rare event. A similar argument was considered for peptides in various instances of the Oparin-Haldane hypothesis, which by a rare event would form self-replicating peptides. It has often been reiterated that prelife is rare and takes long to form, due to the extremely rare events needed to set things in motion. This is, in fact, an assumption, whose popularity goes hand in hand with the presence of such rare events in the RNA world hypothesis and the Oparin-Haldane hypothesis [21]. While often stated as a trivial and intuitive fact, we have no proof that extremely rare events are inherent to abiogenesis or not. At present, there is no substantial evidence that rules out fast abiogenesis, or a slow abiogenesis as a succession of many fast steps.

Let us now explicitly list some popular **assumptions** that have been entertained in the field

- There was an RNA world preceding the DNA/protein world
- The transition from prebiotic chemistry to an RNA world can be understood in terms of RNA
- A major stage in the origins of life timeline is the polymerization of nucleotides
- A major stage in the origins of life timeline is the formation of RNA-based RNA replicases
- Starting prelife is a slow process requiring very rare events
- (chemical) Evolution implies self-replicating polymers
- There are sufficiently proficient and short RNA-based RNA replicases that can form abiotically

Orgel notes that not everyone agrees on what an RNA world is [12]. He also notes (together with G. Joyce) that it is first and foremost a hypothesis of a precursor world for extant biology [22]: *It should be emphasized that the existence of an RNA world as a precursor of our DNA/protein world is a hypothesis.* The RNA world is regularly presented as a historical fact, for which the details of the molecular biochemist's dream simply need some more tinkering. This trend is reflected in the choice of manuscript titles, introductions and overall presentation. For example, in 'Origin of the RNA world: The fate of nucleobases in warm little ponds' [23], the significance statement opens with: *There are currently two competing hypotheses for the site at which an RNA world emerged: hydrothermal vents in the deep ocean and warm little ponds.*

Within communities working on RNA world scenarios, some of the other assumptions in the list

cause discomfort. The abiotic formation and ligation of RNA is far from trivial for thermodynamic and kinetic reasons, which is but one of the obstacles for making RNA-based replicases. The following assumptions, entertained by some, are a direct consequence of this

- There was a pre-RNA world, e.g. an XNA world
- There was no RNA-based replicase
- Chemical evolution implies copolymer sequences
- RNA evolved through cross-catalytic networks

Almost 50 years ago, it was speculated that RNA might have catalytic capacities [7, 8, 9]. This suspicion has now been overwhelmingly confirmed. A 2007 Review [24] already listed 27 reactions for which RNA catalysis had been demonstrated. Nevertheless, we are nowhere near a ribozyme that can maintain a genome in the way it is presented in the corresponding scenarios. In a 2012 perspective called "the eightfold path to non-enzymatic RNA replication" [25], J.W. Szostak identified eight "major problems" that stand in the way of RNA replication in protocell vesicles, namely: 1. Low regiospecificity in polymerization, 2. RNA duplexes stay complexed (too high Tm), 3. Low copying fidelity 4. Low copying rate 5. Unsuitable RNA reactivation chemistry 6. Degradation of RNA and precipitation of fatty acids by divalent metal ions 7. Primers cannot enter the vesicle 8. RNA strands reanneal too fast. The perspective was nevertheless very optimistic, as mechanisms and chemistries were being proposed to individually attack each of these problems and many more have been proposed since.

Of particular interest is a recently reported ribozyme from P. Holliger's lab, that links triplet building blocks instead of monomers [26]. Concentrated triplet building blocks (here promoted by a eutectic water-ice phase that also slows degradation) can bind strongly enough to unfold a ribozyme and this was used to synthesize five segments of the catalytic subunit, which were then partially linked into two fragments that self-assemble to form a new functional subunit. The triplets also served as a primer and permitted addition in both directions, with higher copying fidelity than observed for mononucleotide polymerase ribozymes.

Given the large variety of roles played by RNA, it seems only reasonable to suspect it has played a major role in the course of chemical evolution. Whether that role came early on is unknown. It is (at present) unclear if there was a stage that was so dominated by RNA that we can feel comfortable calling it an RNA world, but clues from molecular phylogeny favor a major role for RNA. The use of triplets and dividing ribozymes in fragments [27] are proving to be fruitful strategies, that start to blur the lines between the replicase approach and the chemical networks approach.

1.2.3 Iron-Sulfur world and metabolic perspectives

The more recent iron-sulfur world (scenarios often come with a 'world' suffix to distinguish themselves) was formulated by G. Wächtershäuser in 1988 [28] and further ideas were developed in upcoming years, including a proposal for a prebiotic metabolic cycle [29, 30]. In his 1988 paper "Before Enzymes and Templates: Theory of Surface Metabolism", the hypothesis of a 'surface metabolism' was developed. In this work, of which some key aspects will be summarized here, Wächtershauser describes a process, in which a mineral surface is in contact with a flowing bulk phase (similar to a mineral surface in a CSTR). Of particular interest are mineral surfaces that develop positive surface charges, such as pyrite (FeS), which can largely be attributed to divalent metal ions (of which are mentioned $Fe_2{}^+, Mg_2{}^+, Mn_2{}^+, Ca_2{}^+$ and $Zn_2{}^+$). These ions can attach and detach, providing species with variable surface lifetimes.

By the action of certain organic molecules, these metal ions can coordinate to stick more effectively to the surface, yielding longer-living species (e.g. phosphate groups are particularly proficient at bonding to positive surface charges). More generally, surface-bound ions may bind a variety of ligands, which endow the resulting complex with various interesting properties. Of key importance is the catalytic repertoire afforded by these surface bound metal-ligand complexes,

which produce new ligands, that in turn form new catalytic complexes. Such a process can then lead to series of complexes synthesizing each-other's ligands. By the appropriate combination of catalyzed pathways, cross-catalytic networks emerge (which are autocatalytic overall, see Fig. 1.2) . These networks constitute a 'surface organism' or surface metabolism, a molecular layer that feeds on the bulk flow and spreads out in two dimensions.

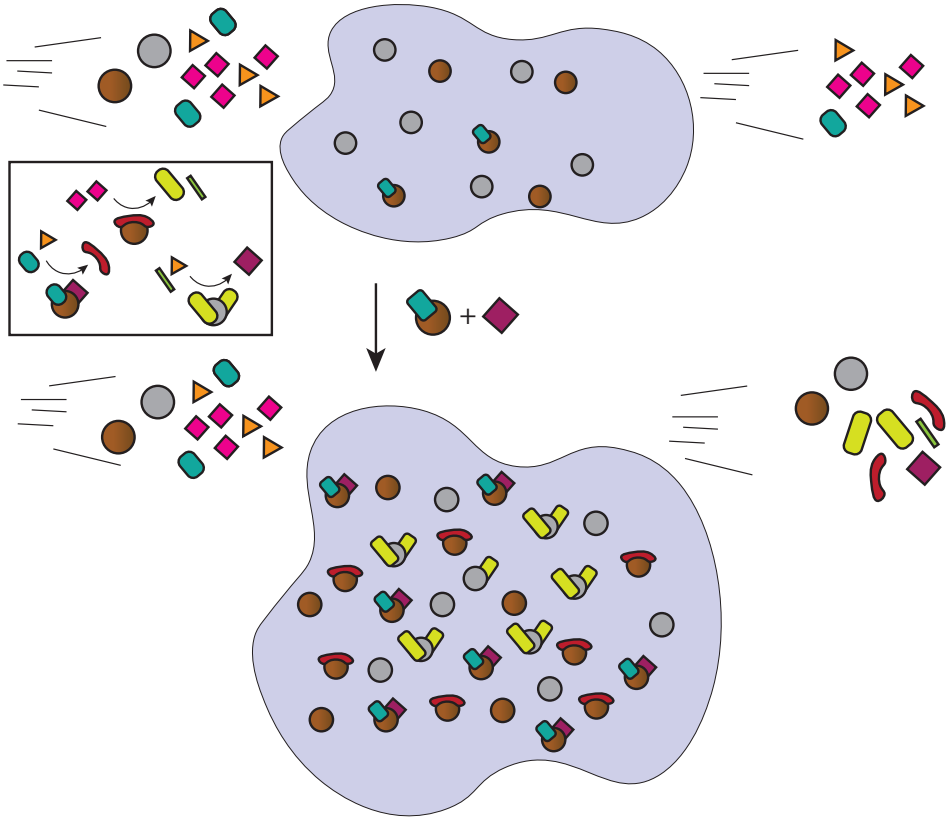


Figure 1.2: An illustration of Wächtershausers autocatalytic evolution mechanism: cross-catalytic ligand production. Metal ions and reactants are flown in, leading to metal adsorption and association with ligands. When the rare purple square species is formed, a new catalytic complex can be formed, which in turn can synthesize the ligands for other catalysts that produce each other's ligands, leading to overall autocatalysis.

In doing so, the layer spread to new surfaces, thereby encountering new environments. Such environments were suggested to lead to new manifestations of the surface organism, but also to be occasionally too hostile, leading to extinction. In turn, the surface organism modifies the environment, by locally modifying the surface and the release of products in the bulk.

Among the proposed compounds that modify the environments, are surfactants. It was suggested that the lipid biosynthesis pathway using isoprenoid phosphates could be mediated by a surface, thereby accumulating an increasing amount of lipophilic material on the surface, which started out as hydrophilic. Since detachment is largely mediated by hydrolysis and protonation, lipophilic patches can considerably increase the lifetime of many species, by locally expelling water. In such an environment (essentially a hydrophobic solvent), new reactions can take place, and e.g. condensation equilibria are considerably shifted, which were argued to lead to di- and triphosphates.

With a sufficient accumulation of surfactants, crowding expels lipids to the bulk and to the water lipid interface, to form a double layer, where hydrolysis may cleave of phosphates to yield

an asymmetric double lipid layer [§]. The initially hydrophilic surface is now largely covered in hydrophobic double layers, with small hydrophilic patches. By the subsequent envelopment of these patches, water droplets covered by a membrane are formed, which are subsequently released as vesicles with a built-in metabolism in its double layer (see Fig. 1.3). From here on, metabolic evolution could continue and gradually build up other biological structures. Wächtershäuser's

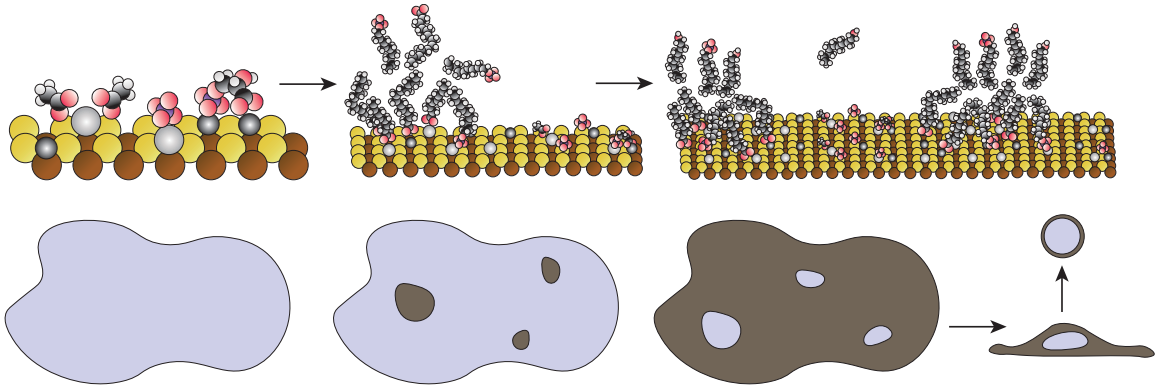


Figure 1.3: An illustration of some steps in Wächtershauser's scenario[28]. After a series of autocatalytic evolution steps, lipid synthesis takes off, forming lipophilic islands among the hydrophilic surface metabolism, enabling new chemistry and locally displacing equilibria. The growth of the lipophilic phase (dark grey) leads to small hydrophilic islands (light blue), which are subsequently covered by an asymmetric lipid bilayer and detach to form protocells.

approach is characterized by attention to detail in extant biochemical pathways, which he attempts to deduce from a surface-bound state. This is done with the clear reservations that should accompany a prebiotic scenario:

The proposed models are grossly simplified and deficient and at best in need of drastic revision and improvement. It would be preposterous to assume that, with this first attempt, the true historic chain of events could be reinvented

An interesting feature of Wächtershäuser's chain of event is that they are gradual, it supposes an absence of rare events. It is supposed that there is a clear trajectory that is set by thermodynamics and kinetics, that can be considered deterministic. The molecules that are invoked are (partly) justified by their context, with phosphate being a prime example of a key species, responsible for gluing the surface metabolism together and facilitating its later detachment through phospholipids. Its extant central role in biology (e.g. as an energy carrier) is described as 'windfall', a pleasant surprise.

The molecular detail of description of Wächtershäuser's work has provided a number of interesting predictions, that are in principle testable. Using iron-sulfur chemistry, interesting experimental results have been advanced by Wächtershäuser and colleagues, such as an $(\mathrm{Fe},\mathrm{Ni})\mathrm{S}$-dependent CO-driven peptide cycle [31]. What is still missing, however, is strong experimental evidence for metabolic evolution. In the words of Orgel [32] *The demonstration of the existence of a complex, nonenzymatic metabolic cycle, such as the reverse citric acid, would be a major step in research on the origin of life, while demonstration of an evolving family of such cycles would transform the subject.* Providing such evidence is a daunting experimental challenge. It is thought to require the monitoring of a large diversity of a priori unknown chemical species. Addressing this point is only now starting to seem feasible [33] with the introduction of non-targeted analytical methods with unprecedented resolution[34], such as Ultrahigh-resolution ion cyclotron resonance

[§]Wächtershauser argued that this explained the asymmetric nature of cell membranes today, with nonionic surfactants being most abundant on the outside

mass spectrometry (FT-ICR-MS).

There are some clear assumptions in the work. By being very explicit, Wächtershäuser aims to provide a testable scientific theory along criteria laid out by the Philosopher K. Popper. Of course, every extra detail is an extra element that could be 'wrong', and we will not go into all of them. Let us here just make some of the broader assumptions explicit:

- Prelife started on surfaces
- Prelife starts with autocatalytic evolution
- Autocatalytic surface evolution is proficient enough to provide all transitions until a genetic mechanism takes over
- Autocatalytic surface evolution is a deterministic process without rare events
- Protocells are an early invention
- Prelife relies on FeS chemistry
- Prebiotic surface chemistry mimics extant biochemical pathways
- We can conceive abiogenesis as a process that repurposes extant biomolecules to their modern use

The final two assumptions allow to formulate an elegant, gradual trajectory to extant biochemistry. At the same time, it does so while only requiring a small number of extra ingredients. This makes for a theory with considerable explanatory power, and many testable elements.

Here, two delicate points come up. One point comes from the alkaline hydrothermal vents community[35], which criticizes the lack of CO_2 sequestration, which is a central to modern biology. Since metabolic evolution is a central feature of the hypothesis, a transition towards CO_2 use can be argued to follow from a metabolic takeover[36]. Sojo et al argue, however, that CO_2 sequestration should be considered from the start, as it would provide a more continuous trajectory.

Another delicate point comes up with the latter two assumptions. Let us paraphrase Krishnamurthy's critique[37]: our extant biochemistry may not be the unique destiny of prebiotic chemistry, and extant biochemistry may not give adequate information on prebiotic chemistry and its evolution. This argument is equally valid for surface chemistry: abiogenesis may have involved a plethora of intermediate stages, structures and chemistries that we do not find back in modern biochemistry.

When we try to connect a set of well-known states, such as in cladistic approaches, we look for the most parsimonious trajectory to connect them. What is problematic for prebiotic scenarios in general, is that the starting point and its intermediate stages are largely unknown. The concept of parsimony becomes blurred by our ignorance of the states we need to connect and how that happens. In fact, we neither know how much intermediate states we need to consider, nor how much time they have to transition (if that is even an appropriate way to conceptualize the problem).

Presently, abiogenesis reaching unicellular life is attributed a timeframe of several hundreds of millions of years. To appreciate if that is ample time or not, we will need to wait until we understand the mechanisms and timescales of the problem. At that point, more robust arguments can be built on parsimony.

Some complementary metabolic perspectives

Other authors have expressed views that are often grouped as 'metabolic' or 'metabolism-first', insisting in particular on continuity[35]. The works of H. Morowitz and E. Smith underline the importance of nonequilibrium systems, in which free energy gradients fuel cycles [38]. These 'more organized' nonequilibrium states may in turn organize further to yield ever more ordered processes and new pathways to dissipate free energy. In [39] "The Origin and Nature of Life on Earth: The Emergence of the Fourth Geosphere" it is argued that life can be seen as a manifestation of an inherent self-organizing drive towards the increasing dissipation of free energy. This interesting and provocative view is supplied with ample nuance, since one should not conflate entropy production with life. Indeed, from early[38] on, Morowitz has warned against the uncritical use of *entropy production as an organizing tendency*, highlighting that we should consider that a large variety of

mechanisms, dynamics and kinetics come into play.

In his monograph 'Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis' [38], Morowitz discusses in length the conditions, constraints and clues that would inform a prebiotic scenario. In conclusion, a scenario is proposed, with the reservations that fit such a proposal:

The best we can do at this stage is to write a scenario in accordance with the known laws of science and consistent with our knowledge of the geological history of the Earth. Of course, the problem is that there are many possible scenarios within the broad constraints that have just been outlined. I opt for the most plausible one, with a caveat that plausibility is in the eye of the beholder.

Morowitz proposes that organic matter was brought in by meteors, comets, and solar winds, which in turn formed more complex molecules by the action of sparks, lightning, radioactivity and UV. Some of these compounds were amphiphilic, leading to colloidal species, which under environmental agitation formed the first vesicles.

Within these vesicles, compositional differences could develop and chromophores in vesicles would absorb visible and near-UV light to form polarized species, which would generate transmembrane potentials and be coupled to chemical reactions and transmembrane transport. These protocells develop, in a subsequent step, a self-sustaining metabolism based on phosphates, CO_2 and exchangeable two-proton carriers (AH_2). Morowitz then states that a self-sustaining protocell requires the synthesis of (1) amphiphiles, (2) chromophores and (3) energy transduction molecules.

In a subsequent stage, Morowitz hypothesizes that prebiotic amino acids and peptides could "absorb unto the surface of the protocell membrane and have specific catalytic functions". These amino acids were thought to cross-catalyze, leading to a net autocatalysis that would change the rate of growth of vesicles. By competition between protocells containing different autocatalytic networks, a basis for Darwinian evolution is set. Morowitz envisions the production of a genetic code as a late event, first in the form of RNA, then transitioning towards DNA.

In his 1992 scenario, Morowitz goes into considerably less molecular detail than Wächtershauser, but some clear analogies can still be drawn. Both scenarios see the formation of vesicular protocells as a key step and invoke an early self-sufficient metabolism. Early chemical evolution is described in terms of autocatalytic metabolic networks.

A clear difference is that Wächtershauser explicitly starts with autocatalytic evolution as the mechanism that supplies the necessary chemical ingredients for subsequent transitions. This evolution can be thought of as a form of 'chemical evolution', in which the notion of an 'individual' is initially absent. Morowitz only considers an evolution in the more classical Darwinian sense, operating between compartments.

In later work with E. Smith, the origins of life is conceptualized as a series of phase transitions, with an emphasis on the ecological and geochemical nature of these events. They advance the caveat that the details of these transitions may be blurred and not all transitions may be captured by their present work. Their mechanistic scenario is captured very succinctly in the abstract of chapter 6 in of their book[39]

The first carbon fixation was mineral hosted. Feedbacks, initially via cofactors and later via oligomer catalysis, lifted core metabolism ''off the rocks". The emerging identity of the biosphere reflected the growth of autonomy as much as of chemical invention. Passage to an oligomer phase corresponding to the ''RNA World" was a complex and heterogeneous transition, which transpired, and froze into place, in an already ordered organisynthetic context. We propose that cellularization occurred relatively late, and relied on functions of oligomers established in a mineral-hosted environment. The emergence of ribosomal translation originated in two parallel worlds of iron-RNA condensation-catalysis and template-directed ligation, which came together to form the first translation apparatus from mRNA to peptides. The refinement of translation fidelity, together with more precise RNA or DNA replication, ushered in the era of vertical descent along lines first appreciated by Carl Woese. Even in the era of evolution of effectively modern cells, many of the

major transitions have been determined by biogeochemical reorganizations.

Particular attention is dedicated in the works of Morowitz and Smith to the metabolic structure of modern life, and how it can be retraced to a prebiotic rTCA cycle. In order to provide a continuous account from prebiotic chemistry to modern life, this is a desirable property and is motivated by a similar philosophy that drives Wächtershäusers attempts to connect extant pathways to prebiotic environments. Illustrative is the following statement[39]: ... *explaining how bicarbonate addition entered biochemistry will be a crucial problem for any metabolism-first theory that purports to account for the selection of the biosynthetic precursors in life as we know it.*

In terms of assumptions, Smith and Morowitz share many assumptions with Wächtershäuser, and their perspectives can be regarded as complementary. A distinctive feature is the focus on modern metabolism. Some notable assumptions are:

- Prelife relied on an early rTCA metabolism.
- The origins of life can be captured as succession of a large number of phase transitions
- A prebiotic scenario can be formulated, by figuring out how modern metabolic tasks were performed prebiotically.

The assumption can be questioned, again, by Krishnamurthy's point. The fact that modern metabolism seems to be quite universal speaks volumes about its success. However, that does not require it to be an early invention, a large number of (evolutionarily) relevant developments may have preceded it. All modern organisms also use DNA and RNA, whose modern use is one of the latest transitions in the Smith-Morowitz scenario. Here, the term metabolism-first seems to refer strictly to an early metabolism with strong ressemblance to a modern one.

If abiogenesis can be considered as a succession of a large number of phase transitions, many of which Smith and Morowitz claimed not to have found yet [39], we may wonder if a large number of them may precede the invention and dominance of the rTCA metabolism. At that point, we can reconsider the meaning of metabolism-first, but also the priority of "the selection of the biosynthetic precursors in life as we know it", which, again, may be a selection that was not set in stone from the start.

Computational metabolic models

An early idea on metabolism, adsorption and phase transitions can be found in F. Dyson's 1985 work 'Origins of Life' [40], in which a toy model is proposed for amino-acid adsorption on a membrane in a system small enough to exhibit large fluctuations in molecular populations. Dyson proposes a bistability, with a default state of some adsorbed monomers ('dead') and a second more active state, with the jump process to the active state being called 'the origins of life'. At first glance, the model does not seem 'chemically realistic'. It was praised for its provocative elegance: it treats the problem as a low dimensional toy model with a single phase transition. The Morowitz-Smith scenario insists on many distinct phase transitions on different scales and provides more chemically explicit arguments. In a revised version of Dyson's book [40], Dyson makes the case that his model can be seen in a larger tradition chemically explicit metabolic computer models that yield similar results. What sets his model apart is that it is a low-dimensional description solved by pen and paper.

Of particular prominence in computational metabolic models, is the GARD (Graded Autocatalysis Replication Domain) model[41], whose main proponents are D. Segré and D. Lancet. The model follows the dynamic evolution of species, described by a population vector $\boldsymbol{n}$. In a deterministic description (the dynamics is actually treated stochastically), the population evolution is presented by the authors as:

$$\frac{dn_i}{dt} = \left(k_i^+ \rho_i N - k_i^- n_i\right)\left(1 + \frac{1}{N}\sum_{j=1}^{s} \beta_{ij} n_j\right). \tag{1.1}$$

The equation can capture a variety of networks, and was upon introduction applied on a network of activated monomers (exchanged with an environment) forming dimers that can cross-catalyze each other's formation (see Fig. 1.4), which is characterized by cross-catalysis term β_{ij}. These are placed within exponentially growing vesicles that divide.

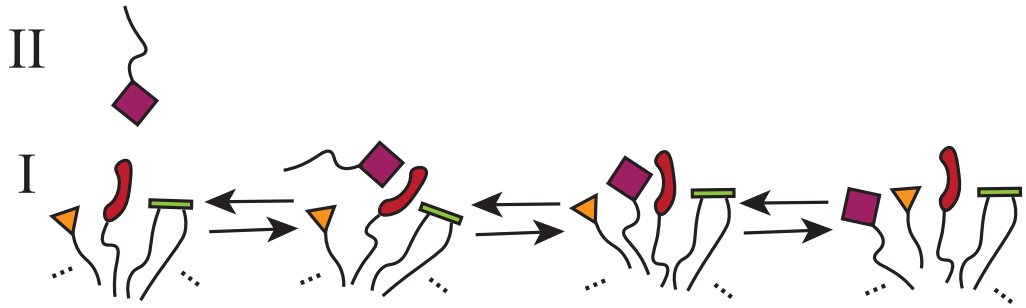


Figure 1.4: A simple illustration of a catalytic incorporation pathway. The red hollow head group complexes the square purple one, thereby facilitating the entry of the surfactant in the assembly.

Later formulations of GARD have merged the description of a compartment and autocatalysis: Eq. 1.1 then describes populations of different surfactants that can cross-catalyze each other's incorporation in a micelle[42] (see Fig. 1.4), that divides upon reaching a critical size (see Fig. 1.5).

For a large diversity of surfactants, different nonequilibrium compositions may be stabilized by strong enough cross-catalysis. These distinct attractors and their vicinity is referred to as a composomes, by analogy to a genome, as different compositions can correspond to different chemistries, which may be selected for. The number of surfactants needed to form a micelle (100) allows for fluctuations in population composition, facilitating transitions from one composition to another, similar to F. Dyson's model.

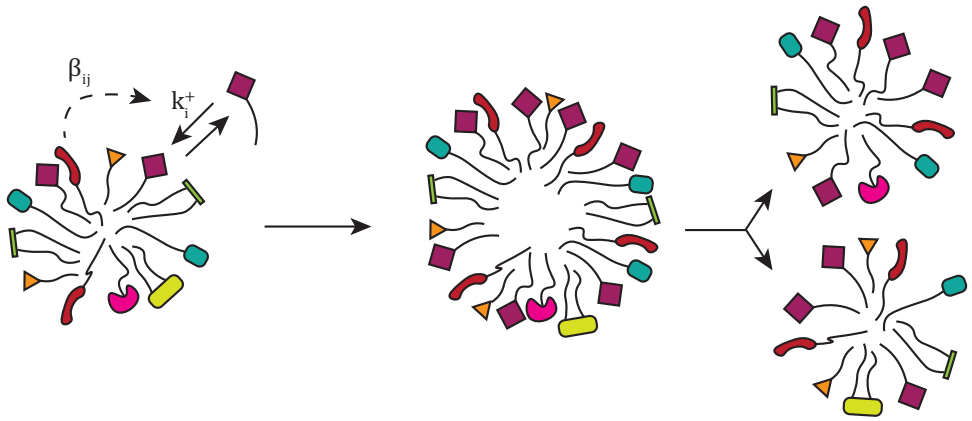


Figure 1.5: An illustration of amphiphile-GARD in a micelle. Incorporation continues until a critical size is achieved, after which two smaller micelles are formed among which the amphiphiles are partitioned. Illustration inspired by a particularly clear representation in Ref. [43].

A criticism that is raised is that the fluctuations may in fact be too large [44] to lead to a meaningful genotype and phenotype on an evolutionary timescale, while simultaneously admitting a relatively narrow space of states. By writing the model in terms of a master equation, they found that a stationary population composition of all compositions was quickly achieved, precluding further open-ended evolution. They then argued that such a critique is expected to hold for other proposals of autocatalytic evolution: *If (and what a big IF) there can be in the same environment distinct, organizationally different, alternative autocatalytic cycles/networks, as imagined for*

example by Gánti and Wächtershäuser, then these can also compete with each other and undergo some Darwinian evolution. But, even if such systems exist(-ed), they would in all probability have limited heredity only and thus could not undergo open-ended evolution.

Lancet et al. point out that the appearance of composomes depends not only on the appropriate strength of cross-catalysis, but also the diversity of the amphiphiles in the reservoir, and that fluctuations can be tempered by formulating GARD for larger structures than micelles, such as vesicles, which can consist of 10^6 surfactants. The latter can admit a larger amphiphile diversity and can have more stable attractors ¶. Furthermore, it is advanced that versions of GARD have been formulated that remain open ended for quite some time. It is expected that more elaborate versions containing covalent chemistry and a metabolism, MGARD[45], will do even better.

Lancet et al see the origins of life as starting with cross-catalytic assembly of lipids. Their 'protocells' start with surface incorporation, like in the Smith-Morowitz scenario and Wächtershauser's proposal, but has evolutionary characteristics reminiscent of Morowitz's 1992 scenario. It is imagined that such surfaces evolve and may at a later stage incorporate new chemistry in them, e.g. due to the arrival of peptides. It differs considerably from the Smith-Morowitz position, by not insisting on an early establishment of modern metabolism and highlighting the informational aspect of a composome as analogue of a genome. Nevertheless, GARD has often been classified as a metabolic theory.

The GARD scenario as formulated in e.g. Ref. [42, 45] comes with some clear assumptions

- Prelife started with lipid protocells.
- A prebiotic environment provided many thousands of different types of amphiphiles.
- These amphiphiles were sufficiently numerous and catalytically competent to yield stable composomes capable of open-ended evolution for a sufficient time.
- Diversity in lipid composition is enough to scaffold further transitions.
- Early chemical evolution occurred exclusively due to competition between composomes that assemble noncovalently.

Not all of these assumptions are necessary for a GARD-based scenario. Dimer GARD and MGARD show that covalent reactions work in GARD as well. If we can suppose an abundance of thousands of different amphiphiles, it seems surprising (and unnecessary) that the same soup is so impoverished in other compounds that covalent chemistry is effectively forbidden.

Simultaneously, other proposed forms of autocatalytic evolution (e.g. ligand decoration [28], nucleation of simple autocatalytic cycles [46], compartmentalized autocatalytic sets [47]) are excluded, as well as other chemistries and compartment types. This exclusion places the full burden of early chemical evolution on amphiphiles. While that makes the scenario more elegant to formulate and simulate, it places a lot of nontrivial requirements on lipids. This critique is not unique to GARD: any scenario that emphasizes the action of a single type of chemical species, a single type of compartment or a single type of environment risks the needless exclusion of other factors that are inadvertently there. These are most plausibly relevant and helpful to the scenario.

Different meanings of metabolism-first and metabolic theories

Amphiphile GARD has a replicating surface that uses exactly the compounds in the reservoir. It is what it eats, and thereby needs no further chemical processing (i.e. metabolism) of the food. The characterization of amphiphile-GARD (which is often what is referred to when GARD is discussed) as a metabolic or metabolism-first theory is therefore inadequate on the scale of the replicating unit (micelles, vesicles, etc.), although Lancet et al. note that such a label has been used for GARD [45] by some authors. Dimer-GARD, MGARD and other formulations with some covalent chemical modifications can perfectly well be considered as metabolic. If we do not limit scenarios to a pure

¶One should expect a trade-off here between the number of attractor states and their stability (lifetime), since one requires more stability and cross-catalysis, the other less. To the best of our knowledge, this tradeoff has not been assessed quantitatively.

amphiphile-GARD starting point, one can consider GARD scenarios that are metabolic from the start.

We may wonder how we should interpret the label 'metabolic': genetic polymer scenarios use covalent chemistry to copy their material, just like any other scenario. In practice, the terms 'metabolic scenario' and 'metabolism-first' are regularly used to underline a self-sufficient chemistry in the absence of genetic polymers, which seems a bit more general than one might with for (non-genetic would be more appropriate).

Some link metabolic scenarios exclusively to an early prebiotic rTCA cycle, like in the Smith-Morowitz case, by analogy to modern metabolism. Such an interpretation is frequently found in 'metabolism-first vs genetics-first' debates and provides a relatively clear position that can be argued for or against. However, it does not capture what some of the literature is now calling 'metabolic scenarios'. Many objections that are advanced against metabolism-first refer only to a subset of those scenarios.

A systems-level perspective

An important general critique on metabolic scenarios, is their apparent lack of transfer of hereditary information and evolution. It is important to appreciate that all metabolic models discussed in this section do perpetuate and replicate their nonequilibrium state, at least approximately. Unfortunately, information is often conflated with 'a genetic copolymer sequence'. Composomes and autocatalytic sets are but some examples of alternative means of information transfer. Whether the evolutionary potential of these mechanisms is sufficiently open-ended is a very different question. It is far from obvious that any of these mechanisms alone (GARD, catalyst decoration, evolving autocatalytic sets) can deliver on such a large promise.

At the same time, none of these proposed autocatalytic mechanisms requires the exclusion of other ones. It may be fruitful to approach the problem of abiogenesis from a systems perspective, as elucidated by Ruiz-mirazo et al[48]. Their perspective invites us to consider how these mechanisms may operate simultaneously and construct more general 'hybrid' scenarios. Such an approach would allow to relax most of the strong assumptions that are tied to any individual approach.

More provocatively, it lays bare that there may be many other relevant mechanisms, that one would not consider in a scenario based on one isolated mechanism or compound. In Ch.5 and Ch.6, we will demonstrate that there are many more autocatalytic mechanisms that have not been considered yet, and in Ch.8 we consider another type of evolutionary compartment dynamics that does not require lipids or cell division. In Ch.9 we consider how autocatalytic feedbacks can operate on higher-order cross-catalytic structures. All of these mechanisms can be bundled to bolster a systems-level scenario with strong evolutionary capacities. It may be practical to avoid labelling such a scenario as 'metabolic' and insist more on its 'autocatalytic' aspects instead.

1.2.4 X-first

Various authors advocate that a certain ingredient or aspect is so essential to life, that it must have been there 'first', from 'the start' of abiogenesis. Let us list here but some of them:

- Lipids first [42]
- RNA first [49, 9, 8, 7] (recently extended with viruses first [50])
- Peptides first [21, 51]

There is no reason to think that these are our only choices. Increasingly, a systems vision is emerging in which it is admitted that collectives of molecules can be more than the sum of their parts [48]. For example, peptides can confer considerable protection against hydrolysis to RNA molecules, leading some to consider they had to tag along from the beginning [52].

We should also entertain the option that not every molecule that was relevant for the origins of life, is still present today. Cairns-Smith made this argument by the analogy with a scaffold, and argued that clays could play this role [53]. Recently, it has been proposed that the space between

mica sheets may provide a breeding pool for RNA and other molecules [17]. The following assumptions rely on species that are not extant biomolecules:

- Mica first [17]
- Minerals first [53, 28]
- XNA (e.g. PNA) first [54]

Some authors have focused on particular features that were 'first' when a certain object emerged, e.g. for the composition of the membranes of protocells one can consider

- terpenes first [55, 28]
- lipids first [45, 42]

Given i) the large space of imaginable chemicals and the completely different environment that constituted the ancient earth, ii) the increasing number of new proposals of things that came first, it is plausible that nothing proposed so far actually 'came first'.

1.2.5 Apparent dilemma's and multilemma's

Origins of life research has advanced various scenarios, that are not always mutually compatible. This happens in particular when the suffix '-first' is introduced: not everything can come first. A dilemma that is popular today in certain communities of OOL concerns where life started:

- (pre)Life started in ponds vs hydrothermal vents

This dilemma is consistently repeated in science communication, where new arguments for either proposition are presented as a major advance. It should be kept in mind that numerous other environments have been proposed, even recently. A more complete list would be:

- (pre)life started between mica sheets vs wet rock pores vs on crystals vs deep under the surface of the earth vs ponds vs natural radioactive reactors vs silica compartments vs ponds vs hydrothermal vents vs etc.

A small number of dilemma's advanced in the literature are given here:

- RNA-first[49, 14] vs peptides first[21, 51]
- Metabolism-first[28, 38] vs genetics-first [49, 14]

RNA-first vs peptides first is highly likely to be a false dichotomy: perhaps neither was there from the start. Whether metabolism-first vs genetics-first is a true dilemma depends on our interpretation of 'metabolism-first'. If it only refers to an early rTCA cycle, then we are ignoring many scenarios. Unless we can justify why all other alternatives can be ignored, this alone immediately makes the dilemma a false dichotomy.

This way of presenting the problem, pervasive in popular accounts, but also in academic works, gives the false suggestion that there are no alternatives. Thereby a vast amount of scholarship is obscured, and logical fallacies are promoted (e.g. an argument against RNA world is suggested to imply metabolism-first, and vice versa).

If metabolism is interpreted in its broadest literature use, where even amphiphile-GARD is considered metabolic, then the dilemma essentially seems to condense to 'genetics-first or not'.

In arguing for a clay scenario, Cairns-Smith advances the following illustrative quote from a Sherlock Holmes story in his work: *It is an old maxim of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth.*

This statement seems logical taken at face value, but there is a crucial hidden assumption:

- Of all the options we have considered so far, at least one of them is 'the truth'

As time proceeds, more options are advanced, and this assumption becomes more likely to be true. For most of our multilemma's, however, we do not know if we have reached a point where this starts to become a reasonable assumption, especially since new ideas are still coming in.

1.2.6 Prebiotic Plausibility

The term 'prebiotically plausible' is mainly used in prebiotic chemistry, often to indicate that a chemical synthesis or process is in accord with geochemical or astrophysical lines of evidence. In particular, this means that we can make the following assumption

- the required substrates can reasonably be argued to have been present, sufficiently abundant and to encounter circumstances that admit the process under consideration.

In his criteria for what would constitute 'prebiotic plausibility', Orgel argued that it required chemistry that occurs in water [12].

What can be considered prebiotically plausible has changed considerably over time. The Urey-Miller experiment used an atmospheric composition (CH_4, NH_3, H_2O, H_2) that, at the time of the experiment, was considered prebiotically plausible by many and was part of Oparin's argument for prebiotic amino acids. In later years, evidence accumulated that ammonia and methane may not have been abundant atmospheric constituents, thus stripping the original experiment of its prebiotic plausibility status. Later versions of the experiment have addressed this critique using N_2 and CO_2.[‖]

'Prebiotically plausible' is, at least presently, not a perennial label. The term has recently received strong criticism [57] for a much more pressing problem: authors have increasingly started applying the label in a loose fashion. Benner illustrates the problem with the case of HCN and its derivatives, key ingredient in many prebiotic chemical synthetic schemes: ... *the prebiotic plausibility of HCN, the other molecules, and adenine long ago vanished as Earth-made species, even though literature too voluminous to cite here continues to assume otherwise.*

Some authors address this point by considering meteorites carrying HCN and derivatives as sources[58]. While Benner considers that meteorites provide insufficient material to maintain prebiotic synthesis of e.g. RNA, he notes that iron-rich meteorites may have have transiently (*perhaps for 100 million years*) endowed the crust with enough material to reduce the atmosphere for a considerable time. This would then have provided an abundant, transient source of HCN and its derivatives. Benner notes:

The planetary model may be evaluated under its own standards-of-proof, and will rise or fall based on criteria quite independent of criteria that are used to evaluate chemical models. Nothing is ever proven in science. However, a network of models, each subject to independent test in their own fields, makes the big picture more, shall we say, plausible.

Wrap-up

We have seen a considerable number of strong assumptions that come up in OOL research. By construction, many of these assumptions can only be true if other assumptions in the field are false. Demonstrating which ones are false is therefore of great interest, it allows to focus our attention on the most promising directions.

Making explicit what is sometimes taken for granted is a useful exercise. Prebiotic scenarios are rich in strong assumptions, some of which get conflated with facts. Upon reading the OOL literature, one gets the impression that a lot more is known about the origins of life than is truly the case.

It is in this instance where this chapter's opening quote by Socrates rings particularly true. There are many things that are still unknown, and we should treat them as such. To say we have considered all possibilities, requires an exhaustive demonstration that there are indeed no other options. A hunch or a clue can be an instructive guide for further research. Promoting assumptions and hunches to facts, however, is an academic faux pas that slows down the scientific process.

[‖]The experiment has been repeated using different atmospheric compositions in contact with different water mixtures. When N_2 and CO_2 were used instead of ammonia and methane, the experiment synthesizes appreciable amounts of amino acids as well. This was only recently recognized, because the simultaneous formation nitrites lead to amino acid degradation, unless these nitrites are sequestered. This was done using ferrous ion[56].

1.3 Paradigms and Paradigm shifts in Origins of Life Research

Concepts and practices considered pertinent to a scientific question are not fixed in time. For age-old questions like abiogenesis, the paradigm has shifted on many occasions. Kuhn argues that such shifts of the scientific paradigm occur through revolutions, for example when general relativity replaced Newtonian gravity[59]. Structurally, this is reminiscent of Gould's evolutionary theory of punctuated equilibrium[60], which divides evolution in short periods of rapid evolution and long periods of slow evolution ('equilibrium').

The scientific perspective regarding abiogenesis has shifted considerably over time. Developments in microscopy, (micro)biology, (bio)chemistry, geology and other fields provided major new ways to look at the problem. In this section, a number of points of view are considered along with the scientific milestones that preceded them. Not all points of view covered were equally dominant, but all are well-known schools of thought.

This exposé is far from exhaustive, it serves merely to demonstrate something: prebiotic scenarios are products of their time.

Pasteurization

It has long been believed that macroscopic life formed spontaneously, e.g. from mud or food. This pervasive 'generatio spontanea'-perspective was dismantled progressively. A series of rigorous experiments by L. Pasteur are widely regarded as the proof that laid the theory to rest for good. In Oparin's work "origin of life",[21] an extensive discussion is provided on these developments. Oparin notes that Pasteur's insight that life does not form spontaneously was necessary, but perhaps somewhat extreme. Abiogenesis would require such an event to have occurred at least once.

Oparin-Haldane Hypothesis

The prebiotic broth theory states that a mix of organic chemicals, formed by all sorts of geological activity, concentrated in protocells (coacervates), where they started forming oligopeptides, until eventually a large self-replicating peptide was assembled. It was arguably one of the first attempts to provide a serious chemical account of life's origins[21].

The Urey-Miller experiment and the prebiotic soup

The Ur-soup theory rose considerably in prominence after S. Miller's discovery of amino acid formed by reaction of a reducing atmosphere under electric discharge. The discovery led to a Nobel prize and the Ur-soup theory became part of many high-school history and science textbooks. Although the prebiotic atmosphere model Miller followed was accepted in his day, later insights suggest that the atmosphere would have a very different composition. Many chemists in OOL today consider the experiment as historically significant for marking the starting point of prebiotic chemistry. The quantities of prebiotic chemicals produced this way, however, are deemed small and the field has largely moved to other chemical sources, such as meteorites.

Genetics

Discovered one year before Miller's well-known experiment, the discovery of DNA started rapidly unveiling new ideas on templates and replication. At the end of the 60s, DNA and RNA started to be considered as an alternative for the self-replicating enzyme in the Oparin-Haldane scenario, starting the first speculations on an early polynucleotide world [7, 8, 9].

Ribozymes and RNA world

The idea of an RNA world has been entertained with some caution at the end of the 60's[7, 8, 9], relying on a speculation that certain shapes of RNA might possess some catalytic activity, but that modern life stopped using RNA for catalytic purposes. To everyone's surprise, it was not only found that catalytic RNA exists, but also that organisms still use catalytic RNA today.

In a 1986 editorial article discussing the recent publications on catalytic RNA, Gilbert raised speculations akin to those entertained in the 60's. In his article, he coined the term RNA World[49], denoting a time where RNA would be performing both catalysis and genetic tasks.

The RNA-World scenario rapidly established dominance after increasing experimental proof for catalytic ribozymes was found. The scenario has entered biology textbooks and much scholarly work can be found that explicitly assumes an RNA World existed.

Martian Meteorite ALH84001

In 1996, the publication of 'Search for Past Life on Mars: Possible Relic Biogenic Activity in Martian Meteorite ALH84001' [61] stirred the scientific community. This meteorite, thought to come from Mars, displayed lifelike morphologies under an electron microscope. Polyaromatic hydrocarbons were found, as well as carbonate globules containing magnetite and iron sulfides. These were interpreted as likely Martian microbe fossils.

The discovery was quickly picked up by president Bill Clinton's science advisor **. Within a week of the publication of the article, a press statement was released, in which it was made clear that Mars exploration and the search for life would become a top priority. In 1995, NASA's exobiology research was expected to be largely discontinued, and its nexus at the Nasa Ames Research Center was expected to close down [62]. The sudden Mars hype in 1996 drastically changed exobiology's fortune. A new astrobiology institute was founded, and Ames suddenly saw its funding increased to 5M dollars per year, which got bumped up to 15M in 2000. It also led to the reinstatement of Mars exploration programs.

Today, not many believe that ALH84001 contained Martian microbes. Extensive research suggests that the biogenic signatures and morphologies can also plausibly be generated by abiotic processes [63].

Chemical gardens and hydrothermal vents

Chemical gardens are growing semipermeable inorganic structures with a liquid interior. They acquire water by osmosis, which increases the pressure on the permeable walls. When the pressure becomes too high, the structure tears locally and the tear quickly reacts via displacement reactions, leading to further growth of the structure [64].

Chemical gardens were popularized at the beginning of the 20th century by Leduc in his treatise 'The mechanism of life', in which he opposed vitalism and argued that lifelike processes such as growth and development follow from inorganic chemistry and osmosis. Although popular at first, chemical gardens lost favor in OOL quickly with the advent of genetics. Renewed interest had to wait for the end of the century. With the discovery of hydrothermal vents, chemical gardens were suddenly promoted to an important geological phenomenon. Hydrothermal vents have become a popular environment for prebiotic scenarios, especially in conjunction with Wächtershäuser's iron-sulfur world[28] and related metabolic theories.

Gold's Deep Rock hypothesis

In 1992, Thomas Gold wrote an influential paper called 'The Deep Hot Biosphere' [65], in which it was argued that life should extend to many kilometers deep under the surface of the earth. Since radioactivity and temperature gradients constantly maintain gradients of reactive chemicals, no photosynthesis is needed. Well-adapted hyperthermophiles could then be expected at great depths, which Gold estimated to be 5-10 km for a temperature limit of $110 - 150^{\circ}$C

Gold argued that an origin of life may well have occurred in deep, porous rock, since chemotrophy requires much less sophistication than photosynthesis. Inspired by astrophysical observations such as alkane lakes, Gold had developed a theory that coal and oil reservoirs on earth need not be biogenic, which he supported with apparently anomalous compositions of these reservoirs. The

**see: https://www2.jpl.nasa.gov/snc/clinton.html

upward percolation of hydrogen, water, methane, higher alkanes and other fluids were argued to maintain nonequilibrium steady states that formed the breeding grounds for life. These ideas were developed in more detail in his book, also titled 'The Deep hot Biosphere'.

Gold's claims of life at great depths have been confirmed, and are still extensively being explored[66]. New geological processes have been discovered that generate H_2 and CH_4 as well as other important metabolites. The debate concerning mantle-sourced hydrocarbons has largely settled, with evidence suggesting that this is only a minor source of hydrocarbons. Whether the subsurface microbiome is endemic remains an open question. The subsurface is still relevant for prebiotic scenarios, although the focus is on hydrothermal vents.

1.4 Interdisciplinarity and origins of life communities

An exhaustive bibliometric study [67] found that the field consists of large subcommunities, some of which hardly know or cite each other, with 'microfossils and evidence of life on the early earth' being the most disconnected. The disciplinary makeups of these communities are very different. For example, RNA-world related research is largely populated by life scientists. Geologists are more abundant in other branches of research, e.g. hydrothermal vents, microfossils and minerals.

Increasingly, origins of life programs are acknowledging the challenges inherent to the multidisciplinarity of the field. Researchers can simply not be expected to be well-versed in all relevant (insofar as we know which ones are relevant) branches of biology, chemistry, physics, mathematics, philosophy, and so forth. This problem was assessed and (partly) addressed in an Astrobiology winter school on knowledge and attitudes, of which a report [68] aptly summarizes this situation

Even though astrophysics, geology, and biology are all natural sciences, they approach problems unique to their fields, with different tools and techniques, idiosyncratic jargon and different meanings of common terms, and even different philosophies about what constitutes good science. The central questions of astrobiology require interdisciplinary integration: a process of learning the research questions in fields other than your own, their jargon and essential theories, and a respect for the way that science is conducted in these sister fields (Boix Mansilla et al., 2016). These collaborations are slow to occur. For example, a bibliometric analysis of 1210 papers produced by scientists affiliated with the NAI between 2008 and 2012 revealed that 34.5 % of papers published were described by two or more Web of Science journal categories; but the two largest category clusters, Astronomy & Astrophysics and Geochemistry & Geophysics, had no links in common (Taskin and Aydinoglu, 2015).

This traditional divide between the sciences has a direct effect on the quality of publications in origins of life. On the one hand, a multidisciplinary work must live up to the standards and scrutiny of all the fields it involves. At the same time, it must strive to communicate its findings to a broad readership of different disciplines. These are nontrivial challenges with a common root. From their questionnaire among students in astrobiology programs, Burnam-fink et al note [68] measurable gaps in fundamental knowledge:

The communication gaps between researchers in these fields are sometimes stark. Fundamental yet specialized knowledge may not be covered in undergraduate scientific requirements and not learned as students specialize in graduate school. Astronomers may not know that photosynthesis can occur that does not produce O2 (Kulp et al., 2008). Geophysicists may not be aware that astronomers observe stellar abundance ratios to differ by factors of 2 from the Sun's (Suda et al., 2008). Biologists may not be aware that, even in the best-case scenario, astronomers will only observe exoplanets on a single pixel of a detector, and that disk-integrated atmospheres and surfaces are all that can be observed (Stone et al., 2015).

In the context of this manuscript, we find it important to point to a number of simultaneous developments in the literature, where authors from various fields (e.g. prebiotic chemistry [69], theoretical biology [70] and complexity [71].) are explicitly addressing aspects of nonequilibrium

thermodynamics and arguing for extensions. The theories they are suggesting to be in need of extension, however, are either equilibrium thermodynamics or the linear regime of nonequilibrium thermodynamics treated by Prigogine. While these works provide interesting insights in their own right, some of the questions they raise are already understood quite well and are captured by quantitative theoretical frameworks. Nonequilbrium thermodynamics and stochastic thermodynamics have made tremendous progress in the last 40 years. This should be seen as a wake-up call to researchers working in stochastic and nonequilibrium thermodynamics.

This trend has been noticed by other authors. Smith and Morowitz [39] made a similar observation in 2016, when confronted with the concept of dynamic kinetic stability: *The view that life must have emerged through a sequence of stages, that the intermediate stages must have been incrementally stable, and that the stability is essentially a process stability of self-regeneration, is being expressed more widely and more explicitly among origins researchers from many backgrounds. Our effort in Chapter 7 and Chapter 8 has been to show that the conceptual foundations (and a considerable body of technique) for such a theory exist and that these grow continuously out of the thermodynamics that accounts for our hierarchy of matter.*

A recent article pointed towards similar tendencies in nanotechnology and supramolecular chemistry[72], highlighting that different authors have started using the lingo of nonequilibrium systems in a divergent and confusing manner. The article also dispelled the still commonly perpetuated myth that thermodynamics can only be used for equilibrium systems, by explicitly highlighting some of the theoretical nonequilibrium frameworks. A particular illustration was given of open chemical networks and the efficiency of nonequilibrium energy storage by chemical synthesis. A call by the editors of Nature Nanotechnology[73] prompted a rigorous theoretical article on that exact problem soon after, by some of the theoreticians responsible for that framework [74]. These are encouraging developments, indicative of a renewed dialogue.

We wish to further promote this dialogue, by reiterating the message. Nonequilibrium thermodynamics is now a quite mature field, being taught in university courses to scientists and engineers[75]. The younger field of stochastic thermodynamics has made great strides in recent decades, accompanied with convincing experimental verifications [76]. We encourage further efforts to extend thermodynamics out of equilibrium towards new directions. These efforts would benefit - like any other scientific venture - from fruitful exchanges with the existing body of knowledge.

To provide support for such a development, the first chapters in this manuscript are dedicated to the introduction of some key concepts in nonequilibrium thermodynamics and chemical networks, in a way that is adapted to chemists: networks are illustrated, examples are given in terms of real chemistry, and some key concepts (e.g. chemostats) are extended towards their more typical experimental situations.

In the critical reading of OOL research, it is instructive to bear in mind the academic background of authors in relation to the claims they make. This manuscript is no exception: the author is first and foremost a chemist and a physicist, and the most convincing claims we hope to advance in this work pertain to these disciplines.

1.5 Summary of the content of this manuscript

In the preceding sections, a quick introduction has been given to origins of life as a field, illustrated by some of its intellectual history and the challenges of such a multidisciplinary venture. The goal here is not to be exhaustive, but to provide instructions on how one should interpret the literature and make explicit some of the key assumptions that underly popular scenarios for the origins of life.

We thereby lay the groundwork to interpret the results of the chapters that are to follow. The intellectual contribution of this manuscript to the field concerned with the origins of life can be considered the following:

i) We analyze in detail some **new mechanisms of prebiotic relevance**, such as a recently proposed form of multilevel selection: transient compartmentalization. We also consider new nonequilibrium mechanisms for the prebiotic formation of long polymers, and derive an abundant new form of multicompartment autocatalysis. This new form of autocatalysis leads to elegant ecological mechanisms of autocatalytic evolution. Furthermore, we find new chemical networks for highly efficient synthesis that can rival kinetic proofreading. We also find a new fundamental parasite problem, called a complexation catastrophe.
ii) We introduce **theoretical tools and frameworks for the study of origins-of-life problems**. These are derived from nonequilibrium thermodynamics, the theory of branching processes and chemical networks. Our framework for autocatalytic evolution generalizes existing approaches (GARD, autocatalytic sets, CESSPOOL, etc.), but also opens new vistas by its general applicability to new mechanisms (multicompartment autocatalysis). By this generalization and expansion, evolution without genes becomes considerably more powerful, general and plausible.

The framework for transient compartmentalization permits to assess a large variety of scenarios in great statistical detail with a low-dimensional approach. Our general nonequilibrium analysis of ligation-hydrolysis-activation cycles puts thermodynamic bounds on dissipative sequence exploration, which rigorously rules out prebiotic scenarios that rely on chemically fueling the search for sequences that are too rare.

iii) We provide a **provocative new prebiotic scenario, that relies on general mechanisms, not extant biochemistry**. The scenario is grounded on the structure of autocatalytic evolution and other general mechanisms found in this manuscript and in systems chemistry. It takes the ideas of R. Krishnamurthy ('extant biochemistry as a destination, not destiny') and systems chemistry ('organizing molecules to be more than the sum of their parts') as guiding principles, while insisting on gradual transitions. In doing so, we avoid the most assumptions that other scenarios make. In this approach, chemical actors can be found a posteriori, following from their suitability for the various essential qualities that are needed on a systems-level for autocatalytic evolution.

With a wealth of new mechanisms at our disposal, it becomes possible to envisage an increasing degree of chemical organization due to small incremental steps (and setbacks). The systems-level view of autocatalytic evolution highlights the wide range of prebiotic functions that molecules can play to become indispensable. Such insights provide a rationale for their introduction, which is indispensable when we wish to connect extant biochemistry to earlier stages of abiogenesis.

1.5.1 Summary of chapters

In this manuscript, we hope to contribute to the field in a number of ways.

In **Chapter 2**, we will introduce chemical networks, drawing largely on results from the Feinberg group, the Esposito group and others. We illustrate these frameworks with systems closely familiar to the experimental chemist, to make concepts such as cycles, conservation laws and deficiency intuitive for a wider audience. The framework is also reformulated for phase transfer and diffusion. We show that, to maintain the original structure of the theory, a kinetic description is best given in terms of molecule numbers instead of concentrations.

In the analysis of real chemical systems, interesting subtleties come up that require us to delve deeply into the exact meaning of (allo)catalysis, autocatalysis and chain reactions (which are explored further in Chapter 5), and their associated chemical network structures. Often, a description of a chemical network is highly coarse-grained, for very practical reasons. The network properties that characterize these networks are then made implicit. The net stoichiometry of reactions in a network, given by the stoichiometric matrix, may correspond to different net stoichiometries of forward and backward steps. In a formalism with reversible reactions, this generally requires two separate matrices of coefficients for an unambiguous characterization.

New results and tools are introduced to adequately treat chemical networks in general. Notably,

we show that chemical networks can always be described on a level of detail that leads to an unambiguous representation (nonambiguity), which we can build up from reversible unimolecular and bimolecular reaction steps.

In **Chapter 3**, thermodynamic aspects of open chemical networks are considered. A key point we wish to illustrate, is that a chemical network can be open in a variety of different ways. The details of how a system is opened up to an environment are of crucial importance to its behavior, which is illustrated by the treatment of simple chemostats and composite chemostats, a CSTR reactor, serial transfer and osmotically coupled compartments. Placing a chemical system in any of these ways subjects it to very distinct dynamics and conservation laws.

The ideal thermodynamic chemostat describes exchange of a chemical species with a large bath, which due to rapid exchange fixes its chemical potential and concentration fluctuations. Conceptually, an ideal bath may be seen as an infinite reservoir that is separated from the system by a perfectly specific membrane. In experimental practice, the role of a 'bath' can be fulfilled by compounds within the system that serve as a buffer (homogeneous chemostats), but this has clear limitations in size. A bath can also be due to compounds in other phases (external chemostat), such as a nonmixing fluid phase, precipitates, gases, etc. Often, the chemical potential of reservoir species is coupled, leading to a composite chemostat (e.g. dissolved ions in equilibrium with their salt). Such chemostats fix products of concentrations instead of individual concentrations, which means chemical species considered as 'food' are no longer independent. We also show how thermodynamic chemostats require a more detailed treatment in stochastic thermodynamics and imply a modified zeroth law for the exchanged of conserved integer quantities (e.g. atoms).

For CSTR systems, a chemical mixture is flown in and out of a well-stirred reactor. The outflow serves as an inherent degradation pathway and under its action all conservation laws of the network are broken. Asymptotically, however, the balance of influx and outflux lead to similar nonequilibrium constraints as the conservation laws for the closed system, which fixes the mass of the reactor contents and strongly constrains its possible behaviors. A similar way of opening the system is transferring a fraction of the reaction mixture to a fresh chemical supply repeatedly. This serial transfer is often used as an easy experimental simulation of a CSTR and we show in what limit they become equivalent and under what circumstances they start to differ.

Finally, we look at osmotically coupled compartments (droplets), which may exchange other small molecules via diffusion. Such systems are open and can grow and divide. However, to persistently grow and divide, the appropriate gradients need to be maintained. We show for two simple situations (exchange with a reservoir, exchange with a regularly refreshed neighbor droplet), that this requires autocatalysis.

In **Chapter 4**, the concept of information in thermodynamics is discussed in a chemically explicit way. The concept of information has caused considerable confusion, notably because it has often been conflated with (genetic) codes. To demonstrate the explicit thermodynamic character of information, it is shown (using equilibrium thermodynamics) how a macroscopic information engine can extract work from the racemization of pure enantiomers, in addition to some other examples. In origins of life and chemical synthesis, we are often interested in the reverse of this protocol: creating pure substances.

In the final part of the chapter, we look at chemical network structures that improve the synthesis of a compound, by adding irreversible reactions or dynamic (reversible) interchange. A prominent example from biology is formed by networks that perform kinetic proofreading, but there are many other networks, some not described before in the literature, with very different tradeoffs.

Several important networks are discussed in detail. The classical dichotomy between a kinetic product or a thermodynamic product holds only for the simplest of networks, and systems chemistry is rapidly uncovering elegant new strategies to push reactions to new extremes in efficiency. The simplicity of these error-correcting chemical networks (some involve only one extra reaction)

is provocative: abiotic functional networks may be considerably more prevalent than currently considered and it will be instructive to take this into account in conceptualizing chemical evolution.

In **Chapter 5**, allocatalysis, autocatalysis are defined in rigorous stoichiometric terms for chemical networks. This is made possible by imposing the properties derived in chapter 2, such as nonambiguity. This allows, for the first time, to readily identify these features for chemical networks in general, through the use of submatrix techniques. We also show that exchange between compartments can lead to emergent new types of autocatalytic cycles that are impossible for single compartments.

In **Chapter 6**, the concept of chemical evolution is discussed, a hypothetical mechanism that has served for over a century as the deus ex machina of origins-of-life scenarios to explain how simple molecules gradually became more complex.

Some of these models and theories from the literature are reviewed and some confusion in the literature is cleared up by pointing out some critical assumptions on the level of coars-graining. Then, it is shown that these different approaches to autocatalysis can be unified in the framework of stoichiometric autocatalysis.

We then proceed by rederiving a literature model for an evolving metabolism in a CSTR, based on structural features of reaction networks. This derivation leads to a more general model, with a larger diversity of autocatalytic cycles.

The probability of triggering an autocatalytic cycle that subsequently survives is shown to correspond to a survival probability in the theory of branching processes. Some 'evolutionary tradeoffs' for autocatalytic evolution in a CSTR are discussed.

We then consider the multicompartment case of the evolutionary model. Under a variety of circumstances, the equations simplify to those found for a CSTR. Survival now depends on the presence and composition of other compartments in the environment. Collectively, they may stabilize pathways that would not be viable otherwise. Autocatalytic evolutionary trade-offs for such systems are detailed.

In **Chapter 7**, we consider a number of nonequilibrium mechanisms that generate long polymers, such as recombination coupled with reservoir exchange, adsorption, and chemically driven ligation chemistry and discuss some heuristics and physical arguments to understand some of the typical polymer length distributions we can encounter in the literature. Our treatment of reversible recombination in closed systems is largely drawn from its corresponding publication in Ref. [77]. A discussion of length-dependence in RNA adsorption on mineral surfaces is based on our theoretical analysis in Ref. [78] (see in particular the Supp. Matt.). The original publication provides a broader perspective, supported with further experiments.

We also look at the mechanisms which allow polymers to acquire new sequences, such as activation ligation fragmentation cycles or recombination reactions. By finding the explicit thermodynamic cost of forming new sequences in a steady state, we find lower bounds for the typical free energy dissipated in the time needed to find a sequence. Many prebiotic scenarios rely on forming rare long polymers with special sequences. They form a key motivation to search for prebiotic mechanisms that form and enrich large polymers, so that large special sequences are more likely to appear.

Using our bound for dissipative sequence exploration, we can test the thermodynamic feasibility of such scenarios. Dissipatively finding one or more 'long, functional sequences' at random is generally a scenario that verges on the impossible, even under the most favorable circumstances that are physically possible.

The demonstrable problems with sequence search scenarios invites us to think deeper on the prebiotic role of polymers and sequence evolution. We propose a broader outlook on the potential roles of polymers in the origins of life, noting that many interesting properties do not require long specific sequences.

In **Chapter 8**, we study transient compartmentalization dynamics for molecular populations. In prebiotic scenarios, compartments are now mainly considered of interest for two reasons: i) the concentration of prebiotically relevant chemicals and ii) multilevel selection. Our discussion will focus on the latter. The contents of Ch. 8 are a synthesis of a published paper[79] and a preprint [80] in submission, and ongoing work (namely a subsection on noise-induced cooperation and a section on the complexation catastrophe).

By compartmentalizing molecules, the fates of molecules can be aligned by selection mechanisms acting on compartments and their populations. A molecular population that favors the survival of its compartment may thereby become more abundant. This can happen through cooperative mechanisms among molecules, favoring compositions that would not survive in a bulk phase through a process of group selection.

A similar mode of group selection is found in cell division, as exemplified by the stochastic corrector model[81]. Such mechanisms have been considered as an important means of saving replicase populations from parasites formed by replication errors, and thereby information maintenance. While replicases pertain to highly specific branches of scenarios, the mode of group selection proposed is a scenario-independent mechanism. A very similar group selection mechanism (but using cross-catalytic incorporation instead of template replication) is used in amphiphile-GARD.

Another recently proposed and experimentally tested mechanism (to which Szathmary contributed as well), is transient compartmentalization, in which (sub)populations are encapsulated, grown, selected and released, after which the cycle can be repeated. As the compartment is no longer divided in two but is destroyed after one cycle, the enclosed molecules are no longer constrained in their multiplication by a replication cycle. They can thus multiply by factors much larger than 2 (in the experiments 10^6 was used).

We derive a formalism for transient compartmentalization which considers the statistics of compartmentalization and growth, and illustrate it for competing independently growing species. Functional replicators (in the experiment: ribozymes) can be stabilized by selection and a phase diagram for this ribozyme-parasite case is derived as function of inoculum size and relative growth. By changing the 1D fitness function, we can treat the case of cooperation. By adding deterministic mutations to the model, the dimension of the model is increased by one, and it is shown that transient compartments can overcome error catastrophes.

In growing from of a small number of replicators to a large population, noise in the rate of replication is exponentially amplified, which can lead to giant fluctuations in the final population composition. Using the theory of branching processes, we derive the noise in population composition for competing replicators as a function of replication rates and polymer lengths. Given the harsh selection on composition that we can impose in transient compartmentalization, considerable fitness gains can be realized in reducing compositional noise. Small-molecule autocatalytic networks, which typically have around one rate-limiting step, would suffer large fluctuations. Polymers, on the other hand, can strongly reduce their compositional noise: the successive incorporation of monomers can lead to a highly peaked waiting time distribution.

When a replicase copies itself, but also parasites, things become more delicate. If polymerization is rate-limiting, then the formation of a replicase-parasite complex prevents the replicase involved from copying other replicases. The theory of branching processes can no longer be used here, because the populations are interacting and can thus not be treated as independent in the renewal theorem. However, by looking at simulations and the possible replication trajectories, it becomes clear that the typical treatment with 2-species mass-action models becomes highly inappropriate: such models do not take into account that molecules are sequestered in complexes. In a more detailed treatment, parasites become dramatically more dangerous: they inherently copy faster (even if they have the same length as the replicase) and if they accumulate they can quickly occupy all available replicases. In such a situation, a freshly released replicase will rapidly encounter

another parasite to form a complex with (replication is rate-limiting, so complexation is relatively quick), and encountering another free replicase becomes an increasingly rare event. We refer to such a takeover as a 'complexation catastrophe'.

In **Chapter 9**, a new scenario is formulated, based on the results of the previous chapters. The purpose if the scenario is to be provocative, we highlight that there is still plenty of room for new ideas and scenarios.

It has been suggested that abiogenesis may be a process for which our biochemistry is only one particular outcome. Our scenario is tailored to this philosophy and speculates on the mechanisms that come into play in such a process.

The scenario considers an out-of-equilibrium multicompartment chemical network with transport barriers. Autocatalytic evolution triggers permanent modification of transport and chemistry, modifying the self-sorting and leading to new network-level innovations (like dynamic kinetic resolution or various forms of proofreading).

By construction, the scenario becomes increasingly far-fetched: as we move beyond single molecules to higher-order structures, a lot of things can happen. It will be interesting to see how these advanced stages can be treated with rigor.

The main point remains unchanged: we argue that a variety of autocatalytic mechanisms, among which multicompartment autocatalysis, provided an ecological chemical evolution, that started assembling new compartments and modifying the environment. This in turn promoted multilevel selection of molecular populations and compartments with particular contents (e.g. error-correcting networks). These compartments depended on multicompartment catalytic pathways for their formation and higher-order structures emerged to favor these pathways. Thereby, new layers of selection were introduced, and new selection pressures started to act. This increasingly had to be addressed by flexible, evolvable chemistries, and culminated in an evolutionary arms race towards evolvability.

Bibliography

Articles

[1] E. Willerslev et al. "Long-term persistence of bacterial DNA". In: *Curr. Biol.* 14.1 (2004), pp. 9–10.

[2] G Brent Dalrymple. "The age of the Earth in the twentieth century : a problem (mostly) solved". In: *Geol. Soc.* 190 (2001), pp. 205–221.

[3] F.G. Houtermans. "Das Alter des Urans". In: *Z. Naturforschg.* 2a (1947), pp. 322–328.

[4] Bridgwater D et al. "Microfossil-like objects from the Archean of Greenland: a cautionary note". In: *Nature* 289 (1981), pp. 51–53.

[5] K Sugitani et al. "Early evolution of large micro-organisms with cytological complexity revealed by microanalyses of 3 . 4 Ga organic-walled microfossils". In: *Geobiology* 13 (2015), pp. 507–521.

[6] Christophe Malaterre. "Chemical evolution and life". In: *Bio Web Conf.* 4 (2015), pp. 1–8.

[8] F.H.C. Crick. "The origin of the genetic code". In: *J. Mol. Biol.* 38 (1968), pp. 367–379.

[9] L. E. Orgel. "Evolution of the genetic apparatus: A review". In: *J. Mol. Biol.* 38 (1968), pp. 381–393.

[10] Allan M Maxam and Walter Gilbert. "A new method for sequencing DNA". In: *Proc. Natl. Acad. Sci. U.S.A.* 74.2 (1977), pp. 560–564.

[11] Niles Lehman. "A Case for the Extreme Antiquity of Recombination". In: *J. Mol. Evol.* 56.6 (June 2003), pp. 770–777.

[12] Leslie E. Orgel. “Prebiotic chemistry and the origin of the RNA world”. In: *Crit. Rev. Biochem. Mol. Biol.* 39.2 (2004), pp. 99–123.

[13] Julien Derr et al. “Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences”. In: *Nucleic Acids Res.* 40.10 (2012), pp. 4711–4722.

[14] Gerald F Joyce. “The antiquity of RNA-based evolution”. In: *Nature* 418 (2002), pp. 214–221.

[15] Meng Wu and Paul G Higgs. “The origin of life is a spatially localized stochastic transition”. In: *Biol. Direct* 7.42 (2012), pp. 1–15.

[16] Alexei V Tkachenko and Sergei Maslov. “Spontaneous emergence of autocatalytic information-coding polymers”. In: *J. Chem. Phys.* 143.2015 (2015), p. 045102.

[17] Helen Greenwood Hansma. “Possible origin of life between mica sheets: Does life imitate mica?” In: *J. Biomol. Struct. Dyn.* 31.8 (2013), pp. 888–895.

[18] Christof B Mast et al. “Escalation of polymerization in a thermal gradient”. In: *Proc. Natl. Acad. Sci. U.S.A.* 110.20 (2013), pp. 8030–5.

[19] Alexander V. Vlassov et al. “The RNA world on ice: A new scenario for the emergence of RNA information”. In: *J. Mol. Evol.* 61.2 (2005), pp. 264–273.

[20] Christopher P. Lepper et al. “Effects of Pressure and pH on the Hydrolysis of Cytosine: Implications for Nucleotide Stability around Deep-Sea Black Smokers”. In: *ChemBioChem* 19.6 (2018), pp. 540–544.

[23] Ben K D Pearce et al. “Origin of the RNA world : The fate of nucleobases in warm little ponds”. In: *Proc. Natl. Acad. Sci. U.S.A.* 114.43 (2017), pp. 11327–11332.

[24] Xi Chen, Na Li, and Andrew D Ellington. “Ribozyme Catalysis of Metabolism in the RNA World”. In: *Chem. Biodivers.* 4 (2007), pp. 633–655.

[25] Jack W Szostak. “The eightfold path to non-enzymatic RNA replication”. In: *J. Syst. Chem.* 3.1 (2012), pp. 2–14.

[26] James Attwater et al. “Ribozyme-catalysed RNA synthesis using triplet building blocks”. In: *Elife* 7 (2018), p. 35255.

[27] Nilesh Vaidya et al. “Spontaneous network formation among cooperative RNA replicators”. In: *Nature* 491.7422 (Oct. 2012), pp. 72–77.

[28] Günter Wächtershäuser. “Before Enzymes and Templates: Theory of Surface Metabolism”. In: *Microbiol. Rev.* 52.4 (1988), pp. 452–484.

[29] G. Wachtershauser. “Evolution of the first metabolic cycles”. In: *Proc. Natl. Acad. Sci. U.S.A.* 87.1 (1990), pp. 200–204.

[30] Günter Wächtershäuser. “Groundworks for an evolutionary biochemistry: The iron-sulphur world”. In: *Prog. Biophys. Mol. Biol.* 58.2 (1992), pp. 85–201.

[31] Claudia Huber et al. “A Possible Primordial Peptide Cycle”. In: *Science (80-.).* 301.August (2003), pp. 938–940.

[32] Leslie E Orgel. “The implausibility of Metabolic Cycles on the prebiotic earth”. In: *Plos Biol.* 6.1 (2008), pp. 5–13.

[33] Thomas Geisberger et al. “Evolutionary Steps in the Analytics of Primordial Metabolic Evolution”. In: *Life* 9.2 (2019), p. 50.

[34] Sara Forcisi et al. “Liquid chromatography-mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling”. In: *J. Chromatogr. A* 1292 (2013), pp. 51–65.

[35] Victor Sojo et al. “The Origin of Life in Alkaline Hydrothermal Vents”. In: *Astrobiology* 16.2 (2016), pp. 181–197.

[36] G. Wachtershauser. “Origin of life: Life as we don’t know it”. In: *Science (80-.).* 289.5483 (2000), pp. 1307–1308.

[37] Ramanarayanan Krishnamurthy. “Life’s Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry?” In: *Chem. Eur. J.* 24 (2018), p. 16708.

[41] Daniel Segré et al. “Graded autocatalysis replication domain (GARD): kinetic analysis of self-replication in mutually catalytic sets”. In: *Orig. Life Evol. Biosph.* 28 (1998), pp. 501–514.

[42] Daniel Segré, Dafna Ben-eli, and Doron Lancet. “Compositional genomes : Prebiotic information transfer in mutually catalytic noncovalent assemblies”. In: *Proc. Natl. Acad. Sci. U.S.A.* 97.8 (2000), pp. 4112–4117.

[43] E Szathmáry. “The origin of replicators and reproducors”. In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361 (2006), pp. 1761–1776.

[44] Vera Vasas, Eörs Szathmáry, and Mauro Santos. “Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life”. In: *Proc. Natl. Acad. Sci. U.S.A.* 107.4 (2010), pp. 1470–1475.

[45] D Lancet, R Zidovetzki, and O Markovitch. “Systems protobiology: origin of life in lipid catalytic networks”. In: *J.R.Soc.Interface* 15 (2018), p. 20180159.

[47] Vera Vasas et al. “Evolution before genes”. In: *Biol. Direct* 7 (2012), pp. 1–14.

[48] Kepa Ruiz-mirazo, Carlos Briones, and De Escosura. “Prebiotic Systems Chemistry : New Perspectives for the Origins of Life”. In: *Chem. Rev.* 114 (2014), pp. 285–366.

[49] Walter Gilbert. “The RNA world”. In: *Nature* 319 (1986), p. 618.

[50] Karin Moelling and Felix Broecker. “Viruses and evolution - Viruses first? A personal perspective”. In: *Front. Microbiol.* 10.MAR (2019), pp. 1–13.

[51] Stuart A. Kauffman. “Autocatalytic Sets of Proteins”. In: *J. Theor. Biol.* 119 (1986), pp. 1–24.

[52] Peter T.S. van der Gulik and Dave Speijer. “How amino acids and peptides shaped the RNA world”. In: *Life* 5.1 (2015), pp. 230–246.

[54] Peter E. Nielsen et al. “Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide”. In: *Science (80-.).* 254.5037 (1991), pp. 1497–1500.

[55] Guy Ourisson and Yoichi Nakatani. “The terpenoid theory of the origin of cellular life: the evolution of terpenoids to cholesterol”. In: *Chem. Biol.* 1 (1994), pp. 11–23.

[56] H James Cleaves et al. “A Reassessment of Prebiotic Organic Synthesis in Neutral Planetary Atmospheres”. In: *Orig. Life Evol. Biosph.* 38 (2008), pp. 105–115.

[57] Steven A Benner. “Prebiotic plausibility and networks of paradox-resolving independent models”. In: *Nat. Commun.* 9 (2018), p. 5173.

[58] John D Sutherland. “The Origin of Life — Out of the Blue Angewandte”. In: *Angew. Chem. Int. Ed.* 55 (2016), pp. 104–121.

[61] David S. McKay et al. “Search for past life on Mars: Possible relic biogenic activity in martian meteorite ALH84001”. In: *Science (80-.).* 273.5277 (1996), pp. 924–930.

[62] Derek Sears, Ed Scott, and Paul Warren. “The legacy of Allan Hills 84001”. In: *Meteorit. Planet. Sci.* 33 (1998), pp. 545–546.

[63] T. J. Lapen et al. “A younger age for ALH84001 and Its geochemical link to shergottite sources in mars”. In: *Science (80-.).* 328.5976 (2010), pp. 347–351.

[64] Laura M Barge et al. “From Chemical Gardens to Chemobrionics”. In: *Chem. Rev.* 115 (2015), pp. 8652–8703.

[65] Thomas Gold. “The deep, hot biosphere”. In: *Proc. Natl. Acad. Sci. U.S.A.* 89 (1992), pp. 6045–6049.

[66] Daniel R Colman et al. “The deep , hot biosphere : Twenty-five years of retrospection”. In: *Proc. Natl. Acad. Sci. U.S.A.* 114.27 (2017), pp. 6895–6903.

[67] Arsev Umur Aydinoglu and Zehra Taskin. “Origins of Life Research : a Bibliometric Approach”. In: *Orig. Life Evol. Biosph.* 48 (2018), pp. 55–71.

[68] Michael Burnam-fink et al. “Impact of the Arizona NExSS Winter School on Astrobiology Knowledge and Attitudes”. In: *Astrobiology* 18.3 (2018), pp. 365–375.

[69] Robert Pascal, Addy Pross, and John D. Sutherland. “Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics”. In: *Open Biol.* 3.NOV (2013), pp. 1–9.

[70] Maël Montévil and Matteo Mossio. “Biological organisation as closure of constraints”. In: *J. Theor. Biol.* 372 (2015), pp. 179–191.

[72] Giulio Ragazzon and Leonard J. Prins. “Energy consumption in chemical fuel-driven self-assembly”. In: *Nat. Nanotechnol.* 13.10 (2018), pp. 882–889.

[73] “Self-assembling life”. In: *Nat. Nanotechnol.* 11 (2016), p. 909.

[74] Emanuele Penocchio and Massimiliano Esposito. “Thermodynamic ef fi ciency in dissipative chemistry”. In: *Nat. Commun.* 10 (2019), p. 3865.

[76] S Ciliberto. “Experiments in Stochastic Thermodynamics : Short History and Perspectives”. In: *Phys. Rev. X* 7 (2017), p. 021051.

[77] Alex Blokhuis and David Lacoste. “Length and sequence relaxation of copolymers under recombination reactions Length and sequence relaxation of copolymers under recombination reactions”. In: *J. Chem. Phys.* 147 (2017), p. 094905.

[78] Ryo Mizuuchi et al. “Mineral surfaces select for longer RNA molecules”. In: *Chem. Commun.* 55 (2019), pp. 2090–2093.

[79] A Blokhuis et al. “Selection Dynamics in Transient Compartmentalization”. In: *Phys. Rev. Lett.* 120.15 (Apr. 2018), p. 158101.

[80] Alex Blokhuis et al. “The generality of transient compartmentalization and its associated error thresholds”. In: *BioRxiv* (2018). arXiv: `521211 [10.1101]`.

Books

[7] C. Woese. *The Genetic Code, the Molecular Basis for Genetic Expression*. New York: Harper and Row, 1967.

[21] A I Oparin. *Origin of Life*. Dover, 1952.

[38] Harold J. Morowitz. *Beginnings of cellular life: Metabolism recapitulates biogenesis*. London: Yale University Press, 1992.

[39] Eric Smith and Harold Morowitz. *The origin and nature of life on earth: The Emergence of the Fourth Geosphere*. Cambridge, UK: Cambridge University Press, 2016.

[40] Freeman Dyson. *Origins of Life Revised Edition*. Cambridge, UK: Cambridge University Press, 1999.

[53] A.G. Cairns-Smith. *Seven Clues To The Origin of life: a scientific detective story*. Cambridge, UK: Cambridge University Press, 1985.

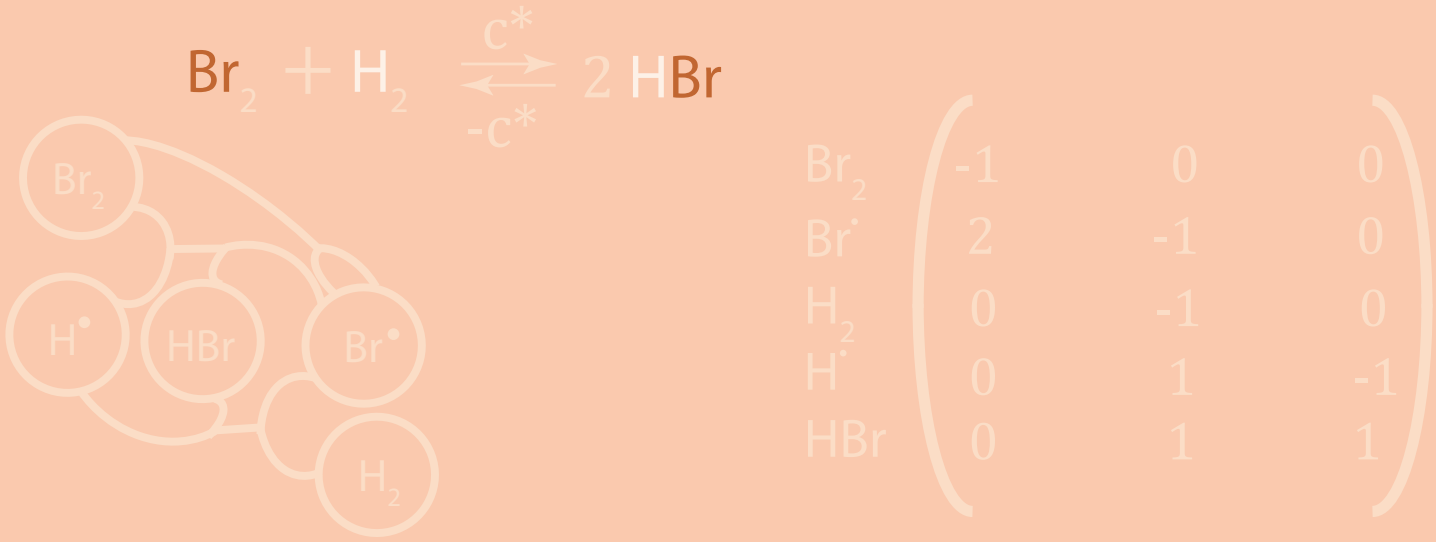
[59] Thomas S. Kuhn. *The structure of scientific revolutions*. The University of Chicago Press, 1962.

[60] Stephen Jay Gould. *The Structure of Evolutionary Theory*. Cambridge, Massachusetts: Harvard University Press, 2002.

[71] Stuart A. Kauffman. *A world beyond physics: the emergence & evolution of life*. Oxford University Press, 2019.

[75] S. Kjelstrup et al. *Non-Equilibrium Thermodynamics for Engineers*. World Scientific, 2010.

[81] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford: Oxford University Press, 1995.



2. Chemical networks: the Stoichiometric matrix

In this chapter, the stoichiometric matrix formalism for chemical networks will be introduced and placed in a broader context. This approach was pioneered in the 60s[1, 2]. The conventions and terminology will largely follow those introduced in recent years by the Esposito group, that is building towards a general overarching framework of nonequilibrium thermodynamics and stochastic thermodynamics. For an excellent introduction to the formalism and its application to nonequilibrium thermodynamics in chemical networks, see Refs [3] and [4]. For its implementation in stochastic thermodynamics, see Ref[5]. We also refer to M. Feinberg's impressive compendium (Ref.[2]) of over 50 years of work on chemical reaction networks, offering a broad perspective on its foundations and the particular networks seen in biochemistry (see also [6]).

It is hard to overstate the importance of making the study of chemical networks systematic and general and many results in upcoming chapters are a consequence of the availability of this framework. In this Chapter, we will also introduce some extensions of our own. A publication is in preparation based on the contents of these extensions.

In Sec. 2.2, we will show that any chemical network can be described in sufficient detail, using reactions that are at most bimolecular, to yield a stoichiometric matrix that is unique. In Sec. 2.3.2 we discuss the description of multicompartment systems and diffusion, which points to the need for treating particle numbers instead of concentrations, even for macroscopic systems. In Sec. 2.4, we introduce the notion of a reaction vector and its reaction coefficients. We provide a distinction between mass-like conservation laws and mixed conservation laws. In Sec. 2.5.2 we will introduce the powerful tool of submatrices to the framework. Implicitly, such a tool is already in use for chemostatting, in which reactants are removed from the description of the system. The submatrix approach also allows to remove reactions, which will prove to be an essential tool in upcoming chapters.

In providing these extensions, we also cover their interpretation for the strongly related framework by the Feinberg group, where reacting species are bundled in 'chemical complexes' (See also 'Foundations of Chemical Reaction Network Theory' [2]).

2.1 The Stoichiometric Matrix

The stoichiometric matrix ν describes the stoichiometry of all chemical reactions in a reaction network[1]. The sign of ν_{ki} indicates whether a species k is consumed (-) or produced (+) by performing reaction i in forward direction. Its absolute magnitude ν_{ki} provides the stoichiometric factor, such that we have a net reaction

$$\sum_k \nu_{ki}^{(+)} \mathrm{Z_k} \rightleftharpoons \sum_k \nu_{ki}^{(-)} \mathrm{Z_k}. \tag{2.1}$$

where the stoichiometric coefficient ν_{ki} is given by a consumption matrix $\nu_{ki}^{(-)}$ and a production matrix $\nu_{ki}^{(+)}$

$$\nu_{ki} = \nu_{ki}^{(-)} - \nu_{ki}^{(+)}. \tag{2.2}$$

As an illustration, let us look at a triple of simple bimolecular reactions: i) the formation of an imine and water from ammonia with an aldehyde and ii) acid-base reaction between ammonia and water to form ammonium ion and hydroxide ion iii) acid-base reaction between two molecules of water to form hydroxide and hydronium ions

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{CH_3CHO} \\ \mathrm{NH_3} \\ \mathrm{CH_3CHNH} \\ \mathrm{H_2O} \\ \mathrm{NH_4^+} \\ \mathrm{OH^-} \\ \mathrm{H_3O^+} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ -1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & -2 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{2.3}$$

$$\begin{aligned} \mathrm{CH_3CHO + NH_3} &\overset{1}{\rightleftharpoons} \mathrm{CH_3CHNH + H_2O} \\ \mathrm{NH_3 + H_2O} &\overset{2}{\rightleftharpoons} \mathrm{NH_4^+ + OH^-} \\ \mathrm{2\,H_2O} &\overset{3}{\rightleftharpoons} \mathrm{OH^- + H_3O^+} \end{aligned}$$

It is often instructive to draw reaction networks as graphs. This can e.g. be in the form of a directed hypergraph, where reactants are nodes linked by directed edges linked to multiple nodes. Often, one uses arrows in directed hypergraphs to link two sets of nodes and explicitly describe a preferred direction (see Sec. 3.2.3 for examples of such hypergraphs).

Here, we do not wish to imply that a reaction proceeds occurs in only one direction, and instead separate the links between nodes by a single piece of line, as shown in Fig. 2.1a. Such a hypergraph is consistent with any particular choice of a reaction direction taken in the stoichiometric matrix, it is a 'coarse-grained' representation of them. By adding arrows, a particular choice of signs for $\boldsymbol{\nu}$ can be made explicit.

In Fig. 2.1a, a representation of the hypergraph for network (2.4) is provided. Another common practice is the use of bipartite graphs, with one type of nodes corresponding to reactions, and another type to reactants, which are linked by edges that indicate their participation. It is a particular instance of a petri net. These graphs can be constructed from the same stoichiometric matrix $\boldsymbol{\nu}$ and are often equivalent to the hypergraph. However, if a stoichiometric coefficient of 2 or more appears in $\boldsymbol{\nu}$, this information remains implicit in a directed graph, as can be seen for reaction 3 in Fig. 2.1b. To make this information explicit, one can replace the directed graph by a directed multigraph, in which pairs of nodes can be linked by more than one edge.

2.1.1 The incidence matrix

Another common way to draw a reaction network [2] is to draw a simple graph with nodes corresponding to pairs of reactants or products, called complexes. Such a simple graph is shown in

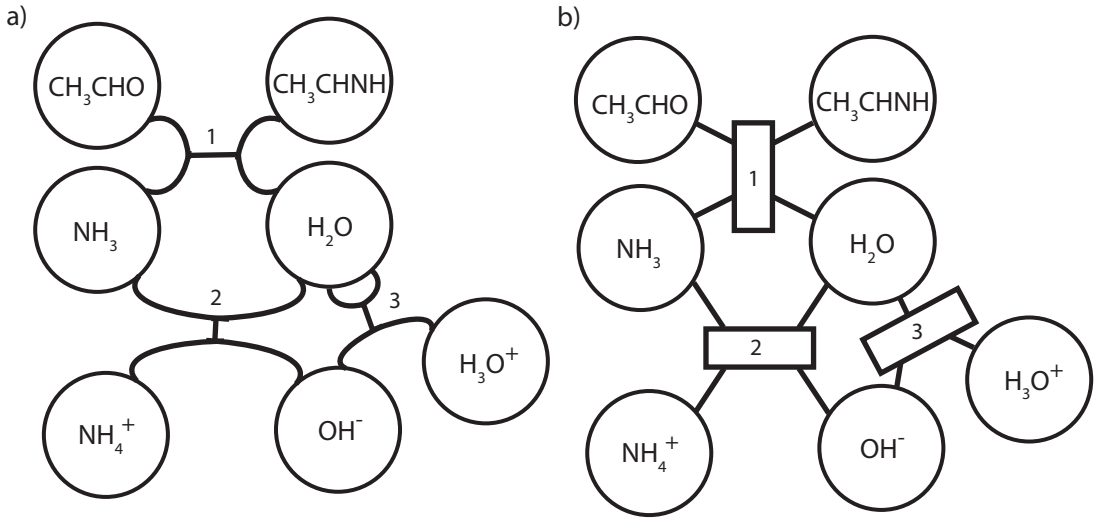


Figure 2.1: a) 'coarse-grained' directed hypergraph representation. Chemical species occurring on the same side of the reaction are joined on the same side of the edge. A reactant that occurs twice is joined by a fork-shaped edge (2 H_2O in reaction 3). b) bipartite simple graph representation. Square nodes represent a reaction. Circular nodes represent chemical species. Species that react together join the reaction node on the same side. Numbers in both graphs correspond to reactions as enumerated in the stoichiometric matrix.

Fig. 2.2 for network (2.4). The latter reposes on a different framework, with transitions between complexes encoded in an incidence matrix $\boldsymbol{\partial}$.

$$\begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} \mathrm{CH_3CHO{+}NH_3} \\ \mathrm{CH_3CHNH{+}H_2O} \\ \mathrm{H_2O{+}NH_3} \\ \mathrm{OH^-{+}NH_4^+} \\ \mathrm{H_2O{+}H_2O} \\ \mathrm{OH^-{+}H_3O^+} \end{array} & \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{2.4}$$

$$\boldsymbol{\partial} = \begin{array}{c} \mathrm{CH_3CHO{+}NH_3} \\ \mathrm{CH_3CHNH{+}H_2O} \\ \mathrm{H_2O{+}NH_3} \\ \mathrm{OH^-{+}NH_4^+} \\ \mathrm{H_2O{+}H_2O} \\ \mathrm{OH^-{+}H_3O^+} \end{array} \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.5}$$

This incidence matrix can directly be obtained from the stoichiometric matrix [7], but the converse is not true: the stoichiometric matrix $\boldsymbol{\nu}$ contains more detailed information than the incidence matrix $\boldsymbol{\partial}$. It is only because explicit labels have been provided for the nodes that we can cross-identify our examples. Polettini et al make this cross-identification explicit, by the relation

$$\boldsymbol{\nu} = \frac{\partial Y}{\partial X}\boldsymbol{\partial}. \tag{2.6}$$

Here, $\left(\frac{\partial Y}{\partial X}\right)_{ij}$ quantifies the stoichiometric contribution of species $\mathrm{X_i}$ in complex $\mathrm{Y_j}$.

If we would not have specified the content of these complexes, the incidence matrix given by $\boldsymbol{\partial}$ in Eq. (2.5) could equally well represent the following set of unimolecular reactions

$$\mathrm{A} \overset{1}{\rightleftharpoons} \mathrm{B}, \quad \mathrm{C} \overset{2}{\rightleftharpoons} \mathrm{D}, \quad \mathrm{E} \overset{3}{\rightleftharpoons} \mathrm{F}, \tag{2.7}$$

as well as many other combinations. The shift to an elegant description in terms of simple graphs comes at the cost of a reduction in explicit detail.

The complexes used as nodes in Fig. 2.2 should not be confused with a complex in the chemical sense, which refers to a single species, often denoted within brackets. Such species

Figure 2.2: Reaction network as a simple graph linking complexes.

form the cornerstone of coordination chemistry, with metal-ligand complexes like $[Cu(NH_3)_6]^{2+}$, but they are also found in many other branches of chemistry. Chemical complexes need not be stable isolable species. They can be theoretical, transient species, such as an encounter complex $[AB]$, $A+B \rightleftharpoons [AB]$, or activated complex close to or at a transition state $OH^- + CH_2BrCH_3 \rightarrow [HO-CH_2CH_3-Br]^- \rightarrow Br^- + CH_2(OH)CH_3$. In the remainder of this text, a complex will refer to a complex in the chemical sense, unless strictly indicated otherwise.

The most suitable choice for drawing a reaction network as a graph depends on context and network size. Hypergraphs rapidly become hard to read due to their large number of interconnected nodes and overlapping edges when drawn on a flat plane.

Complexes, on the other hand, yield disjoint simple graphs or connected components, e.g. in Fig. 2.2 there are three connected components. Such networks are considerably easier to draw. The properties we will investigate in this text mostly require a description of networks with more detail than such graphs permit. Consequently, we will require the use of a stoichiometric matrix $\boldsymbol{\nu}$, illustrated by hypergraphs when it is instructive to highlight small motifs.

2.2 Conventions

In the chemical networks that will be discussed in upcoming sections, we will at all times assume that reactions are 'reversible' in the chemical sense: a well-defined reverse reaction exists, that is the exact opposite of the forward reaction in all molecular populations.

In the mathematical literature, a network in which all reverse reactions exist is referred to as a 'reversible chemical network'. Whether a reaction does appreciably occur in both directions is a question of kinetics and thermodynamics. To have a complete description of entropy production, it is necessary to describe all reverse reactions, no matter how rare they are. To monitor species concentrations, however, such rare reactions become redundant. Consequently, we may use the notation

$$A \rightarrow B, \tag{2.8}$$

to denote a reaction that has a well-defined reverse reaction in the thermodynamic sense, but which occurs rarely enough to neglect it in its kinetic description.

The stoichiometric matrix and Chemical Reaction Networks can be extended in various ways that may no more pertain to chemistry, but capture similar constraints of transformations in a system. The input-output model in economics uses an object sometimes referred to as a Leontief* matrix[8], which captures the interdependence of industrial production. In stochastic thermodynamics, the formalism has been generalized by having a matrix characterizes the exchange of other conserved quantities such as energy between discrete levels [9].

For our purposes, it will be important to respect the constraints set by chemistry, and describe that chemistry in sufficient detail. In doing so, useful properties emerge that are general to chemistry. In principle, one could relax this structure to non-chemical situations, e.g. by using noninteger stoichiometries or by not conserving atoms. This generality comes at a cost: we lose some of the defining features of chemistry.

*Leontief was awarded the Nobel Prize in economics for his work on this approach in 1973.

To provide a clear illustration of what the structural constraints in chemistry lead to, we will need a description that is sufficiently precise to exploit it. As will become more clear as we progress, this requires us to have conventions that make the framework unambiguous and universal. To this end, we will introduce two conventions that our reaction networks will respect: i) reaction steps involve at most two species, ii) every chemical species involved in a reaction does so either as a reactant or as a product (nonambiguity).

2.2.1 Unimolecular and bimolecular reactions

Our first convention states that at most two species react in either sense. This means that the stoichiometric coefficients $\nu_{ki}^{(+)}, \nu_{ki}^{(-)}$ of reaction i can sum to at most two species

$$\forall i \quad \sum_k \nu_{ki}^{(+)} \leq 2, \sum_k \nu_{ki}^{(-)} \leq 2. \tag{2.9}$$

Physically, this corresponds to the fact that one can always decompose a reaction in many substeps involving at most two reactants[10, 11]. True 'termolecular' collisions are rare, and a termolecular transition state can be considered as arising from sequential bimolecular steps to yield the same dynamics.

As an example, consider the formation of a bromonium from an alkene, whose kinetics sometimes correspond to a termolecular reaction[12, 13]

$$\mathrm{Br_2} + \mathrm{R_1R_2C{=}CR_3R_4} \overset{\mathrm{CCl_4}}{\rightleftharpoons} \mathrm{R_1R_2CBrCBrR_3R_4} \tag{2.10}$$

$$\frac{d[\mathrm{R_1R_2CBrCBrR_3R_4}]}{dt} = k[\mathrm{Br_2}]^2[\mathrm{R_1R_2C{=}CR_3R_4}] \tag{2.11}$$

Intuitively, such a rate law could be thought of as a 'three-body' collision process. However, the same rate law can be built up from a series of bimolecular reactions, as is observed by detailed kinetic studies. A reaction mechanism for the latter situation is given in Fig. 2.3

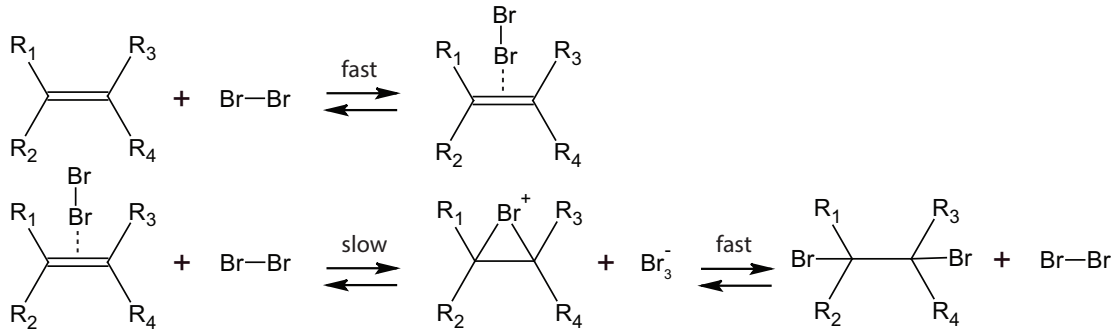


Figure 2.3: An example of a termolecular process, composed of biomolecular substeps.

Since the first step is fast, we can assume local equilibrium, to write

$$[\mathrm{R_1R_2C{=}CR_3R_4 \cdot Br_2}] = K[\mathrm{R_1R_2C{=}CR_3R_4}][\mathrm{Br_2}], \tag{2.12}$$

Where K is an equilibrium constant. In turn, this can be injected in the rate equation Eq. (2.11) to give

$$\frac{d[\mathrm{R_1R_2CBrCBrR_3R_4}]}{dt} = k'[\mathrm{Br_2}][\mathrm{R_1R_2C{=}CR_3R_4 \cdot Br_2}] = k'K[\mathrm{Br_2}]^2[\mathrm{R_1R_2C{=}CR_3R_4}]. \tag{2.13}$$

Which means the termolecular rate constant can be written in terms of a bimolecular rate constant and an equilibrium constant $k = k'K$.

In practice, one can always decompose higher order reactions in unimolecular and bimolecular steps with an appropriate separation of timescales. If one wishes to do this in a rigorous manner to describe genuine chemical species, detailed mechanistic and numerical studies may be required. Such studies have recently revealed an important new class of termolecular reactions, which are important in e.g. combustion [14]. In these reactions, a collision complex is formed by two molecules, which subsequently encounters a radical species. The IUPAC gold book[†] refers to a collision complex as [15]: *An ensemble formed by two reaction partners for which the distance is the sum of their van der Waals radii. As such it constitutes a subclass of the species indicated as encounter complex*. New techniques to uncover reaction mechanisms are still being elaborated, e.g. by monitoring the response to small temperature oscillations[16] .

In general, we can construct reaction networks from second-order reactions involving complexes to reproduce the kinetics of higher order reactions. In the end, experiments will tell what the most realistic choice for those reactions is.

2.2.2 Nonambiguity

It is important to underline that a stoichiometric matrix $\boldsymbol{\nu}$ can represent different reaction networks, since Eq. (2.2) admits multiple solutions for $\boldsymbol{\nu}^{(+)}$ and $\boldsymbol{\nu}^{(-)}$, even when constrained by Eq. (2.9). In upcoming sections, it will be very useful if a reaction network is uniquely defined by $\boldsymbol{\nu}$.

Definition 2.2.1 — Nonambiguity. A stoichiometric matrix $\boldsymbol{\nu}$ respects **nonambiguity**, if for each reaction i the equation

$$\nu_{ki} = \nu_{ki}^{(-)} - \nu_{ki}^{(+)} \tag{2.14}$$

has one unique solution for $\nu_{ki}^{(-)}, \nu_{ki}^{(+)}$, which means

$$\forall k,i \; \nu_{ki}^{(-)} \nu_{ki}^{(+)} = 0. \tag{2.15}$$

Without loss of generality, the ambiguity in Eq. (2.2) can be removed, by introducing the condition given in Eq. (2.15), which implies that every chemical species involved in a reaction does so either as a reactant or as a product. With this constraint we can still represent every chemical reaction network we could represent before, but some reactions need to be described in more detailed substeps to verify Eq. (2.15).

Situations that generally require modification due to Eq. (2.15) are simplified equations for autocatalysis, catalysis and chain reactions. Let us first consider a simple catalytic reaction between an enzyme E and a substrate S, to yield a product P

$$\mathrm{S} + \mathrm{E} \rightleftharpoons \mathrm{P} + \mathrm{E}. \tag{2.16}$$

For enzyme catalysis, we can write an enzyme-substrate complex ES and enzyme-product complex EP, to observe a typical enzymatic reaction network

$$\mathrm{E} + \mathrm{P} \rightleftharpoons \mathrm{ES}, \qquad \mathrm{ES} \rightleftharpoons \mathrm{EP}, \qquad \mathrm{EP} \rightleftharpoons \mathrm{E} + \mathrm{P}. \tag{2.17}$$

A prototypical model equation for autocatalysis has a molecule B convert a ‘food’ molecule A to another copy of B

$$\mathrm{A} + \mathrm{B} \overset{1}{\rightleftharpoons} 2\mathrm{B}. \tag{2.18}$$

[†]The correct and unambiguous use of terminology in chemistry is promoted by the international union of pure and applied chemistry (IUPAC), which publishes a variety of colored reference books. The blue book, for example, contains the IUPAC nomenclature rules for organic compounds. The Gold book is a compendium of chemical terminology, consultable online at https://goldbook.iupac.org/

In the absence of other food or waste, this reaction constitutes a net isomerization from A to B, which in an uncatalyzed fashion would take the form

$$\mathrm{A} \overset{2}{\rightleftharpoons} \mathrm{B}. \tag{2.19}$$

In a stoichiometric matrix $\boldsymbol{\nu}$, both pathways lead to the same entries

$$\begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \boldsymbol{\nu} = \begin{array}{c} \mathrm{A} \\ \mathrm{B} \end{array} & \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \end{array} \quad \begin{array}{c} \mathrm{A}+\mathrm{B} \rightleftharpoons 2\mathrm{B} \\ \mathrm{A} \rightleftharpoons \mathrm{B}. \end{array} \tag{2.20}$$

This is an example of ambiguity due to Eq. (2.2). These pathways can be distinguished, by a decomposition in nonambiguous reaction steps. For autocatalysis, we can extend our description, by including a reaction intermediate, e.g. an encounter complex [AB] or a transition-state complex. We can then write a two-step process

$$\mathrm{A}+\mathrm{B} \rightleftharpoons [\mathrm{AB}], \qquad [\mathrm{AB}] \rightleftharpoons 2\mathrm{B}. \tag{2.21}$$

Physically, such a decomposition makes it possible to distinguish between reaction-limited and diffusion-limited reactions. In addition, it means that no single reaction step is catalyzed or autocatalytic: only collections of reaction steps are. This seemingly small distinction has useful consequences: nonambiguity makes chain reactions and (auto)catalysis an explicit property of the network topology. This will be discussed in detail in the upcoming sections.

2.3 Stoichiometric Matrix and dynamics

2.3.1 Mass-action for a well-mixed reactor

Let us denote by c_k the concentration of a species Z_k, such that

$$c_k = \frac{n_k}{V} \tag{2.22}$$

where n_k is the number of molecules of Z_k and V the reactor volume. According to the mass action law, the reaction rates are proportional to the concentrations of all the species entering in the reaction. It is convenient to make the distinction between the forward and backward reactions so that for a reaction i we can define rates $w_{\pm i}$

$$w_{\pm i} = k_{\pm i} \prod_k \left(\frac{c_k}{c^0} \right)^{\nu_{ki}^{(\pm)}}, \tag{2.23}$$

where $k_{\pm i}$ are the rate constants, $\nu_{ki}^{(\pm)}$ the indices of the stoichiometric matrix as defined in Eq. (2.1) and c^0 the standard concentration of one mole per liter. The net rate of the reaction is given by

$$w_i = w_{+i} - w_{-i}. \tag{2.24}$$

We shall assume all reactions to be reversible, which means that for all i, $k_{\pm i} > 0$, a condition which is needed for a thermodynamically consistent description.

The kinetic rate equations can now be rewritten in matrix form as follows:

$$\frac{d\mathbf{c}}{dt} = \boldsymbol{\nu} \cdot \mathbf{w}, \tag{2.25}$$

2.3.2 Mass-action for communicating multicompartment systems

When a system is composed of subsystems with different volumes (e.g. two coupled, well-mixed reactors), exchange between the subvolumes can equally well be treated as a chemical reaction (an observation that was also made for the complex networks framework[2]), e.g.

$$\mathrm{A^I} \rightleftharpoons \mathrm{A^{II}}, \tag{2.26}$$

which can be modeled as a first-order process with corresponding exchange rates

$$w_+ = k_+ \left(\frac{c_A^{\mathrm{I}}}{c^0} \right), \tag{2.27}$$

$$w_- = k_- \left(\frac{c_A^{\mathrm{II}}}{c^0} \right). \tag{2.28}$$

The rates $w_\pm$ have dimension $[\mathrm{mol}/\mathrm{m}^3 \cdot \mathrm{s}]$. Let us introduce the molecular rates of A transport between I and II, in dimensions $[\mathrm{mol}/\mathrm{s}]$

$$w'_+ = w_+ V^I, \tag{2.29}$$

$$w'_- = w_- V^{II}. \tag{2.30}$$

Denoting $w' = w'_+ - w'_-$, it follows that

$$\frac{dc_{\mathrm{A}}^{\mathrm{I}}}{dt} = -\frac{w'}{V^I} = -\frac{w_+ V^I - w_- V^{II}}{V^I}, \tag{2.31}$$

$$\frac{dc_{\mathrm{A}}^{\mathrm{II}}}{dt} = \frac{w'}{V^{II}} = \frac{w_+ V^I - w_- V^{II}}{V^I}. \tag{2.32}$$

For reactions within a single reactor, such volume contributions cancel, reducing the system to Eq. (2.25). In a multicompartment setup, we can have $V^I \neq V^{II}$, leading to a general form that is not compatible with Eq. (2.25). A similar situation occurs for surface reactions and a number of more complex transport mechanism (e.g. by symporters or antiporters), as well as other multi-compartment or multi-phase situations, e.g. evaporation or crystallization.

For such situations, it is then convenient to replace concentrations c_k with numbers of molecules n_k, and use molecular rates w'_i, so that the defining equations of mass-action become

$$w'_{\pm i} = k'_{\pm i} \prod_k n_k^{\nu_{ki}^{(\pm)}}, \tag{2.33}$$

$$\frac{d\mathbf{n}}{dt} = \boldsymbol{\nu} \cdot \mathbf{w}', \tag{2.34}$$

where $k'_{\pm i}$ are rate constants containing all volume (or surface) dependence. Note that the number of molecules n_k is described in a deterministic continuum limit. For small numbers of molecules, (2.34) needs to be replaced by its master equation[5] analogue.

2.3.3 Diffusion

When reactions are not slow enough with respect to diffusion, the approximation of a well-mixed homogeneous system breaks down. Then, we can model kinetics with a reaction-diffusion equation, which takes the form

$$\frac{d\mathbf{c}(\boldsymbol{x})}{dt} = \boldsymbol{\nu} \cdot \mathbf{w}(\boldsymbol{x}) + \boldsymbol{D} \cdot \Delta \boldsymbol{c}(\boldsymbol{x}), \tag{2.35}$$

where $\boldsymbol{D} = (D_1, ..., D_s)$ is a vector of diffusion constants for the species Z_1 to Z_s, $\boldsymbol{x}$ is a position vector and Δ is the Laplacian operator

$$\Delta c(\boldsymbol{x}) = \sum_{i=1}^{d} \frac{\partial^2 c(\boldsymbol{x})}{\partial x_i^2}, \tag{2.36}$$

with d the dimension of the space. Numerically, calculation of such a Laplacian on a lattice proceeds by taking a stencil of points surrounding $\boldsymbol{x}$. A popular, low order approximation is to take the nearest neighbors, which for 2D ($\boldsymbol{x} = (x,y)$) yields

$$\frac{\partial^2 c(x,y)}{\partial x^2} = \frac{c(x+h,y) + c(x-h,y) - 2c(x,y)}{h^2} + \frac{1}{12}\frac{\partial^4 c(x,y)}{\partial x^4} h^2 + O(h^4), \tag{2.37}$$

$$\frac{\partial^2 c(x,y)}{\partial y^2} = \frac{c(x,y+h) + c(x,y-h) - 2c(x,y)}{h^2} + \frac{1}{12}\frac{\partial^4 c(x,y)}{\partial y^4} h^2 + O(h^4). \tag{2.38}$$

$$\tag{2.39}$$

Provided the higher-order gradients are small and h is chosen adequately, the Laplacian is well-approximated by its nearest neighbors. For each dimension, the first order term can be decomposed in two difference terms corresponding to unimolecular reactions like (2.26), since

$$\frac{c(x+h,y) + c(x-h,y) - 2c(x,y)}{h^2} = \frac{c(x+h,y) - c(x,y)}{h^2} + \frac{c(x-h,y) - c(x,y)}{h^2}. \tag{2.40}$$

If we label each reactant with its local discretized box location, e.g. $A_{x,y}$, we can model diffusion by local reactions of the form

$$A_{x,y} \rightleftharpoons A_{x+h,y} \tag{2.41}$$

with molecular rates

$$w'_{+} = k'_{+} n_{A_{x,y}} = \frac{D}{h^2} n_{A_{x,y}}, \tag{2.42}$$

$$w'_{-} = k'_{-} n_{A_{x+h,y}} = \frac{D}{h^2} n_{A_{x+h,y}}. \tag{2.43}$$

For an N by N grid, and a chemistry involving s species, we can consequently construct an extended stoichiometric matrix, $\boldsymbol{\nu}_*$. While there are only s chemically distinct compounds, we have now labelled them by the box in the grid they occupy, $\boldsymbol{\nu}_*$ contains $N^2 s$ species. For periodic boundary conditions, we have rN^2 chemical reactions (r per box) and $2dN^2 s$ transport reactions. The reaction-diffusion equation then simplifies to

$$\frac{d\mathbf{c}'}{dt} = \boldsymbol{\nu}_* \cdot \mathbf{w} \tag{2.44}$$

where $\boldsymbol{c'} = (..., \boldsymbol{c}_{x,y}, \boldsymbol{c}_{x+h,y}, ...)^T$ is a concentration vector containing the concentrations of each discretized box. This generalizes what was found in the multicompartment case: a coarse-grained diffusion process can reasonably be treated on the same footing as a chemical process, dynamically and topologically. In a master equation framework, treating reactions and diffusion on the same footing is a very natural choice. In Refs. [17, 18, 19], the deterministic and master equation approaches are combined and applied to the Schlögl[20] reaction (discussed in more detail in Sec. 5.5) for the study of the pattern formation.

2.4 Properties and operations on reaction networks

By putting the stoichiometry of chemical reactions in a matrix, a number of important network properties are unveiled, such as chemical cycles and conservation laws. Even more can be learned, by considering submatrices of the network. Upon removing rows (species), affinities and broken conservation laws[3] from contact with reservoirs are found. Upon removing reactions as well, network submotifs can be made explicit. Doing so will reveal which structures are fundamentally responsible for catalysis, chain-reactions and autocatalysis (see Sec. 5.3).

2.4.1 Reaction vectors

Definition 2.4.1 — Reaction vector. We will define a reaction vector $\boldsymbol{g}$ as linear combination of reactions such that

$$\boldsymbol{g} = (g_1, g_2, ..., g_r). \tag{2.45}$$

The change in the number of molecules of a chemical species $\mathrm{Z_i}$ resulting from such a combination of reactions is

$$\Delta n_i = (\boldsymbol{\nu} \cdot \boldsymbol{g})_i \,. \tag{2.46}$$

Given a reaction vector $\boldsymbol{g}$, the corresponding net chemical reaction takes the following form:

$$\sum_k n_k^{(+)} \mathrm{Z_k} \rightleftharpoons_{-\boldsymbol{g}}^{\boldsymbol{g}} \sum_k n_k^{(-)} \mathrm{Z_k}, \tag{2.47}$$

where the vector $\boldsymbol{g}$ (resp. $-\boldsymbol{g}$) is used to explicitly denote the forward (resp. backward) direction.

To perform a single reaction i, $\boldsymbol{g}$ corresponds to $\hat{\boldsymbol{e}}_i$, where $\hat{\boldsymbol{e}}_i$ is the ith unit vector:

$$\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j = \delta_{i,j}, \tag{2.48}$$

with the Kronecker delta $\delta_{i,j}$ defined by

$$\delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2.49}$$

For example, taking the autocatalytic reaction 2 in Eq. (2.21), we can write

$$\begin{pmatrix} \Delta n_\mathrm{A} \\ \Delta n_\mathrm{B} \end{pmatrix} = \boldsymbol{\nu} \cdot \hat{\boldsymbol{e}}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{2.50}$$

2.4.2 Overall reaction coefficients

We will now outline how the values of $n_k^{(\pm)}$ in the reaction balance can be extracted from $\boldsymbol{\nu}$.

For a reaction i proceeding forward g_i times, we have $\boldsymbol{g} = g_i \hat{\boldsymbol{e}}_i$. The coefficient of the net chemical reaction (Eq. (2.47)) then verifies $n_k^{(\pm)}(\boldsymbol{g}) = (\boldsymbol{\nu}^{(\pm)} \cdot g_i \hat{\boldsymbol{e}}_i)_k$. Had we performed a tranformation in the reverse sense, such that $\boldsymbol{g'} = g'_i \hat{\boldsymbol{e}}_i = -\boldsymbol{g}$, the coefficient of the reverse reaction would have to be used:

$$n_k^{(\pm)}(\boldsymbol{g'}) = (\boldsymbol{\nu}^{(\mp)} \cdot g_i \hat{\boldsymbol{e}}_i)_k = (\boldsymbol{\nu}^{(\mp)} \cdot |g'_i| \hat{\boldsymbol{e}}_i)_k \tag{2.51}$$

which more generally can be written as

$$n_k^{(\pm)}(\boldsymbol{g'}) = (\boldsymbol{\nu}^{(\pm \mathrm{sgn}(\mathrm{g'_i}))} \cdot |g'_i| \hat{\boldsymbol{e}}_i)_k. \tag{2.52}$$

where

$$\mathrm{sgn(x)} = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \\ 0 & x = 0 \end{cases} \tag{2.53}$$

A coefficient for a more general reaction vector $\boldsymbol{g} = \sum_i g_i \hat{\boldsymbol{e}}_i$ is then found by summing the components of every reaction following Eq. (2.52)

$$n_k^{\pm} = \sum_i (\boldsymbol{\nu}^{(\pm \mathrm{sgn(g_i)})} \cdot |g_i| \hat{\boldsymbol{e}}_i)_k. \tag{2.54}$$

If $\forall i \; g_i \geq 0$, this expression simplifies to

$$n_k^{\pm} = (\boldsymbol{\nu}^{(\pm)} \cdot \boldsymbol{g})_k. \tag{2.55}$$

In a reaction network where all reactions are reversible, our choice of the 'forward' and 'backward' reaction is arbitrary: exchanging the indices $\nu_{ki}^{(+)}, \nu_{ki}^{(-)}$ for all species k for a reaction i yields a modified matrix $\boldsymbol{\nu}_*$, that still encodes the exact same reaction network.

By the appropriate initial choice of directions for reactions, we can always construct a $\boldsymbol{\nu}_\diamond$ such that $\forall i \; g_{\diamond i} \geq 0$. To see this, we can construct a transformation for the pair $\boldsymbol{g}$, $\boldsymbol{\nu}$ for which Eq. (2.55) is true. We define

$$\begin{matrix} \boldsymbol{g}_\diamond = \boldsymbol{P} \cdot \boldsymbol{g} \\ \boldsymbol{\nu}_\diamond = \boldsymbol{\nu P} \end{matrix}, \quad P_{ij} = \begin{cases} 0 & i \neq j \\ \mathrm{sgn}(g_i) & i = j \end{cases}, \tag{2.56}$$

which guarantees $g_{\diamond i} \geq 0$. By taking $\boldsymbol{\nu}_\diamond = \boldsymbol{\nu P}$, we have $\boldsymbol{\nu}_\diamond \cdot \boldsymbol{g}_\diamond = \boldsymbol{\nu I} \cdot \boldsymbol{g}$, where we used that $\boldsymbol{P}^2 = \boldsymbol{I}$ with $\boldsymbol{I}$ the identity matrix. Therefore, this particular choice of the stoichiometric matrix verifies

$$n_k^{\pm} = (\boldsymbol{\nu}_\diamond^{(\pm)} \cdot \boldsymbol{g}_\diamond)_k. \tag{2.57}$$

This property will allow to make some upcoming expressions considerably more elegant. In upcoming examples, we will choose the appropriate $\boldsymbol{\nu}_\diamond$ as a starting point, making this step implicit.

∅-reactions

If a reaction i has no reactant we have $\sum_k \left(\boldsymbol{\nu}^{(+)} \cdot \hat{\boldsymbol{e}}_i\right)_k = 0$. Similarly, if this reaction has no product, we can write $\sum_i \left(\boldsymbol{\nu}^{(-)} \cdot \hat{\boldsymbol{e}}_i\right)_k = 0$. Physically, such reactions can correspond to an influx (resp. dilution) in e.g. a CSTR. E.g. we can consider

$$\emptyset \rightleftharpoons^{1} \mathrm{A}, \quad \mathrm{A} \rightleftharpoons^{2} \mathrm{B}, \quad \mathrm{B} \rightleftharpoons^{3} \emptyset. \tag{2.58}$$

where the empty set symbol ($\emptyset$) denotes the absence of chemical species. ∅-reactions are easily identified, since their corresponding columns are zero vectors in $\boldsymbol{\nu}^{(+)}$ or $\boldsymbol{\nu}^{(-)}$

$$\boldsymbol{\nu}^{(+)} = \begin{matrix} \\ \mathrm{A} \\ \mathrm{B} \end{matrix} \begin{matrix} 1 & 2 & 3 \end{matrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\nu}^{(-)} = \begin{matrix} \\ \mathrm{A} \\ \mathrm{B} \end{matrix} \begin{matrix} 1 & 2 & 3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.59}$$

An important case when ∅-reactions appear, discussed in Sec. 2.5.3, is when the description is coarse-grained.

2.4.3 Conservation laws

A conserved quantity

$$L \equiv \boldsymbol{\ell} \cdot \boldsymbol{n}, \tag{2.60}$$

is a linear combination of species numbers, with $\boldsymbol{\ell}$ a left nullvector of the stoichiometric matrix

$$\boldsymbol{\ell} \cdot \boldsymbol{\nu} = \mathbf{0}^T. \tag{2.61}$$

This nullvector can be linked to L by considering that

$$\frac{dL}{dt} = \boldsymbol{\ell} \cdot \frac{dn}{dt} = \ell \cdot \boldsymbol{\nu} \cdot \boldsymbol{w}' = 0 \tag{2.62}$$

Such a conserved quantity constrains the chemical compositions that the system can attain. Indeed, any reaction vector $\boldsymbol{g}$ must respect such a conservation law, since

$$\boldsymbol{\ell} \cdot \Delta \boldsymbol{n} = \boldsymbol{\ell} \cdot (\boldsymbol{\nu} \cdot \boldsymbol{g}) = 0. \tag{2.63}$$

A particular class of reaction vectors are those that follow from the dynamic equations (kinetics).

The dimension of the nullspace is the number of linearly independent conservation laws, which we denote by ℓ

$$\ell = \dim(\ker(\nu^{\mathrm{T}})) \tag{2.64}$$

For a particular choice of linearly independent bases, the indexed nullvectors $\boldsymbol{\ell}^{(i)}$ span the left nullspace (cokernel) $\mathscr{L}$ of the stoichiometric matrix $\boldsymbol{\nu}$

$$\mathscr{L} = \mathrm{null}(\nu^{\mathrm{T}}) = \ker(\nu^{\mathrm{T}}), \qquad \boldsymbol{\ell}^{(\mathrm{i})} \in \mathscr{L}. \tag{2.65}$$

Among such conservation laws, we can distinguish two kinds of conservation laws:
i) A mass-like conservation law follows from a left nullvector $\boldsymbol{\ell}^{+}$ with strictly nonnegative components:

$$\forall i,\ \ell_i^{+} \geq 0. \tag{2.66}$$

ii) A mixed conservation law follows from a left nullvector $\boldsymbol{\ell}^{\square}$ with positive and negative components:

$$\exists i, j\ (\ell_i^{\square} > 0, \text{and } \ell_j^{\square} < 0,\ i \neq j). \tag{2.67}$$

For the stoichiometric matrix in Eq. (2.4), $\ell = 4$. One possible choice of four independent conservation laws is

$$L^{(1)} = n_{\mathrm{NH_3}} + n_{\mathrm{NH_4^+}} + n_{\mathrm{CH_3CHNH}}, \tag{2.68}$$

$$L^{(2)} = n_{\mathrm{H_2O}} + n_{\mathrm{H_3O^+}} + n_{\mathrm{OH^-}} + n_{\mathrm{CH_3CHO}}, \tag{2.69}$$

$$L^{(3)} = n_{\mathrm{CH_3CHO}} + n_{\mathrm{CH_3CHNH}}, \tag{2.70}$$

$$L^{(4)} = 4n_{\mathrm{CH_3CHO}} + 5n_{\mathrm{CH_3CHNH}} + 3n_{\mathrm{NH_3}} + 4n_{\mathrm{NH_4^+}} + 2n_{\mathrm{H_2O}} + 3n_{\mathrm{H_3O^+}} + n_{\mathrm{OH^-}}, \tag{2.71}$$

which are all mass-like. Alternatively, we could have constructed a mixed conservation law by taking some combination of the mass-like conservation laws. For example, by choosing $L^{(\square)} = L^{(1)} - L^{(2)}$, we can choose the quadruplet of conservation laws $\{L^{(1)}, L^{(\square)}, L^{(3)}, L^{(4)}\}$.

The distinction between mass-like conservation laws and mixed conservation laws may often not be important, since one can construct one from the other. In upcomping chapters, however, we will encounter situations where the distinction becomes essential. This can happen for subnetworks or when $\ell = 1$.

Mass-like conservation laws

Mass-like conservation laws are very natural in chemistry, since chemical reactions preserve every atom and isotope[‡]. A chemical reaction rearranges the atoms and isotopes. Let us denote by $\chi^{\mathrm{E}}_{\mathrm{Z}}$ the element-counting function, with E denoting an element or particular isotope and Z a chemical species. For a glucose molecule, we can then count the number of carbon, hydrogen, oxygen and xenon atoms:

$$\chi^{\mathrm{C}}_{\mathrm{C_6H_{12}O_6}} = 6, \qquad \chi^{\mathrm{H}}_{\mathrm{C_6H_{12}O_6}} = 12, \qquad \chi^{\mathrm{O}}_{\mathrm{C_6H_{12}O_6}} = 6, \qquad \chi^{\mathrm{Xe}}_{\mathrm{C_6H_{12}O_6}} = 0. \tag{2.72}$$

The preservation of atoms and isotopes implies that we can write

$$\sum_i \chi^{\mathrm{E}}_{\mathrm{Z}_i} (\boldsymbol{\nu} \cdot \boldsymbol{g})_i = 0 \quad \forall\, \mathrm{E}. \tag{2.73}$$

A closed chemical network without hidden reactions or species must therefore have a mass-like conservation law $\boldsymbol{\ell}^{(+,E)}$ for every element E, such that

$$\ell_i^{(+,E)} = \chi^{\mathrm{E}}_{\mathrm{Z}_i}. \tag{2.74}$$

This only yields a useful conservation law for the (few) elements that are actually present.

For Eq. (2.4), only four elements are present, Hydrogen (H), Carbon (C), Nitrogen (N) and Oxygen (O), for which we can construct the conservation laws

$$L^{\mathrm{H}} = 4n_{\mathrm{CH_3CHO}} + 5n_{\mathrm{CH_3CHNH}} + 3n_{\mathrm{NH_3}} + 4n_{\mathrm{NH_4^+}} + 2n_{\mathrm{H_2O}} + 3n_{\mathrm{H_3O^+}} + n_{\mathrm{OH^-}} = L^{(4)} \tag{2.75}$$

$$L^{\mathrm{C}} = 2n_{\mathrm{CH_3CHO}} + 2n_{\mathrm{CH_3CHNH}} = \frac{1}{2} L^{(3)}, \tag{2.76}$$

$$L^{\mathrm{N}} = n_{\mathrm{NH_3}} + n_{\mathrm{NH_4^+}} + n_{\mathrm{CH_3CHNH}} = L^{(1)}, \tag{2.77}$$

$$L^{\mathrm{O}} = n_{\mathrm{H_2O}} + n_{\mathrm{H_3O^+}} + n_{\mathrm{OH^-}} + n_{\mathrm{CH_3CHO}} = L^{(2)}, \tag{2.78}$$

which are all linearly independent. For most reaction networks, however, conservation laws extracted from Eq. (2.74) will not all be linearly independent. When there are more elements than linearly independent conservation laws (ℓ), this is necessarily the case. For example, consider the deprotonation of trichloroacetic acid in water

$$\mathrm{CCl_3COOH + H_2O \rightleftharpoons CCl_3COO^- + H_3O^+} \tag{2.79}$$

whose stoichiometric matrix can be written $\nu = (-1,1,-1,1)^T$. The nullspace yields $\ell = 3$ linearly independent conservation laws, whereas four elements are present.

Conservation laws for unimolecular reactions

A single, connected network containing only unimolecular reactions has a single, mass-like conservation law. On the single-reaction level, this is intuitive, since for such a reaction

$$\nu = (-1,1)^T, \qquad \mathrm{A} \rightleftharpoons \mathrm{B}, \tag{2.80}$$

the nullspace corresponds to solutions of $\ell \cdot \nu = 0$, and thus

$$-\ell_{\mathrm{A}} + \ell_{\mathrm{B}} = 0. \tag{2.81}$$

The solution requires that $\ell_{\mathrm{A}} = \ell_{\mathrm{B}}$. we can extend this conclusion to any pair of species $\mathrm{Z_i}, \mathrm{Z_j}$ linked by a unimolecular reaction: $\ell_{\mathrm{Z_i}} = \ell_{\mathrm{Z_j}}$, leading to a mass-like conservation law $\ell^+ = (1,1,...,1,1)$.

[‡]Of course, nuclear reactions such as $\mathrm{^{10}B + n \longrightarrow {}^{7}Li + \alpha}$ break such conservation laws. For them, a more general treatment is required, which should provide interesting extensions to the framework.

If there are n connected components (disjoint graphs) engaging only in unimolecular reactions, we have n linearly independent conservation laws of the form $\boldsymbol{\ell}^{+} = (0,...,0,1,1,1,1,0,...,0)$. More exactly, let us denote Ω_k the set of species in the kth disjoint network, then a mass-like conservation law for the kth network implies a nullvector $\boldsymbol{\ell}^{+,k}$ such that

$$\ell_i^{+,k} = \delta_{Z_i}^{\Omega_k} \tag{2.82}$$

Where

$$\delta_{Z_i}^{\Omega_k} = \begin{cases} 1 & Z_i \in \Omega_k \\ 0 & Z_i \notin \Omega_k \end{cases} \tag{2.83}$$

For a closed system, Eq. (2.73) implies that there must be at least one conservation law ($\ell \geq 1$).

Linkage classes

In the framework where complexes are used, with links described by an incidence matrix $\boldsymbol{\partial}$, a reversible chemistry yields undirected graphs[§]. Following the terminology introduced by Feinberg, each connected component (disjoint network) in such a graph is called a linkage class[2]. In Fig. 2.2 there are 3 linkage classes. Letting λ denote the number of linkage classes, we observe that

$$\lambda = \dim\left(\ker\left(\boldsymbol{\partial}^{\boldsymbol{T}}\right)\right). \tag{2.84}$$

Structurally, the incidence matrix is equivalent to a stoichiometric matrix with only unimolecular reactions. The left nullspace, spanned by solutions to

$$\boldsymbol{\lambda} \cdot \boldsymbol{\partial} = 0, \tag{2.85}$$

will thus be equivalent to Eq. (2.82). Having again Ω_k denote the set of complexes in the kth linkage class, the coefficient $\lambda_i^{(k)}$ of a complex C_i in the corresponding conservation law $\boldsymbol{\lambda}^{(k)}$ is

$$\lambda_i^{(k)} = \delta_{C_i}^{\Omega_k}. \tag{2.86}$$

In closed systems, linkage classes are strongly tied to conservation laws on the level of complexes. Since a complex contains all chemical species involved in a reaction, a reaction between two complexes respects the conservation of every element. Let us apply the element-counting function on complexes, $\chi_{\mathrm{C_i}}^{\mathrm{E}}$, such that

$$\chi_{\mathrm{H_2O+CH_3CHO}}^{\mathrm{O}} = 2, \quad \chi_{\mathrm{H_2O+CH_3CHO}}^{\mathrm{C}} = 2, \quad \chi_{\mathrm{H_2O+CH_3CHO}}^{\mathrm{H}} = 6. \tag{2.87}$$

In a closed system, any reaction connecting a pair of complexes C_i, C_j therefore satisfies

$$\chi_{\mathrm{C_i}}^{\mathrm{E}} = \chi_{\mathrm{C_j}}^{\mathrm{E}} \quad \forall\, \mathrm{E}, \tag{2.88}$$

which can trivially be converted to the conservation law (2.86). Unlike the mass-like conservation laws derived on the level of species, these laws on the level of complexes conserve all elements simultaneously.

When comparing a complex network approach to a stoichiometric matrix approach, we may not always find the same number of conservation laws. This is evidenced by Eq. (2.4), where we have $\lambda = 3$ and $\ell = 4$. In that example, conservation laws were found for each individual element in isolation, whereas a linkage class (2.88) requires the collective of all elements to be conserved.

Note that we do not need explicit knowledge of elements to find conservation laws, we only need to consider the nullspace of the network. Reaction networks can be entirely symbolic without reference to particular elements or composition. Such networks, however, still have to obey the constraints imposed by chemistry, and tracking elements is an instructive way of finding and understanding such constraints.

[§] relaxing the reversibility makes it a directed graph

Mixed conservation laws

Since networks with only unimolecular reactions only have a mass-like conservation law, true mixed conservation laws require bimolecular reactions to exist. On the single-reaction level, this becomes clear by considering that a reaction $\mathrm{A}+\mathrm{B} \rightleftharpoons \mathrm{C}$ with $\boldsymbol{\nu} = (-1,-1,1)^T$, admits conservation laws such that $\boldsymbol{\ell} \cdot \boldsymbol{\nu} = 0$, therefore

$$\ell_\mathrm{A} + \ell_\mathrm{B} - \ell_\mathrm{C} = 0. \tag{2.89}$$

This admits mass-like solutions $(1,0,1)$, $(0,1,1)$, $(1,1,2)$. A possible mixed solution is $(1,-1,0)$. Denoting $\ell^{\square} = (1,-1,0)^T$, we have a mixed conservation law

$$L = \boldsymbol{\ell}^{\square} \cdot \boldsymbol{n} = n_\mathrm{A} - n_\mathrm{B}. \tag{2.90}$$

Since A and B can only be formed and consumed together, the network conserves their difference in abundance.

One simple example where mixed conservation laws naturally appear is in the disproportionation of an electrically neutral species, to form oppositely charged species. Let $\chi^{+e}_{Z_i}$ be the charge counting function that returns the net (elementary) charge of a species. As an example, we can consider

$$\chi^{+e}_{\mathrm{PO_4}^{3-}} = -3, \quad \chi^{+e}_{\mathrm{H_2O}} = 0, \quad \chi^{+e}_{\mathrm{H_3O^+}} = 1. \tag{2.91}$$

For a closed chemical network, we require the conservation of net electrical charge

$$\sum_i \chi^{+\mathrm{e}}_{\mathrm{X}_i} (\boldsymbol{\nu} \cdot \boldsymbol{g})_i = 0. \tag{2.92}$$

From which a conservation law L^{+e} can be found. E.g. for deprotonation of $\mathrm{CCl_3COOH}$ (Eq. (2.79)) we have

$$L^{+e} = n_{\mathrm{H_3O^+}} - n_{\mathrm{CCl_3COO^-}} \tag{2.93}$$

2.4.4 Reaction Cycles

The term 'cycle' is used in chemistry to describe sequences of reactions that 'loop around'. Such reactions collectively describe catalysis, autocatalysis, metabolic pathways and other processes. In the stoichiometric matrix framework, the term 'cycle' is used in a highly similar, but not equivalent manner, to denote right nullvectors of the stoichiometric matrix. In the following sections, the connection between the two will be made explicit.

The right nullspace of $\boldsymbol{\nu}$, $\mathscr{C}$, is spanned by cycles. A cycle $\boldsymbol{c}$ is a linear combinations of reactions $(c_1, c_2, ..., c_r)$ that leaves the system unchanged

$$\boldsymbol{\nu} \cdot \boldsymbol{c} = \boldsymbol{0} \tag{2.94}$$

We denote the number of linearly independent cycles by c, which corresponds to the kernel (nullspace) of the stoichiometric matrix

$$c = \dim(\mathscr{C}) = \dim(\ker(\boldsymbol{\nu})) \tag{2.95}$$

A cycle $\boldsymbol{c}$ is a particular instance of a reaction vector $\boldsymbol{g}$, for which, as follows from Eq. (2.94),

$$\Delta n_i = (\boldsymbol{\nu} \cdot \boldsymbol{c})_i = 0 \ \ \forall\, i. \tag{2.96}$$

As an example, let us consider a more elaborate description of acid-base reactions of CCl_3COOH in water

$$\boldsymbol{\nu} = \begin{array}{c} \\ CCl_3COOH \\ CCl_3COO^- \\ H_2O \\ OH^- \\ H_3O^+ \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & -2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \end{array} \quad \begin{array}{c} CCl_3COOH + H_2O \overset{1}{\rightleftharpoons} CCl_3COO^- + H_3O^+ \\ 2\,H_2O \overset{2}{\rightleftharpoons} OH^- + H_3O^+ \\ CCl_3COO^- + H_2O \overset{3}{\rightleftharpoons} CCl_3COOH + OH^- \end{array} \tag{2.97}$$

This network admits a cycle $\boldsymbol{c} = (1,-1,1)$, which corresponds to performing reaction 1 in forward direction, reaction 2 in reverse, and reaction 3 forwards.

Cycles for complexes

For complexes, the notion of a cycle is often used in the graph-theoretical sense, where a sequence of nodes (complexes) that finishes at the initial node is given. This requires the nodes to be part of the same linkage class, and a complex cycle $\boldsymbol{\gamma}$ is in the nullspace of the incidence matrix $\boldsymbol{\partial}$, and there are γ linearly independent complex cycles:

$$\boldsymbol{\partial} \cdot \boldsymbol{\gamma} = 0, \qquad \gamma = \dim(\ker(\boldsymbol{\partial})). \tag{2.98}$$

For closed systems, the linkage class conserves the full elemental composition between complexes (Eq. (2.88)), which puts an additional constraint on cycles with respect to those defined for a stoichiometric matrix $\boldsymbol{\nu}$. If we write the incidence matrix $\boldsymbol{\partial}$ for Eq. (2.98), we find

$$\boldsymbol{\partial} = \begin{array}{c} \\ CCl_3COOH{+}H_2O \\ CCl_3COO^-{+}H_3O^+ \\ 2H_2O \\ OH^-{+}H_3O^+ \\ CCl_3COO^-{+}H_2O \\ CCl_3COOH{+}OH^- \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \quad \begin{array}{c} CCl_3COOH + H_2O \overset{1}{\rightleftharpoons} CCl_3COO^- + H_3O^+ \\ 2\,H_2O \overset{2}{\rightleftharpoons} OH^- + H_3O^+ \\ CCl_3COO^- + H_2O \overset{3}{\rightleftharpoons} CCl_3COOH + OH^- \end{array} \tag{2.99}$$

which has more complexes than chemical species. Contrary to $\boldsymbol{\nu}$, there is no cycle ($\gamma = 0$), since the reactions that generate $\boldsymbol{c}$ are in different linkage classes.

Deficiency

The stoichiometric matrix $\boldsymbol{\nu}$ and incidence matrix $\boldsymbol{\partial}$ share the same reactions. The nullspace spanned by the reactions in $\boldsymbol{\nu}$ contains at least the nullspace of $\boldsymbol{\partial}$, as can also be inferred from Eq. 2.6.[7]. Equivalently [2, 3], we can write

$$\dim(\ker(\boldsymbol{\nu})) = c \geq \dim(\ker(\boldsymbol{\partial}) = \gamma. \tag{2.100}$$

The difference in the number of cycles in these descriptions is called the deficiency δ

$$\delta = c - \gamma = n. \tag{2.101}$$

The deficiency is notably used in the deficiency zero theorem and the deficiency one theorem[2]. For reversible chemical reaction networks with $\delta = 0$, the deficiency zero theorem states that the concentrations of chemical species (whose dynamics is governed by mass-action kinetics) will relax to a single fixed point.

The deficiency one theorem can be restated as follows: Consider a chemical reaction network, with incidence matrix $\boldsymbol{\partial}$, stoichiometric matrix $\boldsymbol{\nu}$, and a deficiency δ. Furthermore, consider the

subnetworks obtained by taking the submatrices corresponding to all λ linkage classes $\boldsymbol{\partial}_1,...,\boldsymbol{\partial}_\lambda$ and the corresponding stoichiometric submatrices $\boldsymbol{\nu}_1,...,\boldsymbol{\nu}_\lambda$, with deficiencies $\delta_i, i = 1,...,\lambda$. If then, the following are true:
1. $\forall\, i \quad \delta_i \leq 1$,
2. $\sum_{i=1}^{\lambda} \delta_i = \delta$,
the concentrations of chemical species (whose dynamics is governed by mass-action kinetics) will relax to a single fixed point.

While these theorems are highly informative, there are many simple systems for which the theorems are inconclusive. As an example, in the system given by Eq. (2.99), individual linkage classes admit no cycles, and neither do the corresponding submatrices $\boldsymbol{\nu}_i$, which means $\delta_i = 0$ for all i. Consequently,

$$\sum_{i=1}^{\lambda} \delta_i = 0 < \delta = 1, \tag{2.102}$$

which means that we can apply neither the deficiency one nor the deficiency zero theorem. Nevertheless, we know that such a system should relax to a unique steady state, because a closed chemical system in general admits a single minimum to its Gibbs free energy function.

2.5 Subspaces and submatrices

2.5.1 The four subspaces

Each s-by-r matrix A has four fundamental subspaces:
i) the column space or image (im(A))
ii) the nullspace or kernel (null(A) or ker(A))
iii) the row space or coimage (im($\mathrm{A^T}$))
iv) the left nullspace or cokernel (null(A) or ker(A)).

The rank nullity theorem relates the dimension of the domain to the dimension of the image and the kernel (respectively called the rank and the nullity), :

$$\mathrm{rank(A) + nullity(A) = dim(im(A)) + dim(ker(A)) = r}, \tag{2.103}$$

$$\mathrm{rank(A^T) + nullity(A^T) = dim(im(A^T)) + dim(ker(A^T)) = s}. \tag{2.104}$$

Furthermore, we know that $\mathrm{rank(A) = rank(A^T)}$. For a stoichiometric matrix $\boldsymbol{\nu}$, we have related the nullity to cycles and conservation laws:
$c = \mathrm{nullity}(\boldsymbol{\nu})$, with c the number of cycles,
$\ell = \mathrm{nullity}(\boldsymbol{\nu^T})$, with ℓ the number of conservation laws,
Putting the aforementioned together, we obtain the fundamental theorem of linear algebra [3], that is

$$r - c = s - \ell. \tag{2.105}$$

Denoting by σ the number of linkage classes, a similar result applies to the incidence matrix $\boldsymbol{\partial}$,

$$r - \gamma = \sigma - \lambda. \tag{2.106}$$

Note that the two descriptions must have the same number of reactions r, but as shown in previous examples, the number of species, cycles and conservation laws may differ between the two approaches. Using Eq. (2.106), we can write the deficiency in its more common form

$$\delta = \sigma - \lambda - (s - \ell) \geq 0, \tag{2.107}$$

where $s - \ell$ is the stoichiometric subspace (rank), which quantifies the dimensionality of the concentration space.

An example for a discretized grid

Let us return to the discretized N by N grid from Sec. 2.3.3. Let us imagine such a grid with r chemical reactions per box, s species and ms exchange reactions (diffusion) with neighbors (where m is the number of neighbors). An illustration for $N = 2$ is given in Fig. 2.4, for a 2 by 2 grid, with each cell in contact with $m = 2$ neighbors, a single reaction ($r = 1$), and $s = 2$ chemical species. From Eq. (2.105), we then find for an N by N grid

$$c = N^2(r + ms - s) + \ell, \tag{2.108}$$

where we note that the number of conservation laws ℓ is not altered by the discretization operation. Denoting c_N the number of cycles for an N by N grid, the number of cycles due to box partitioning grows as

$$c_N - c_1 = (N^2 - 1)(r + ms - s). \tag{2.109}$$

Note that a cycle can be a trajectory that returns to its origin, e.g.

$$\mathrm{A}_{x,y} \to \mathrm{A}_{x+h,y} \to \mathrm{A}_{x+h,y+h} \to \mathrm{A}_{x,y+h} \to \mathrm{A}_{x,y}. \tag{2.110}$$

The cycle $\boldsymbol{c}_1$ in Fig. 2.4b is an example of such a cycle. A cycle can also contain a mixture of reaction and displacement reactions, as exemplified by the cycle $\boldsymbol{c}_2$. The 4-cell network has one chemical reaction per cell ($r = 1$), two neighbors ($m = 2$) and there are two possible chemical species ($s = 2$). From Eq. (2.108) we then find $c = 7$ cycles.

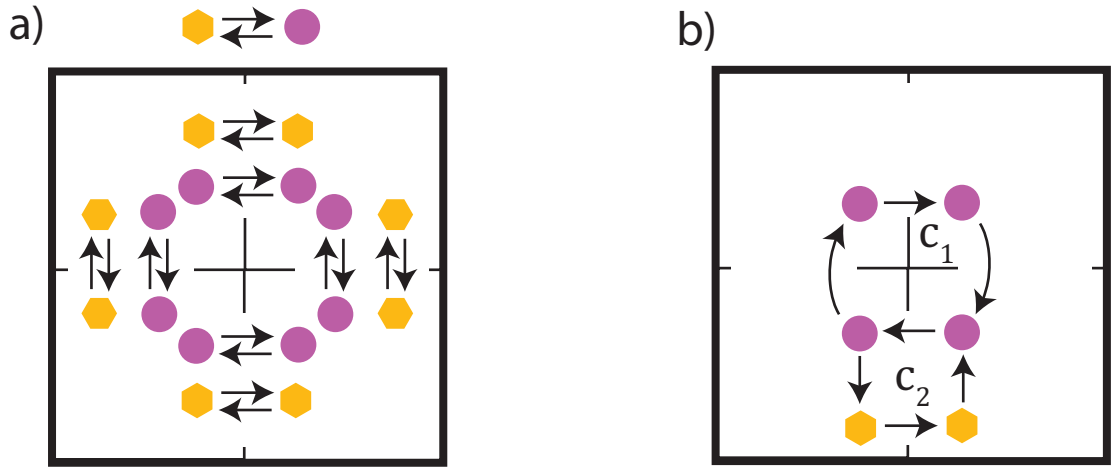


Figure 2.4: Reactions and transport in 4 linked cells. a) There are 8 exchange processes, 4 per species. The purple and yellow species can isomerize to form each other. b) a pure displacement cycle $\boldsymbol{c}_1$, and a mixed cycle $\boldsymbol{c}_2$.

In Ref. [21], cycles coupling displacement and chemistry were realized out-of-equilibrium. By the application of patterned illumination, a reversible photoacid supplied protons in illuminated regions, which would then undergo elaborate reaction-diffusion cycles by reacting with reactants in other regions.

2.5.2 Submatrices

Consider a stoichiometric matrix $\boldsymbol{\nu}$. Let $[s]$ be the set of species indices, $[s] = \{1, 2, .., s\}$ and let $[r]$ be the set of reaction indices $[r] = \{1, 2, .., r\}$, such that the rows of $\boldsymbol{\nu}$ are $\boldsymbol{\nu} = [\boldsymbol{s}_1, \boldsymbol{s}_2, .., \boldsymbol{s}_s]^{\mathrm{T}}$ and the columns $\boldsymbol{\nu} = [\boldsymbol{r}_1, \boldsymbol{r}_2, .., \boldsymbol{r}_r]$. Let us define $x \subset [s]$, $z \subset [r]$ as proper subsets of the indices, with complements $\bar{x} = [s]/x$, $\bar{z} = [r]/z$.

Definition 2.5.1 — Submatrix. A submatrix $\boldsymbol{\nu}_*$ of $\boldsymbol{\nu}$ is a matrix obtained by removing rows (species) and/or columns (reactions) from $\boldsymbol{\nu}$. By convention, $\boldsymbol{\nu}$ is a submatrix of $\boldsymbol{\nu}$.

We define $\boldsymbol{\nu}_* = \boldsymbol{\nu}(x|z)$ as a submatrix, constructed by removing all ν_{ki} with $k \in x$ and $i \in z$ from $\boldsymbol{\nu}$ (remove all species labeled by x and all reactions labeled by z). We define $\boldsymbol{\nu}_* = \boldsymbol{\nu}[x|z]$ as a submatrix, containing only ν_{ki} with $k \in x$ and $i \in z$ from $\boldsymbol{\nu}$ (retain species labeled by x and reactions labeled by z). It follows that

$$\boldsymbol{\nu}(x|z) = \boldsymbol{\nu}[\bar{x}|\bar{z}] \tag{2.111}$$

Eq. (2.112) illustrates $\boldsymbol{\nu}$ with two of its submatrices, with the red-colored matrix obtained by removing the fourth row and third column, and the blue-colored matrix by removing the first two rows and the first and last column:

$$\begin{pmatrix} \nu_{1,1} & \nu_{1,2} & \nu_{1,3} & \nu_{1,4} \\ \nu_{2,1} & \nu_{2,2} & \nu_{2,3} & \nu_{2,4} \\ \nu_{3,1} & \nu_{3,2} & \nu_{3,3} & \nu_{3,4} \\ \nu_{4,1} & \nu_{4,2} & \nu_{4,3} & \nu_{4,4} \end{pmatrix} \begin{matrix} \boldsymbol{\nu}(4|3) \\ \cdot \\ \boldsymbol{\nu}(1,2|1,4) \end{matrix} \tag{2.112}$$

Translating the destructive convention to the constructive convention, we can write, $\boldsymbol{\nu}(4|3) = \boldsymbol{\nu}[1,2,3|1,2,4]$ (given in red) and $\boldsymbol{\nu}(1,2|1,4) = \boldsymbol{\nu}[3,4|2,3]$ (shown in blue).

Submatrices associated with linkage classes play a central role in the deficiency one theorem. Consider the following incidence matrix $\boldsymbol{\partial}$, whose two linkage classes can be extracted as separate submatrices (respectively shown in red and blue)

$$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} \boldsymbol{\partial}(4,5|4) \\ \cdot \\ \boldsymbol{\partial}[4,5|4] \end{matrix} \tag{2.113}$$

Submatrices are also important for the modelling of species whose concentrations have been fixed by chemostats (see Sec. 2.5.3). In a mass action framework, a rate $k' n_{\mathrm{A}} n_{\mathrm{B}}$ depends on the abundance of the involved species A and B. If a chemostat now fixes $n_{\mathrm{B}} = \bar{n}$, we can write a rate only in terms of A: $k' n_A$, with $k' = k\bar{n}$ a pseudo-rate constant. We can therefore remove B from our description. The stoichiometric matrix for this new network is obtained by removing the row corresponding to B from the original matrix.

Let us denote by $y \subset [s]$ the indices of chemostatted species and $x \subset [s]$ and the indices of internal species, such that $\bar{y} = x$. The internal submatrix $\boldsymbol{\nu}_{\mathrm{X}}$ then verifies $\boldsymbol{\nu}_{\mathrm{X}} = \boldsymbol{\nu}[x|[r]] = \boldsymbol{\nu}(y|\emptyset)$. The external submatrix $\boldsymbol{\nu}_{\mathrm{Y}}$ associated to chemostats is then $\boldsymbol{\nu}_{\mathrm{Y}} = \boldsymbol{\nu}[y|[r]] = \boldsymbol{\nu}(x|\emptyset)$.

2.5.3 Chemostatting and subspaces

When a chemostat is introduced, we lower the number of species by 1: $s' = s - 1$. Polettini et al. point out that [3] for $\boldsymbol{\nu}_{\mathrm{X}}$ to be in accord with Eq. (2.105), either i) $c' = c + 1$, an 'emergent cycle' or 'affinity' is generated, or ii) $\ell' = \ell - 1$, a conservation law is broken. More explicitly, we can write

$$s^Y = c^* + l^\times, \tag{2.114}$$

with s^Y the number of chemostats, c^* the number of emergent cycles and $l^\times$ the number of broken conservation laws.

An emergent cycle $\boldsymbol{c}^*$ inhabits the nullspace of $\boldsymbol{\nu}_\mathrm{X}$, but not of $\boldsymbol{\nu}$ and $\boldsymbol{\nu}_\mathrm{Y}$

$$\boldsymbol{\nu}_\mathrm{X} \cdot \boldsymbol{c}^* = 0, \tag{2.115}$$

$$\boldsymbol{\nu}_\mathrm{Y} \cdot \boldsymbol{c}^* \neq 0. \tag{2.116}$$

In a description without chemostats, $\boldsymbol{c}^*$ is a reaction vector. Like any reaction vector, it must respect the conservation laws of the full stoichiometric matrix $\boldsymbol{\nu}$ (Eq. (2.63)), and thus a conservation law L, even if broken, still constrains the emergent cycle, since its corresponding $\boldsymbol{\ell}$ imposes

$$\boldsymbol{\ell} \cdot (\boldsymbol{\nu} \cdot \boldsymbol{c}^*) = 0. \tag{2.117}$$

An emergent cycle requires more than a single chemostat[¶], since Eq. (2.116) would then imply a net accumulation or consumption of the chemostatted compound, with no compensation by the system.

Indeed, emergent cycles require at least two chemostats[3]. On the level of reservoir species $\{\mathrm{Y}_1, ..., \mathrm{Y}_\mathrm{s}\}$, we can write a reaction balance similar to (2.47)

$$\sum_k n_k^{(+)} \mathrm{Y_k} \underset{-\boldsymbol{g}}{\overset{\boldsymbol{g}}{\rightleftharpoons}} \sum_k n_k^{(-)} \mathrm{Y_k}. \tag{2.118}$$

Similarly, we can write a balance in terms of internal species $\{\mathrm{X}_1, ..., \mathrm{X}_\mathrm{s}\}$

$$\sum_k \tilde{n}_k^{(+)} \mathrm{X_k} \underset{-\boldsymbol{g}}{\overset{\boldsymbol{g}}{\rightleftharpoons}} \sum_k \tilde{n}_k^{(-)} \mathrm{X_k}. \tag{2.119}$$

From Eq. (2.115) it follows that for an emergent cycle $\boldsymbol{c}^*$ internal species are as much consumed as they are produced

$$\tilde{n}_k^{(+)} = \tilde{n}_k^{(-)}. \tag{2.120}$$

The currents between reservoirs are constrained by the compositional constraints of chemistry (Eq. (2.73)). For two chemostats Y_1 and Y_2, we find on the level of elements E that a current must satisfy

$$n_1^{(+)} \chi_{\mathrm{Y}_1}^\mathrm{E} = n_2^{(-)} \chi_{\mathrm{Y}_2}^\mathrm{E} \quad \forall\, \mathrm{E}. \tag{2.121}$$

for $n_1^{(+)} = n_2^{(-)}$, the pair of chemostats must either contain the same species or isomers of the same species. If $n_1^{(+)} \neq n_2^{(-)}$, one can also imagine oligomers e.g. 3 dimers becoming 2 trimers, (potentially accompanied by isomerization). For an arbitrary pair of chemical species, Eq. (2.121) is typically not satisfied and the scenarios for currents afforded by two chemostats is fairly limited. In practice, emergent cycles are expected to become prevalent when more chemostats come into play, such that Eq. (2.73) can be satisfied more easily.

Example: HBr chain-reaction

We will now move to an illustration of emergent cycles and broken conservation laws. Let us consider the chain reaction converting molecular bromine (Br_2) and molecular hydrogen (H_2) to hydrobromic acid (HBr):

$$\mathrm{Br_2} \quad \overset{1}{\rightleftharpoons} \quad 2\mathrm{Br}^\bullet \tag{2.122}$$

$$\mathrm{Br}^\bullet + \mathrm{H_2} \quad \overset{2}{\rightleftharpoons} \quad \mathrm{HBr} + \mathrm{H}^\bullet \tag{2.123}$$

$$\mathrm{H}^\bullet + \mathrm{Br_2} \quad \overset{3}{\rightleftharpoons} \quad \mathrm{HBr} + \mathrm{Br}^\bullet \tag{2.124}$$

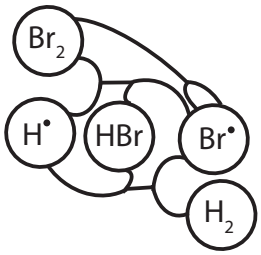


Figure 2.5: A hypergraph for the HBR chain reaction.

A corresponding hypergraph is shown in Fig. 2.5: For this system, we can write the stoichiometric matrix $\boldsymbol{\nu}$

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{Br_2} \\ \mathrm{Br^\bullet} \\ \mathrm{H_2} \\ \mathrm{H^\bullet} \\ \boldsymbol{\nu}_\star \quad \mathrm{HBr} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 2 & -1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \end{array} \tag{2.125}$$

The full matrix $\boldsymbol{\nu}$ has $r = 3$, $s = 5$, $\ell = 2$ and consequently (Eq. (2.105)), no cycles ($c = 0$). Upon inspection of elements, we can directly make a natural choice for two linearly independent conservation laws:

$$L^{\mathrm{H}} = 2n_{\mathrm{H_2}} + n_{\mathrm{HBr}} + n_{\mathrm{H^\bullet}}, \qquad L^{\mathrm{Br}} = 2n_{\mathrm{Br_2}} + n_{\mathrm{HBr}} + n_{\mathrm{Br^\bullet}}. \tag{2.126}$$

Let us now consider the submatrix $\nu_\star = \nu(1|\emptyset)$ (shown in red) obtained by removing $\mathrm{Br_2}$. For this submatrix, L^{Br} is no longer a conservation law, and in total only one conservation law remains. If we now also remove $\mathrm{H_2}$, we also lose the conservation law L^{H}, and the reactions are now

$$\emptyset \overset{1}{\rightleftharpoons} 2\mathrm{Br^\bullet}, \qquad \mathrm{Br^\bullet} \overset{2}{\rightleftharpoons} \mathrm{HBr} + \mathrm{H^\bullet}, \qquad \mathrm{H^\bullet} \overset{3}{\rightleftharpoons} \mathrm{HBr} + \mathrm{Br^\bullet} \tag{2.127}$$

For this system, we can write the stoichiometric matrix $\tilde{\boldsymbol{\nu}}$

$$\tilde{\boldsymbol{\nu}} = \begin{array}{c} \\ \mathrm{Br^\bullet} \\ \mathrm{H^\bullet} \\ \tilde{\boldsymbol{\nu}}_\star \quad \mathrm{HBr} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \end{array} \tag{2.128}$$

$\tilde{\boldsymbol{\nu}}$ is a square matrix with no conservation laws and no cycles. However, if we now finally remove HBr to yield $\tilde{\boldsymbol{\nu}}_\star$ (shown in red), an emergent cycle $\boldsymbol{c}^* = (0,1,1)^T$ is generated, such that $\tilde{\boldsymbol{\nu}}_\star \cdot \boldsymbol{c}^* = \boldsymbol{0}$. Returning to the original matrix $\boldsymbol{\nu}$, the reaction vector $\boldsymbol{c}^*$ yields

$$\Delta \boldsymbol{n} = \boldsymbol{\nu} \cdot \boldsymbol{c}^* = (-1,0,-1,0,2)^T, \tag{2.129}$$

¶here, we are implicitly considering a system that has no other thermodynamic forces applied to it. In Ref[22], emergent cycles were elegantly introduced by introducing spatially localized photochemical reactions. As of yet, the framework has not been extended to treat such a system.

We can apply Eq. (2.118), to find the net reaction balance on the level of removed species (for chemostats: reservoir species)

$$\mathrm{Br_2} + \mathrm{H_2} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} 2\mathrm{HBr}. \tag{2.130}$$

In the chemostatted system, all of these compounds are chemostatted, and the emergent cycle will generate a current between the reservoirs of $\mathrm{Br_2} + \mathrm{H_2}$ and the reservoir of HBr. We can also derive a more detailed reaction balance, using Eq. (2.54), which becomes

$$\mathrm{Br_2} + \mathrm{H_2} + \mathrm{Br}^\bullet + \mathrm{H}^\bullet \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} 2\mathrm{HBr} + \mathrm{Br}^\bullet + \mathrm{H}^\bullet. \tag{2.131}$$

Here, we see that the radicals $\mathrm{Br}^\bullet$ and $\mathrm{H}^\bullet$ appear with the same nonzero stoichiometry on both sides. The participation of radicals here mirrors that of a catalytic intermediate. On the atomic level, however, the bromine radical $\mathrm{Br}^\bullet$ in (2.123) is replaced by a different one in (2.124). The hydrogen radical $\mathrm{H}^\bullet$ is also constantly consumed and resupplied. Sometimes this distinction is highlighted by the use of the word chain-reaction, at other times catalysis is used. To highlight such atomic details, an atomic-level description of connectivity is needed, as displayed e.g. in SMILES notation[23]. As noted at its inception[1], such a level of description goes beyond the stoichiometric matrix framework.

2.5.4 Chemostatting and complexes

When we remove a species from the description, its associated complexes are modified as well. E.g. if we have complexes $C_1 = \{X_1, X_2\}$ and $C_2 = \{X_2\}$, chemostatting X_1 will make C_1 and C_2 the same complex. Since we want complex to be unique, we will perform a 'merger', by merging C_1 and C_2 in a new complex Γ_1, which performs the reactions that C_1 and C_2 are engaged in.

If the operation of chemostatting removes all species from a complex, we obtain an empty complex $\Gamma = \{\emptyset\}$. Such a complex is often used to model influx and outflux.

Chemostatting modifies the existing set of complexes. When this leads to the merger of complexes, the number of removed complexes $\sigma^\times$ corresponds to the number of mergers. Mergers require chemostats, but the number of mergers can be more or less numerous than the number of chemostats introduced.

$$\sigma^\times \geq 0, \quad s^y > 0, \tag{2.132}$$
$$\sigma^\times = 0, \quad s^y = 0. \tag{2.133}$$

When complexes in different linkage classes are merged, their linkage classes are also merged: they now form a single connected component.

Complex cycles can emerge, either by a single merger within an existing linkage class, or due to multiple mergers with other linkage classes.

Fixing the number of reactions $r' = r$, Eq. (2.106) yields

$$\sigma^\times = \lambda^\times + \gamma^*, \tag{2.134}$$

with $\lambda^\times$ the reduction in the number of linkage classes (number of merged linkage classes) and γ^* the number of emergent cycles.

As an example, let us consider the chain reaction between $\mathrm{H_2}$ and $\mathrm{Br_2}$ to yield HBr in Eq.

(2.125), with the following incidence matrix

$$\boldsymbol{\partial} = \begin{array}{c} \\ \mathrm{Br_2} \\ \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet{+}H_2} \\ \mathrm{HBr{+}H^\bullet} \\ \mathrm{H^\bullet{+}Br_2} \\ \mathrm{HBr{+}Br^\bullet} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} \mathrm{Br_2} \overset{1}{\rightleftharpoons} \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet} + \mathrm{H_2} \overset{2}{\rightleftharpoons} \mathrm{HBr} + \mathrm{H^\bullet} \\ \mathrm{H^\bullet} + \mathrm{Br_2} \overset{3}{\rightleftharpoons} \mathrm{HBr} + \mathrm{Br^\bullet} \end{array} \tag{2.135}$$

For $\boldsymbol{\partial}$, we find $\lambda = 3$. Upon chemostatting $\mathrm{Br_2}$, the incidence matrix is not modified:

$$\boldsymbol{\partial} = \begin{array}{c} \\ \emptyset \\ \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet{+}H_2} \\ \mathrm{HBr{+}H^\bullet} \\ \mathrm{H^\bullet} \\ \mathrm{HBr{+}Br^\bullet} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} \emptyset \overset{1}{\rightleftharpoons} \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet} + \mathrm{H_2} \overset{2}{\rightleftharpoons} \mathrm{HBr} + \mathrm{H^\bullet} \\ \mathrm{H^\bullet} \overset{3}{\rightleftharpoons} \mathrm{HBr} + \mathrm{Br^\bullet} \end{array} \tag{2.136}$$

Let us now chemostat the species HBr, which then provides

$$\tilde{\boldsymbol{\partial}} = \begin{array}{c} \\ \emptyset \\ \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet{+}H_2} \\ \mathrm{H^\bullet} \\ \mathrm{Br^\bullet} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} \emptyset \overset{1}{\rightleftharpoons} \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet} + \mathrm{H_2} \overset{2}{\rightleftharpoons} \mathrm{H^\bullet} \overset{3}{\rightleftharpoons} \mathrm{Br^\bullet} \end{array} \tag{2.137}$$

In removing HBr and $\mathrm{Br_2}$ from the description, the complexes $\{\mathrm{HBr} + \mathrm{H^\bullet}\}$ and $\{\mathrm{H^\bullet} + \mathrm{Br_2}\}$ both are reduced to the complex $\{\mathrm{H^\bullet}\}$, the distinct complexes merge into a single one. The complexes were part of distinct linkage classes, which now form a single linkage class in which both the second and third reaction are performed. We find $\sigma^\times = \lambda^\times = 1$.

Let us now remove $\mathrm{H_2}$, to find

$$\boldsymbol{\partial}_\star = \begin{array}{c} \\ \emptyset \\ \mathrm{2Br^\bullet} \\ \mathrm{H^\bullet} \\ \mathrm{Br^\bullet} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} \emptyset \overset{1}{\rightleftharpoons} \mathrm{2Br^\bullet} \\ \mathrm{Br^\bullet} \overset{2}{\rightleftharpoons} \mathrm{H^\bullet} \overset{3}{\rightleftharpoons} \mathrm{Br^\bullet} \end{array} \tag{2.138}$$

The merger of the complexes $\{\mathrm{Br^\bullet} + \mathrm{H_2}\}$ and $\{\mathrm{Br^\bullet}\}$ now yields a complex cycle $\boldsymbol{\gamma}^* = (0,1,1)^T$, such that $\boldsymbol{\partial}_\star \cdot \boldsymbol{\gamma}^* = 0$, and we have $\sigma^\times = 2$, $\lambda^\times = 1$, $\gamma^* = 1$.

Minimal number of chemostats for a cycle

Since we impose reactions to be at most bimolecular (Eq. (2.9)), complexes contain at most two molecules. It can then be shown that in a full description (no hidden species) of a closed chemical system (where no $\emptyset$-complex can exist), the introduction of a first chemostat cannot merge two elements in the same linkage class.

To show this, suppose that there exist two distinct complexes C_1 and C_2 that are merged after chemostatting a species Z_1. For a closed system, this can happen for the following situations:
i) $C_1 = \{2Z_1\}, C_2 = \{Z_1\}$,

ii) $C_1 = \{Z_1, Z_2\}, C_2 = \{Z_2\}$.
In a closed system, linkage classes link complexes with the same elemental composition (Eq. (2.88)). In both situation i) and ii), the compound in C_2 is contained within C_1, but C_1 contains an additional Z_1 species. There are thus inherently more atoms composing C_1 and thus its composition must be distinct. It follows that C_1 and C_2 are not part of the same linkage class. We conclude that connecting a first chemostat to a closed system cannot merge two elements within the same linkage class.

To acquire a complex cycle from two separate linkage classes, two mergers are needed, one to establish a single linkage class and a subsequent one to form a cycle. Starting with a closed system and four distinct complexes C_1, C_2, C_3, C_4, this can be achieved for: $C_1 = \{\mathrm{Z}_1\}, C_2 = \{\mathrm{Z}_2\}, C_3 = \{\mathrm{Z}_1, \mathrm{Z}_1\}, C_4 = \{\mathrm{Z}_1, \mathrm{Z}_2\}$, provided Z_1 is an isomer of Z_2, such that there are two linkage classes. Chemostatting Z_1 then merges C_1 with C_3 in the complex $\{\emptyset\}$ and C_2 with C_4 to give $\{Z_2\}$, to establish a single linkage class with a complex cycle.

Of course, most reaction networks do not have this structure, and one should typically expect more chemostats to be required to yield a complex cycle. For the stoichiometric matrix framework, it was shown that at least two chemostats are needed to obtain an emergent cycle. A complex cycle can be obtained with only one chemostat, which can be explained when we consider the deficiency

$$\delta = c - \gamma \geq 0. \tag{2.139}$$

The complex cycle established by a single chemostat requires the $\boldsymbol{\nu}$ to already have this cycle from the start. Since $c \geq 1$ it follows for the closed system that $\delta \geq 1$.

2.5.5 An alternative perspective to chemostatting: adding reactions

An alternative approach to chemostatting a species in $\boldsymbol{\nu}$ is to introduce a reversible exchange reaction with a (simple) reservoir

$$\mathrm{A}(reservoir) \rightleftharpoons \mathrm{A}(system). \tag{2.140}$$

Since the reservoir is considered stationary, we may simplify our description limiting it to system species, leading to a reaction

$$\emptyset \rightleftharpoons \mathrm{A}. \tag{2.141}$$

If such a reaction occurs on a much faster timescale than the other reactions (and provided the solution is well-mixed), the species A acquires a fixed concentration and its dynamics become equivalent to the former approach.

By adding a reaction $r' = r + 1$, and leaving the number of species untouched $s' = s$, Eq. (2.105) provides an equivalence between the number removed species s^Y and the number of reservoir exchange reactions r^*

$$r^* = \ell^\times + c^*, \tag{2.142}$$

which has the same interpretation as Eq. (2.114). Using this approach, a closed system is described by a submatrix of an open system. When we remove species to introduce a chemostat, open systems are described by submatrices of the closed system.

For complexes, $\{\emptyset\}$ must be introduced as a separate complex first. If we require reservoir exchange processes to be unimolecular (e.g. simple diffusion), then adding reservoir exchange reactions is different from removing species. E.g. if we introduce a reaction $\emptyset \rightleftharpoons \mathrm{Z}_1$, it will be connected to the complex $\{Z_1\}$, but not $\{Z_1, X_1\}$. Since many reactions of interest are not unimolecular, this may often not be a productive way to introduce chemostats in the framework.

With such exchange reactions, chemostatting no longer removes complexes, but establishes new links between the $\{\emptyset\}$ complex and chemostatted species. When linkage classes are linked up in this manner, they do so via the $\{\emptyset\}$ complex. Similarly, emergent cycles must pass through the bridging $\{\emptyset\}$ complex.

2.6 Submatrices and removing reactions

We can also be interested in removing reactions from $\boldsymbol{\nu}$. This is a useful venture when we are looking for particular behaviors such as catalysis, chain reactions and autocatalysis, which are properties of subnetworks, not of the total network. For our purposes, this operation will serve as a mathematical tool to look under the hood and learn something about the network. Like for chemostatting, however, there are physical situations where such submatrices are pertinent.

For example, adding or removing reactions as an operation describes the effect of adding and removing catalytic species in a network. For a heterogeneous catalyst on a solid substrate, it suffices to remove the substrate from the medium. The search for the controlled introduction, activation, removal and modification of homogeneous catalysts in solution is an ongoing endeavor. An interesting recent example makes uses of an optical protocol to convert a chiral enantioselective catalyst to its mirror image[24]. Such developments highlight the diversity of new situations where submatrices can come into play.

Let us again consider the fundamental theorem (2.105) and apply $r' = r - 1$. Keeping the number of species fixed ($s' = s$), it follows that

$$c' - l' = c - l - 1. \tag{2.143}$$

We either have i) a broken cycle $c' = c - 1$ or ii) an emergent conservation law $l' = l + 1$, such that

$$r^Y = c^\times + \ell^* \tag{2.144}$$

with $c^\times$ the number of broken cycles and ℓ^* the number of emergent conservation laws. It forms the natural counterpart of Eq. (2.114). Since the complex networks share the same reactions, we can perform an analogous operation while preserving the number of complexes σ, to find

$$r^Y = \gamma^\times + \lambda^* \tag{2.145}$$

With $\gamma^\times$ the number of broken complex cycles and λ^* the number of emergent linkage classes. Such an operation, however, should create no isolated complexes (complexes engaging in no reaction) to ensure that one can properly use the framework.

2.6.1 Chain Reactions

The chain reaction system in (2.125) is supplied with an initiation reaction

$$\mathrm{Br_2} \overset{1}{\rightleftharpoons} 2\mathrm{Br^\bullet}, \tag{2.146}$$

which provides the first radicals to perform the reaction. Alternatively (and often the case), such a reaction may not be inherent to the chemistry, and the initiation must be performed with other species (e.g. a free radical initiator). At the heart of this chain reaction is that it regenerates its intermediates from the reactant pool while converting $\mathrm{Br_2} + \mathrm{H_2}$ to $2\mathrm{HBr}$

$$\mathrm{Br^\bullet} + \mathrm{H_2} \overset{2}{\rightleftharpoons} \mathrm{HBr} + \mathrm{H^\bullet}, \quad \mathrm{H^\bullet} + \mathrm{Br_2} \overset{3}{\rightleftharpoons} \mathrm{HBr} + \mathrm{Br^\bullet}. \tag{2.147}$$

Via the emergent cycle $\boldsymbol{c}^* = (0,1,1)$, which yields the detailed reaction (using Eq. (2.54))

$$\mathrm{Br_2} + \mathrm{H_2} + \mathrm{Br^\bullet} + \mathrm{H^\bullet} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} 2\mathrm{HBr} + \mathrm{Br^\bullet} + \mathrm{H^\bullet}. \tag{2.148}$$

Note that in our framework, we can clearly distinguish this cycle from a single reaction. The nonambiguity condition (Eq. (2.15)) forbids a species to fulfill the role of both a reactant and a product in a single reaction. A transformation can thus only acquire the functional form of Eq. (5.10) by being composed of at least two different reactions, which is also the minimal requirement for a cycle.

A characteristic of a chain reaction is that it maintains itself by regenerating its intermediates. Consequently, the reactions involved in the chain reaction must satisfy a conservation law that ensure these species are maintained. To see this, consider the submatrix without initiation reaction and only radical species, $\bar{\boldsymbol{\nu}} = \boldsymbol{\nu}(1|1,3,5)$ (shown in red)

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{Br_2} \\ \mathrm{Br^\bullet} \\ \mathrm{H_2} \\ \mathrm{H^\bullet} \\ \bar{\boldsymbol{\nu}}\ \ \mathrm{HBr} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 2 & -1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \end{array} \tag{2.149}$$

The submatrix $\bar{\boldsymbol{\nu}}$ admits an emergent cycle $\boldsymbol{c}^* = (1,1)^T$, and has an emergent conservation law $l^* = (1,1)$. The same result is found on the level of complexes, since the unimolecular species and reaction yield the equivalent incidence matrix

$$\begin{array}{cc} & \begin{array}{cc} 2 & 3 \end{array} \end{array} \tag{2.150}$$

$$\bar{\boldsymbol{\partial}} = \begin{array}{c} \mathrm{Br^\bullet} \\ \mathrm{H^\bullet} \end{array} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{2.151}$$

which means $\boldsymbol{c}^*$ is also a complex cycle $\boldsymbol{\gamma}^*$. Now, let us add the reaction 1, which in the absence of $\mathrm{Br_2}$ behaves like a bimolecular reservoir exchange

$$\emptyset \rightleftharpoons 2\mathrm{Br^\bullet}. \tag{2.152}$$

The resulting stoichiometric matrix is then

$$\begin{array}{cc} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \end{array} \tag{2.153}$$

$$\bar{\boldsymbol{\nu}}_\circ = \begin{array}{c} \mathrm{Br^\bullet} \\ \mathrm{H^\bullet} \end{array} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{2.154}$$

By adding the initiation reaction 1, the conservation ℓ^* is broken and the emergent cycle now takes the form $\boldsymbol{c}^* = (0,1,1)^T$. A key property of the chain-reaction, mass-like conservation of the reactive species, has been obscured by extending the network. In Ch. 5 we will return to this property, when we define the concept of stoichiometric allocatalysis.

2.6.2 Catalysis

Colloquially speaking, catalysts are chemical compounds, which accelerate a chemical reaction while not being consumed themselves in the overall reaction. The various steps in this process form a catalytic cycle. Let us here investigate what that looks like.

It is common to represent a catalytic cycle by its net reaction in the sense of Eq. (2.47). For instance, if there is a single catalyst E, substrate S and product P, standard shorthand notations for such a catalytic cycle are:

$$\mathrm{S} + \mathrm{E} \rightleftharpoons \mathrm{E} + \mathrm{P}, \tag{2.155}$$

$$\mathrm{S} \overset{\mathrm{E}}{\rightleftharpoons} \mathrm{P}. \tag{2.156}$$

These notations are ambiguous on the level of the stoichiometric matrix, since Eq. (2.15) is not satisfied. In order to remove the ambiguity, we turn to a more detailed description with a catalytic pathway and an uncatalyzed reaction

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{S} \\ \mathrm{P} \\ \mathrm{E} \\ \mathrm{ES} \\ \mathrm{EP} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \mathrm{S}+\mathrm{E} \overset{1}{\rightleftharpoons} \mathrm{ES} \\ \mathrm{ES} \overset{2}{\rightleftharpoons} \mathrm{EP} \\ \mathrm{EP} \overset{3}{\rightleftharpoons} \mathrm{E}+\mathrm{P} \end{array} \tag{2.157}$$

Now, this stoichiometric matrix verifies Eq. (2.15).

A convenient reduced description can be obtained by removing the chemical species S and P from $\boldsymbol{\nu}$ to obtain $\boldsymbol{\nu}_X$:

$$\boldsymbol{\nu}_X = \begin{array}{c} \\ \mathrm{E} \\ \mathrm{ES} \\ \mathrm{EP} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array}. \tag{2.158}$$

In the present case, this leads to the emergent cycle $\boldsymbol{c}^* = (1,1,1)^T$. This cycle corresponds to the net reaction

$$\mathrm{S}+\mathrm{E}+\mathrm{ES}+\mathrm{EP} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{E}+\mathrm{ES}+\mathrm{EP}+\mathrm{P}, \tag{2.159}$$

which is built from the stoichiometric matrix $\boldsymbol{\nu}$.

In terms of the stoichiometric matrix $\boldsymbol{\nu}_X$, (resp. $\boldsymbol{\nu}_Y$), we can use Eqs. (2.118) and (2.119) to obtain the net reactions

$$\mathrm{E}+\mathrm{ES}+\mathrm{EP} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{E}+\mathrm{ES}+\mathrm{EP}, \tag{2.160}$$

$$\mathrm{S} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{P}. \tag{2.161}$$

We can now distinguish between the single reaction (2.161) and the composition of reactions in Eq. (2.160). For an arbitrary catalytic cycle, the net reaction should contain catalytic species (E, ES, EP in our example) with the same coefficients on the reactant and product side.

Catalysis maintains its catalytic species in their cyclic conversion and accelerates a reaction (Eq. (2.161)). This implies that we can associate a subnetwork with catalysis, that has a mass-like conservation law $\boldsymbol{l}^+$ for its catalytic species and an emergent cycle $\boldsymbol{c}^*$ that leaves those species unchanged, while performing the net reaction.

For $\boldsymbol{\nu}$ in Eq. (2.158), this is achieved by removing non-catalytic species S and P, and removing the non-catalytic reaction 4, such that $\boldsymbol{\nu}_\star = \boldsymbol{\nu}(4,1,2)$, which gives the matrix

$$\boldsymbol{\nu}_\star = \begin{array}{c} \\ \mathrm{E} \\ \mathrm{ES} \\ \mathrm{EP} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array}. \tag{2.162}$$

which has an emergent cycle $\boldsymbol{c}^* = (1,1,1)^T$ and emergent conservation law $\boldsymbol{\ell}^* = (1,1,1)$. The functional form of the conservation law $\boldsymbol{\ell}^*$ directly reveals that we should find the analogous structure for the complexes, since the incidence matrix verifies $\boldsymbol{\partial}_\star = \boldsymbol{\nu}_\star$. For catalysis by a single

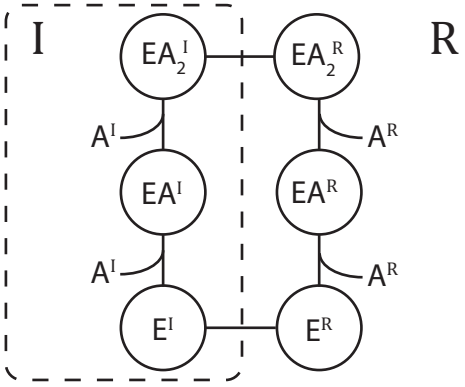


Figure 2.6: A minimalist scheme for a passive nonlinear transport enzyme catalyzing the reaction $2A_I \rightleftharpoons 2A_R$.

catalytic species, we should generally expect that it can be identified by a square matrix with an emergent cycle $\boldsymbol{c}^* = (\boldsymbol{\ell}^*)^T$, for the exact same reason that we can often draw enzymatic cycles as a simple graph (see Fig. 2.6). On the level of the catalyst, reactions simply link various catalytic intermediates or enzyme states, the catalyst hops from one state or intermediate to another. Some authors take this single-catalyst picture as a requirement in their definition of catalysis [25].

In such a picture, the subnetwork of internal catalytic species would only contain unimolecular reactions. This is indeed a very common situation. One can, however, imagine catalytic cycles that do not obey this structure, e.g. because the catalytic cycle contains a bimolecular reaction between two catalytic intermediates.

A catalytic stoichiometric matrix $\boldsymbol{\nu}_\star$ with $\boldsymbol{c}^* = (\boldsymbol{\ell}^*)^T$ has the interesting property that its transpose $\boldsymbol{\nu}_\star^T$ is also a catalytic stoichiometric matrix. Alternatively, one can interpret $\boldsymbol{\nu}_\star^T$ as a dual to $\boldsymbol{\nu}$, where the the columns correspond to species, and rows indicate for a given reaction to which species they are linked. A right nullvector then plays the role of a conservation law, and a left nullvector plays the role of a cycle. Since $\boldsymbol{c}^* = (\boldsymbol{\ell}^*)^T$, these still remain the same.

Bibliography

Articles

[1] Rutherford Aris. “Prolegomena to the Rational Analysis of Systems of Chemical Reactions”. In: *Arch. Ration. Mech. Anal.* 19 (1965), pp. 81–98.

[3] Matteo Polettini and Massimiliano Esposito. “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”. In: *J. Chem. Phys.* 141.2 (2014), pp. –.

[4] Riccardo Rao and Massimiliano Esposito. “Nonequilibrium Thermodynamics of Chemical Reaction Networks : Wisdom from Stochastic Thermodynamics”. In: *Phys. Rev. X* 6 (2016), p. 041064.

[5] Riccardo Rao and Massimiliano Esposito. “Conservation laws and work fluctuation relations in chemical reaction networks”. In: *J. Chem. Phys.* 149.24 (2018).

[7] M. Polettini, A. Wachtel, and M. Esposito. “Dissipation in noisy chemical networks: The role of deficiency”. In: *J. Chem. Phys.* 143.18 (2015), pp. 1–10.

[9] Riccardo Rao and Massimiliano Esposito. “Conservation laws shape dissipation”. In: *New J. Phys.* 20.2 (2018).

[10] R. Aris, P. Gray, and R. Scott. “Modelling cubic autocatalysis by successive bimolecular steps”. In: *Chem. Eng. Sci.* 43 (1988), pp. 207–2011.

[11] G.B. Cook et al. “Bimolecular routes to cubic autocatalysis”. In: *J. Phys. Chem.* 93 (1989), pp. 2749–2755.

[14] Micheal P. Burke and Stephen J. Klippenstein. “Ephemeral collision complexes mediate chemically termolecular transformations that affect system chemistry”. In: *Nat. Chem.* 9 (2017), pp. 1078–1082.

[15] J.W. Verhoeven. “Glossary of terms used in photochemistry (IUPAC Recommendations 1996)”. In: *Pure Appl. Chem.* 68.12 (1996), p. 2223.

[16] F Closa et al. “Identification of two-step chemical mechanisms using small temperature oscillations and a single tagged species”. In: *J. Chem. Phys.* 142 (2015), p. 174108.

[17] A. Lemarchand and B. Nowakowski. “Do the internal fluctuations blur or enhance axial segmentation ?” In: *EPL* 94 (2011), p. 48004.

[18] P. Dziekan et al. “Reaction-diffusion approach to prevertebrae formation : Effect of a local source of morphogen”. In: *J. Chem. Phys.* 139 (2013), p. 114107.

[19] P Dziekan et al. “Effect of a Local Source or Sink of Inhibitor on Turing Patterns Effect of a Local Source or Sink of Inhibitor on Turing Patterns”. In: *Commun. Theor. Phys.* 62 (2014), pp. 622–630.

[20] F Schlögl. “Chemical reaction models for non-equilibrium phase transitions”. In: *Z. Phys.* 253.2 (1972), pp. 147–161.

[21] Matthieu Emond, Le Saux, and Jean-francois Allemand. “Energy Propagation Through a Protometabolism Leading to the Local Emergence of Singular Stationary Concentration Profiles”. In: *Chem. Eur. J.* 18 (2012), pp. 14375–14383.

[22] Thomas Le Saux, Raphaël Plasson, and Ludovic Jullien. “Energy propagation throughout chemical networks”. In: *Chem. Commun.* 50 (2014), pp. 6189–6195.

[23] David Weininger. “SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules”. In: *J. Chem. Inf. Comput. Sci.* 28 (1988), pp. 31–36.

[24] Stefano F. Pizzolato et al. “Bifunctional Molecular Photoswitches Based on Overcrowded Alkenes for Dynamic Control of Catalytic Activity in Michael Addition Reactions”. In: *Chem. Eur. J.* 23.25 (2017), pp. 6174–6184.

[25] Manoj Gopalkrishnan. “Catalysis in Reaction Networks”. In: *Bull. Math. Biol.* 73.12 (2011), pp. 2962–2982.

Books

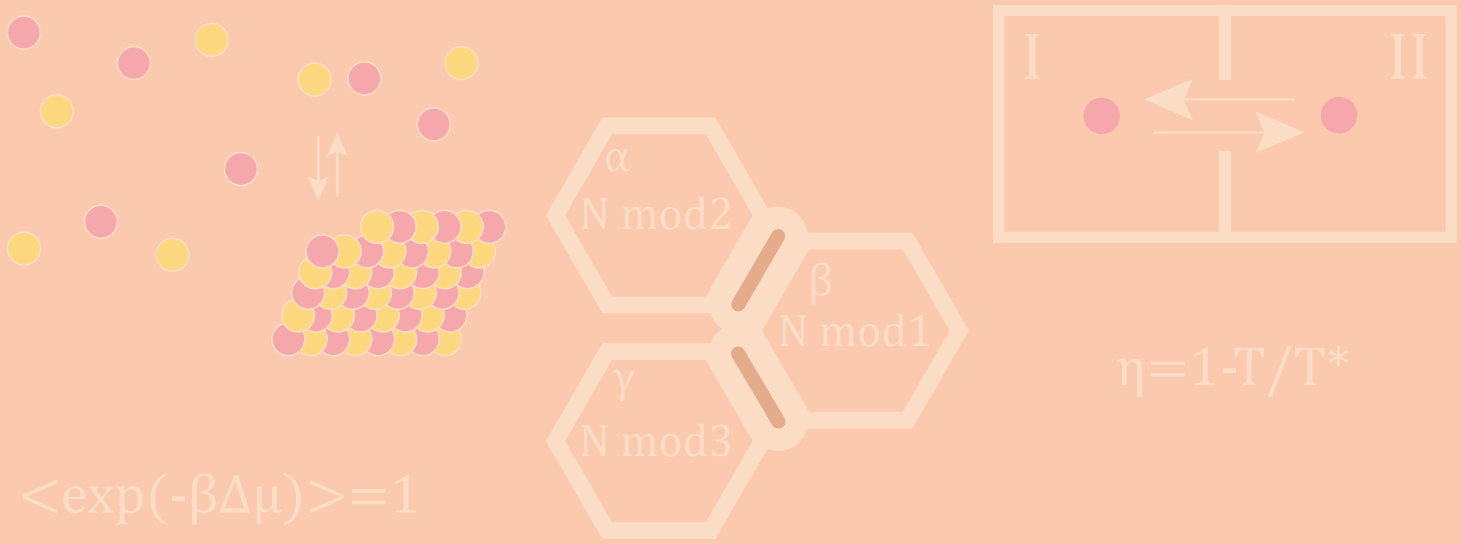
[2] Martin Feinberg. *Foundations of Chemical Reaction Network Theory*. New York: Springer, 2019.

[6] Bernhard Ø. Palsson. *Systems Biology Properties of Reconstructed Networks*. New York: Cambridge University Press, 2005.

[8] Wassily Leontief. *Input-output Economics*. Oxford: Oxford University Press, 1986.

[12] T.W. Graham Solomons. *Organic Chemistry*. Sixth edit. New York: John Wiley & Sons, 1995.

[13] Francis A. Carey and Richard J. Sundberg. *Advanced Organic Chemistry Part A: Structure and Mechanisms*. New York: Springer, 2006.



3. Thermodynamic aspects of open networks

In Origins of Life, one of the few things all scenarios agree on is that systems we consider relevant to abiogenesis are open chemical systems out of equilibrium. Such chemical systems are not confined a minimum free energy state, they can be multistable[1, 2, 3], exhibit chaos [4, 5, 6, 7], form spatial patterns [8, 9, 10], proofread synthesis products [11, 12, 13, 3, 14], make persistent copies of information-carrying molecules [15, 16] and so forth. It is important to realize, however, that there are various ways to have an open (sub)system and these are not equivalent. In this section, we will cover some common examples of open systems.

In experimental setting, we can e.g. introduce chemostats (e.g. buffers, reservoirs), continuously stirred tank reactors (CSTR) or serial transfer to make our subsystem of interest open to matter exchange. It is important to realize that these are fundamentally different ways of opening a system and the exact way a system is opened has important consequences for the system behavior. To give but some examples: Polymer length distributions are strongly affected by the presence and nature of a dilution process. The same is true for chemical evolution, cooperation and the nucleation of autocatalytic cycles. Conservation laws are absent in a flow reactor, but a fixed mass-influx balanced by an outflux yields a somewhat related set of constraints. For chemostats, conservation laws are much more flexible, leading e.g. to unbalanced growth phases (See Ref.[17] and Sec. 7.3.3). Serial transfer has all the conservation laws of a closed system, right until the next transfer.

A growing and dividing cell can live and interact with complex environments, in a manner that is not exactly described as any single one of these. Indeed, a spatially heterogeneous environment may have local patches behaving like chemostats, areas with rapid flows, zones subject to periodic changes (tides, waves, day-night cycles of evaporation) and this environment will over time be modified by the chemistry subject to these ways of opening the system.

In discussions of prebiotic chemistry in open systems, it is therefore important to specify how we open the system and understand what that choice entails and what limitations it introduces. A number of interesting results in the field can be understood much more readily by having such a perspective [18, 19]. To that aim, this section will discuss chemostats (Sec. 3.1), CSTRs (3.2), serial transfer (Sec. 3.3) and coupled growing compartments (Sec. 3.4). The sections on CSTRs and serial transfer (Sec. 3.2, 3.3) are largely based on Ref. [20], which can be read for a more

detailed discussion. Following experiments in the LBC lab, a publication is in preparation, of which Sec. 3.4 and Appendix 10.2 provide some of the underpinnings.

3.1 Chemostats

3.1.1 Thermodynamic Chemostats

We define a thermodynamic chemostat* as a large equilibrium bath of a chemical species that can be exchanged with the system. Physically, such baths can typically be thought to exchange a species A according to a stoichiometric process

$$\mathrm{A}^{system} \rightleftharpoons \mathrm{A}^{bath}. \tag{3.1}$$

By equilibration of the exchange process, the corresponding $\Delta\mu$ vanishes

$$\Delta\mu = \mu_{\mathrm{A}}^{bath} - \mu_{\mathrm{A}}^{system} = 0, \tag{3.2}$$

and thus $\mu_{\mathrm{A}}^{bath} = \mu_{\mathrm{A}}^{system}$. The bath then fixes the chemical potential of A in the system. For dilute solutions, this chemical potential is considered a function of concentration of the form

$$\mu_{\mathrm{A}} = \mu_{\mathrm{A}}^{\circ} + k_b T \ln x_{\mathrm{A}}, \tag{3.3}$$

which then fixes the molar fraction $x_{\mathrm{A}}^{bath} = x_{\mathrm{A}}^{system}$. This allows for the very practical interpretation that a chemostat fixes a concentration or molar fraction[21], in the way a thermostat fixes a temperature. These are more intuitive quantities to work with and this approximation will be adopted throughout this manuscript.

More generally, however, μ can depend on other factors, such as the ionic strength of the solvent. Consequently, if the system has a different ionic strength than the chemostat, or in some other way has its chemical potential altered, we still expect Eq. 3.1 to hold, although now $x_{\mathrm{A}}^{bath} \neq x_{\mathrm{A}}^{system}$. What remains fixed is the chemical potential.

Another remark that should be considered, is that the origin of the coupling between a system and its chemostat may be more elaborate than a simple unimolecular exchange with a bath. For example, consider a typical antiporter and symporter, whose transport catalysis is described by a reaction vector $\boldsymbol{g}$:

$$\mathrm{A}^{system} + \mathrm{B}^{bath} \underset{-\boldsymbol{g}}{\overset{\boldsymbol{g}}{\rightleftharpoons}} \mathrm{A}^{bath} + \mathrm{B}^{system}, \qquad \Delta\mu = \ln\left(\frac{x_a^{bath} x_b^{system}}{x_a^{system} x_b^{bath}}\right), \tag{3.4}$$

$$\mathrm{A}^{system} + \mathrm{B}^{system} \underset{-\boldsymbol{g}}{\overset{\boldsymbol{g}}{\rightleftharpoons}} \mathrm{A}^{bath} + \mathrm{B}^{bath}, \qquad \Delta\mu = \ln\left(\frac{x_a^{system} x_b^{system}}{x_a^{bath} x_b^{bath}}\right). \tag{3.5}$$

The equilibration of a single exchange process introduces one constraint, and as such for $\Delta\mu = 0$ we respectively equilibrate the product

$$x_a^{bath} x_b^{system} = x_a^{system} x_b^{bath}, \tag{3.6}$$

for an antiporter, and

$$x_a^{system} x_b^{system} = x_a^{bath} x_b^{bath}. \tag{3.7}$$

*A thermodynamic chemostat should not be confused with a biochemical reactor containing growing microorganisms, also commonly called a chemostat. In this reactor, fresh medium and sterile air is supplied in a stirred microbial culture. The influx is compensated by an equal outflux of this mix. Such a reactor corresponds to what we will call a CSTR, and the term chemostat will from hereon be reserved to designate chemical baths.

for a symporter.

Unless a second constraint is introduced (e.g. another linearly independent transport process involving A and/or B), we generally have $x_{\mathrm{A}}^{system} \neq x_{\mathrm{A}}^{bath}$, $x_{\mathrm{B}}^{system} \neq x_{\mathrm{B}}^{bath}$. Discrete conserved quantities like atomic matter present a rich variety of constraints in their networks, that distinguishes them from a continuous quantity like temperature.

This behavior is not restricted to transport enzymes. Many chemical species may not exchange by themselves, but only by engaging with other species. For example, metal ions in water are not expected to dissolve well in oil, hindering their exchange with another water phase. Some metal ion complexes (e.g. acetates or diacetyl complexes) however, dissolve quite well, enabling cotransport of the anionic ligands and the metal.

Transport enzymes may exchange an integer multiple of a compound at a time, some species may form soluble dimers that exchange well. These chemostats do not impose the same type of constraints as process (3.1), and this has direct consequences on the validity of the zeroth law of thermodynamics. A detailed discussion is provided in Appendix10.1.

3.1.2 Statistical aspects of chemostats

As a first exploration of chemostats, let us consider a large box, containing m equally sized compartments, linked through small holes (see Fig. 3.1). The compartments contain a chemical species A, either as a gas or a dilute solution. Interactions are considered negligible and the number of particles that can be contained in a compartment is unconstrained. The timescale of mixing τ_{mix} within the boxes is much faster than the timescale of exchange through holes τ_{ex}, such that we can consider the exchange between boxes as exchange between perfectly mixed reactors. Now, let us

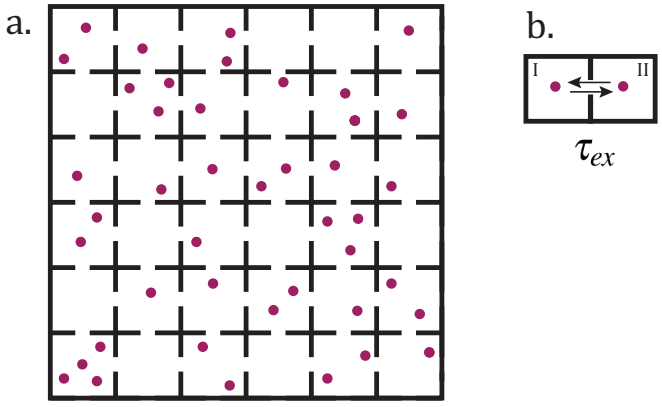


Figure 3.1: a) A large box with $m = 36$ equally sized partitions. b) exchange between two partitions

label the number of A molecules in compartment i as n_{A}^i, and let total number of A molecules be N_{A}

$$N_{\mathrm{A}} = \sum_i n_{\mathrm{A}}^i. \tag{3.8}$$

For noninteracting A molecules, the degeneracy of a particular configuration (see also Sec. 4.1.1) $\boldsymbol{n} = \{n_{\mathrm{A}}^1, n_{\mathrm{A}}^2, ..., n_{\mathrm{A}}^m\}$ is then a multinomial

$$z(\boldsymbol{n}) = \binom{N_{\mathrm{A}}}{n_{\mathrm{A}}^1, n_{\mathrm{A}}^2, ..., n_{\mathrm{A}}^m} = \frac{N_{\mathrm{A}}!}{\Pi_{i=1}^m n_{\mathrm{A}}^i!} \tag{3.9}$$

and the total number of states Z is

$$Z = \sum_{\boldsymbol{n}} z(\boldsymbol{n}) = m^{N_{\mathrm{A}}}. \tag{3.10}$$

The probability for a particular configuration $\boldsymbol{n}$ follows a multinomial law

$$p(\boldsymbol{n}) = \binom{N_\mathrm{A}}{n_\mathrm{A}^1, n_\mathrm{A}^2, ..., n_\mathrm{A}^m} \left(\frac{1}{m}\right)^{n_\mathrm{A}^1} ... \left(\frac{1}{m}\right)^{n_\mathrm{A}^m} = \frac{z(\boldsymbol{n})}{Z} \tag{3.11}$$

The probability to observe $n_A^1 = k$ then becomes

$$p(n_A^1 = k) = \sum_{n_\mathrm{A}^2,....,n_\mathrm{A}^m} \frac{\binom{N_A}{k}\binom{N_A-k}{n_\mathrm{A}^2,...,n_\mathrm{A}^m}}{Z} = \binom{N_A}{k} \frac{(m-1)^{N_A-k}}{m^{N_A}}. \tag{3.12}$$

This quantity is normalized as expected

$$\sum_{k=0}^{N_\mathrm{A}} p(n_A^1 = k) = \left(\frac{m-1}{m}\right)^{N_\mathrm{A}} \sum_{k=0}^{N_\mathrm{A}} \binom{N_A}{k} (m-1)^{-k} = 1. \tag{3.13}$$

Introducing $D = \left(\frac{m-1}{m}\right)^{-N_\mathrm{A}}$, we can then write the average

$$\langle n_A^1 \rangle = \frac{1}{D} \sum_{k=0}^{N_\mathrm{A}} k \binom{N_A}{k} (m-1)^{-k} = \frac{1}{D} \sum_{k=1}^{N_\mathrm{A}} N_A \binom{N_A-1}{k-1} (m-1)^{-k}, \tag{3.14}$$

where we used that $N_A\binom{N_A-1}{k-1} = k\binom{N_A}{k}$ and that the $k=0$ contribution is 0. Introducing $k' = k-1$, we can now write

$$\langle n_A^1 \rangle = \frac{1}{D} \frac{N_A}{m-1} \sum_{k'=0}^{N_\mathrm{A}-1} \binom{N_A-1}{k'} (m-1)^{-k'} = \frac{N_\mathrm{A}}{m}, \tag{3.15}$$

where we used the same sum as in the normalization (3.13). We will refer to this average using the shorthand quantity

$$\lambda = \frac{N_\mathrm{A}}{m}, \tag{3.16}$$

which has the interpretation of mean number of A species per compartment.

Let us now consider a bimolecular process that can occur within a cell, of the form

$$2\mathrm{A} \rightleftharpoons \mathrm{A}_2. \tag{3.17}$$

This reaction is slow with respect to mixing and exchange and proceeds with a forward rate

$$R \propto n_A(n_A - 1) \tag{3.18}$$

The average rate of the forward reaction is then $\langle n_A^1(n_A^1 - 1)\rangle$, which is found to be

$$\begin{aligned} \langle n_A^1(n_A^1-1)\rangle &= \frac{1}{D}\sum_{k=0}^{N_\mathrm{A}} k(k-1)\binom{N_A}{k}(m-1)^{-k} \\ &= \frac{N_\mathrm{A}(N_\mathrm{A}-1)}{D(m-1)^2}\sum_{k'=0}^{N_\mathrm{A}-2}\binom{N_A-2}{k}(m-1)^{-k'} = \frac{N_A(N_A-1)}{m^2}. \end{aligned} \tag{3.19}$$

Or in terms of λ:

$$\langle R_{open}\rangle \propto \lambda^2 - \frac{\lambda}{m}. \tag{3.20}$$

Had we taken a single, closed compartment with exactly the same number λ of molecules, we would have found a rate

$$\langle R_{close} \rangle \propto \lambda(\lambda - 1), \tag{3.21}$$

which is strictly smaller for $m > 1$. When higher moments than the mean come into play, the different statistics for open and closed systems come to expression. Here, this is reflected in the reaction rates.

For small λ, R_{open} can become considerably larger than R_{close}. In particular, when $\lambda = 1$, the closed system will proceed at a zero rate, but the open system will still spend some time in states with $k \geq 2$.

We are typically interested in baths of chemicals that are 'infinite' with respect to the subsystem: $N_\mathrm{A}, m \to \infty$. In this regime, the ongoing injection or reception of large quantities of A will correspond to minor relative changes in the number of molecules in the bath, since the bath size is $m-1$ times larger than the subsystem. Let us denote the mth falling factorial as

$$(N)_m = N(N-1)(N-2)...(N-m+1). \tag{3.22}$$

The number of A molecules in a given compartment (Eq. (3.12)) is then described by a Poisson distribution, since

$$\lim_{N_\mathrm{A}, m \to \infty} \frac{(N_\mathrm{A})_k}{k!} \frac{\left(1 - \frac{\lambda}{N_\mathrm{A}}\right)^{N_A - k}}{m^k} = \frac{\lambda^k \exp(-\lambda)}{k!}, \tag{3.23}$$

where we use the limit

$$e^{-x} = \lim_{n \to \infty} \left(1 - \frac{x}{n}\right)^n. \tag{3.24}$$

In this regime, we can neglect $\frac{1}{m}$ corrections. And it is readily found that higher order propensities verify

$$\langle (N)_m \rangle = \lambda^m. \tag{3.25}$$

3.1.3 Examples of real chemostats

Chemostats are an important theoretical tool in thermodynamics, forming the chemical analogue of a thermostat. These chemical reservoirs are as diverse as chemistry itself, the 'ideal' chemostat drawn in Fig. 3.1 is a poor representation of the chemostats we encounter in practice. Partitions with large holes will in principle exchange (hence: chemostat) everything, whereas one is often interested in chemostatting only a subset of components. Here, we will provide a more general picture of what we should call chemostats.

Chemostat, reservoir, or bath can be used interchangeably to designate an abundant source of certain chemicals. Such a source of chemicals can be in a different phase, solvent or form, than encountered in the system, which leads to a considerably variety of ways in which chemostats arise, as illustrated by Fig. 3.2.

Homogeneous chemostats, buffers

Let us first consider homogeneous chemical chemostats, which means chemostats established by certain abundant buffer compounds present in the volume of the system. For example, when a large concentration of acetic acid is present with an equally large concentration of acetate ion at concentration $\bar{c}$, 'free' protons in the solution are chemostatted by the reaction

$$\mathrm{CH_3COOH} + \mathrm{H_2O} \rightleftharpoons \mathrm{CH_3COO^-} + \mathrm{H_3O^+}. \tag{3.26}$$

Where H_3O^+ is a simplified notation for a variety of water-proton complexes. The equilibrium constant of this process is expressed with water absorbed in the constant:

$$K_A = \frac{c_{CH_3COO^-} c_{H_3O^+}}{c_{CH_3COO^-}} = 1.754 \cdot 10^{-5} \text{Mol/L}. \tag{3.27}$$

Since we chose $c_{CH_3COO^-} = c_{CH_3COOH} = \bar{c}$, we have

$$c_{H_3O^+} = K_A. \tag{3.28}$$

Now, suppose a reaction with a base B consumes the free protons according to a reaction

$$B + H_3O^+ \rightleftharpoons BH^+ + H_2O. \tag{3.29}$$

Let this base irreversibly consume c_{BH^+} protons, and let us consider the (approximate) conservation laws for protons and for the acetic acid acetate pair

$$L_1 = c_{BH^+} + c_{CH_3COOH} + c_{H_3O^+} \tag{3.30}$$

$$L_2 = c_{CH_3COOH} + c_{CH_3COO^-} = 2\bar{c}. \tag{3.31}$$

We can now use these values to solve Eq. (3.27) for H_3O^+ with the removed protons

$$c_{H_3O^+} = K_A \frac{\bar{c} - c_{BH^+}}{\bar{c} + c_{BH^+}}. \tag{3.32}$$

We can speak of a chemostat, when this addition of base hardly affects $c_{H_3O^+}$, which means there are a lot more protons in the reservoir than are consumed by the system, which requires $\bar{c} \gg c_{BH^+}$. In that regime, we can linearize Eq. (3.32)

$$c_{H_3O^+} \approx K_A \left(1 - \frac{c_{BH^+}}{\bar{c}}\right)^2 \approx K_A \left(1 - 2\frac{c_{BH^+}}{\bar{c}}\right). \tag{3.33}$$

As long as $\bar{c} \gg c_{BH^+}$, the $\text{pH} = -\log c_{H_3O^+}$ is maintained at a very stable level. In biology, this is of key importance. Although the volume of e.g. an E. Coli cell typically contains $\approx 80 - 100$ protons [22] they are rapidly exchanged with millions of weakly acidic and basic groups, thus preventing large fluctuations and storing a considerably quantity of protons in buffering organic acids.

Another example of an important homogeneous chemostat in biology is glutathione (GSH) and its dimer (GSSG), which forms a reservoir for the extraction or donation of electrons. These are important, e.g. to form or break disulfide bonds in peptides, but also to break down reactive oxygen species[22]

$$2GSH + RSSR' \rightleftharpoons GSSG + RSH + R'SH. \tag{3.34}$$

In principle, any appropriately chosen equilibrium reaction can function as a homogeneous chemostat for a compound. To have a high buffering capacity, this equilibrium should disfavor the compound, so that the chemostat is abundant and can withstand large fluctuations in the compound consumption/production.

An inherent limitation to homogeneous chemostats is that the buffering capacity is bounded by system size. As a consequence, the consumption of chemostatted compounds of interest is limited to this size. In addition, chemostatting species at a high concentration comes at the cost of decreasing their buffering capacity.

Growing and replicating systems have material requirements that are not covered by homogeneous chemostats. In fact, many of their homogeneous chemostats, needed for proper functioning, need to be replicated and replenished as well. To satisfy these material requirements, contact with external reservoirs is required. These reservoirs are not inherently limited by the system that feeds on it, which provides an essential distinction.

External reservoirs

In Fig. 3.2, a number of external reservoirs are considered. The compound of interest can e.g. be a gas dissolving in a solvent, adsorbing on a surface, or dissociating to enter a metal. It can also be a solid, e.g. a salt with a solubility product K_{sp} or a precipitate of a compound with a solubility K_s. The external reservoir may also be a similar phase, separated by a pore, air, a membrane, a different solvent, etc. In these situations, it is very natural to have 'composite chemostats', which we define below.

Definition 3.1.1 — Composite Chemostat. A composite chemostat is a chemostat that exchanges multiple species with the system in a stoichiometrically coupled fashion

$$\emptyset \rightleftharpoons \sum_k n_k^{(-)} \mathrm{X_k}, \tag{3.35}$$

thereby fixing the combined chemical potential $\bar{\mu}$ of these species in the system:

$$\bar{\mu} = \sum_k n_k^{(-)} \mu_k \tag{3.36}$$

Composite chemostats can exchange reactants like symporters and antiporters do in biology. However, considerably more prevalent and primitive examples exist. As an example of a symporter, a mineral like AgCl can slightly dissolve in the form of Ag^+ and Cl^-

$$\mathrm{Ag^+ + Cl^- \rightleftharpoons AgCl\downarrow} \tag{3.37}$$

The AgCl salt can be of arbitrary size and function as a chemostat, fixing the chemical potentials through

$$\mu_{\mathrm{Ag^+}} + \mu_{\mathrm{Cl^-}} = \mu_{\mathrm{AgCl}}, \tag{3.38}$$

$$k_b T \ln x_{\mathrm{Ag^+}} x_{\mathrm{Cl^-}} = -\Delta\mu^\circ. \tag{3.39}$$

Note that there is no molar fraction term for AgCl, which is in a separate solid phase. Eq. (3.39) is often tabulated as a solubility product, with a constant K_{sp}

$$x_{\mathrm{Ag^+}} x_{\mathrm{Cl^-}} = K_{sp} = \exp(-\beta\Delta\mu^\circ). \tag{3.40}$$

At equilibrium, the product of Ag^+ and Cl^- fractions does not exceed K_{sp}, and any excess precipitates to form AgCl. AgCl therefore acts as a reservoir with a symporter coupling. Since salts have, by definition, cations balanced by anions, their capacity as symporter reservoirs is a general feature.

It is important to consider that composite chemostats do not fix a single concentration. This has consequences for the dynamics of a system: a disproportionate consumption of Cl^- will increase the abundance of Ag^+ and lower that of Cl^-, which means we cannot neglect the dynamics of 'food' from the environment.

In practical situations, one can often get a composite chemostat to mimic a simple one, by making a species that is not consumed abundant. An example of this is given by the homogeneous chemostat (3.26) for H_3O^+, where a high concentration of CH_3COO^- was provided to yield an approximate simple chemostat.

Ions also provide common examples for antiporter reactions. Ion-exchange resins and charged soil particles can initially be loaded with a certain set of ions, which are subsequently displaced by other ones, often with a higher affinity, leading to effective reactions such as

$$\mathrm{Ca^{2+}}(system) + \mathrm{Mg^{2+}}(bath) \rightleftharpoons \mathrm{Ca^{2+}}(bath) + \mathrm{Mg^{2+}}(system), \tag{3.41}$$

$$2\mathrm{Al^{3+}}(system) + 3\mathrm{Mg^{2+}}(bath) \rightleftharpoons 2\mathrm{Al^{3+}}(bath) + 3\mathrm{Mg^{2+}}(system). \tag{3.42}$$

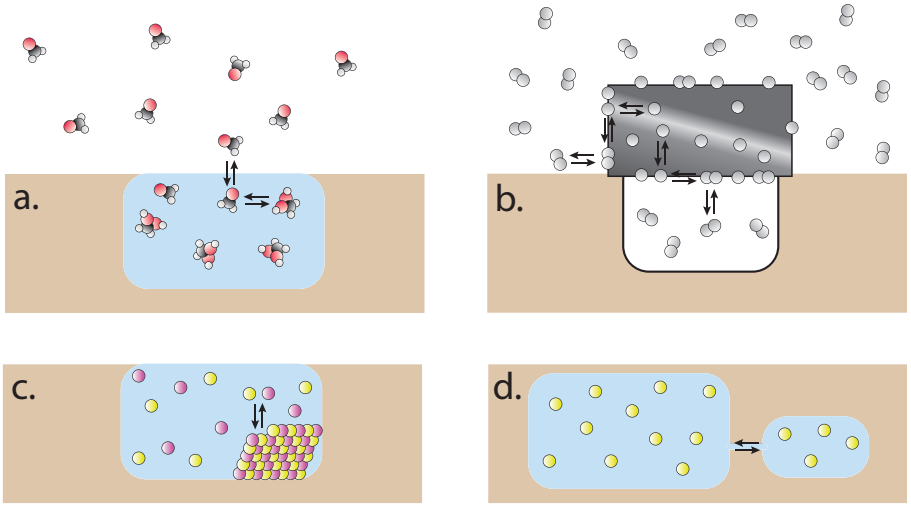


Figure 3.2: Examples of chemostatted systems: a) Gaseous formaldehyde CH_2O exchanging with dissolved formaldehyde, which reversibly incorporates water H_2O to form gem-diol $CH_2(OH)_2$. b) H_2 adsorbs on metal surface, splits in atoms. H atoms enter the metal, and enter the bottom compartment after forming H_2 again. c) Na^+ and Cl^- ions, precipitating to form NaCl. d) Exchange of particles between a reservoir and a compartment through a pore.

which respectively chemostat the quantities $x_{Ca^{2+}}/x_{Mg^{2+}}$ and $x^2_{Al^{3+}}/x^3_{Mg^{2+}}$.

Another common reservoir for a solution is the gas phase that typically covers it, e.g. the reaction

$$O_2(aq) \rightleftharpoons O_2(g). \tag{3.43}$$

Whose equilibrium is characterized by Henry's law constant H^{cp}

$$\frac{c_{O_2}}{p_{O_2}} = H^{cp} = 1.2 \cdot 10^{-5} \text{mol/m}^3\text{Pa}. \tag{3.44}$$

Such reactions are important in atmospheric chemistry where a wide variety of conventions have been used for expressing equilibria of the form (3.43), depending on a preference for units. In the very extensive compilation by R. Sander [23], these are discussed in detail and converted to a common H^{cp}.

To assess the buffering capacity of a gas, it is instructive to express it in the form $N_{O_2(g)}/N_{O_2(aq)}$. Suppose we have a fixed container with size $V = V_{system} + V_{reservoir}$. Using the ideal gas, law, we then have at equilibrium:

$$N_{O_2(aq)}/N_{O_2(g)} = \frac{V_{system}}{V_{reservoir}} \frac{H^{cp}}{RT}. \tag{3.45}$$

Which must be complemented with the conservation law

$$L_{O_2} = N_{O_2(aq)} + N_{O_2(g)}, \tag{3.46}$$

To find the actual amounts of O_2.

Values of H^{cp} vary dramatically as a function of molecular size, charge, and functional groups, and when interpreting a literature value of H^{cp}, one should keep in mind the dynamic equilibria that affect it. For formaldehyde (Fig. 3.2), values around $H^{cp} = 32$ have been reported, but also

$H^{cp} = 0.028 - 0.042$. The former is surprisingly high for such a small aldehyde, compared to the slightly larger acetaldehyde (CH_3CHO) with a reported value[23] of $H^{cp} = 0.13$.

However, formaldehyde readily performs hydration to form a much less volatile gem-diol:

$$H_2CO + H_2O \rightleftharpoons CH_2(OH)_2, \tag{3.47}$$

and over 99.9% of formaldehyde takes this form[24]. The higher value of $H^{cp} = 32$ can now readily be understood as a coarse-grained equilibrium constant for $p(H_2CO)/(c_{H_2CO} + c_{CH_2(OH)_2})$. For acetaldehyde, only around 50% is found in a gem-diol form[24].

Another common situation where caution is required is for organic acids and bases. Ammonia NH_3 is in equilibrium with its protonated ammonium form $NH_4{}^+$

$$NH_3 + H_3O^+ \rightleftharpoons NH_4{}^+ + H_2O. \tag{3.48}$$

As can be expected from common experience, ammonia is very volatile:

$$\frac{c_{NH_3}}{p_{NH_3}} = H^{cp}(NH_3) = 5.9 \cdot 10^{-1} \text{mol/m}^3\text{Pa}. \tag{3.49}$$

For most ionic compounds, the equilibrium is drastically shifted to the aqueous solution and it is typical to completely neglect their evaporation. The fact that an ammonia solution at low pH is odorless is a testament to this fact. A similar situation occurs with acetic acid ($H^{cp} = 40$), which loses its characteristic vinegar smell at high pH, because the reaction (3.26) is shifted to the nonvolatile acetate CH_3COO^-. It should be kept in mind that many reported Henry's law constants are effective constants, measured at a particular pH and subject to dynamic equilibria.

The present discussion covered a small portion of all the thermodynamic chemostats that can be found. In Appendix 10.1, we will develop their treatment in more detail using stochastic thermodynamics. In doing so, we show that some formulations of the zeroth law of thermodynamics can be broken for conserved integer quantities. We then propose a formulation which is not broken.

3.2 Continuously stirred tank reactor (CSTR)

A CSTR (Continuously stirred tank reactor) is a reactor containing a well-stirred solution, in which new reactants are supplied by a constant flow, while the solution volume is kept constant by a compensating outflow [25, 26, 27, 28, 29]. To better distinguish it from a batch reactor that is closed but stirred, it is sometimes referred to as a CFSTR (continuous-flow stirred tank reactor) to underline that there is a constant flow. In keeping with convention, the more commonly used term CSTR will be employed here. The contents of this section, and the following one on serial transfer, are largely a reproduction of earlier published work that can be found in Ref. [20].

3.2.1 Kinetic equations of the CSTR

Continuous-flow stirred tank reactors are open reactors with a continuous feed of reactants and an outflow in order to keep the volume constant inside the reactor (see Fig. 3.3). The reactants are pumped into the reactor at given controlled concentrations $c_{k,\text{in}}$. The solution in the reactor is stirred well so that the concentrations of the different species can be supposed to remain uniform inside the volume of the reactor. In order to establish the evolution equations of the concentrations in the CSTR, we use the balance equations of the concentrations c_k in the flow:

$$\partial_t c_k + \boldsymbol{\nabla} \cdot (c_k \mathbf{v} + \mathbf{j}_k) = \sum_i \nu_{ki} w_i \,, \tag{3.50}$$

expressed in terms of the fluid velocity $\mathbf{v}$, the diffusive current density of species k given by Fick's law $\mathbf{j}_k = -D_k \boldsymbol{\nabla} c_k$, the stoichiometric coefficient ν_{ki} of species k in the reaction i, and the rate w_i

of reaction i. The different species are passively advected by the turbulent velocity field $\mathbf{v}$ of the flow. By stirring, the concentrations rapidly become uniform so that the Fickian diffusive current densities are soon negligible $\mathbf{j}_k \simeq 0$. Integrating the balance equation (Eq. (3.50)) over the volume V of the reactor, we find

$$\int_V \partial_t c_k \, dV + \int_{\partial V} c_k \mathbf{v} \cdot d\mathbf{A} = \int_V \sum_i \nu_{ki} w_i \, dV \,, \tag{3.51}$$

where $d\mathbf{A}$ is the surface element of integration on the border ∂V of the volume V. The surface integral has contributions from the inflow tube of species k entering with concentration $c_{k,\mathrm{in}}$ and the outflow tube where the species k exits at the uniform concentration c_k resulting from stirring:

$$\int_{\partial V} c_k \mathbf{v} \cdot d\mathbf{A} = \int_{\partial V_{k,\mathrm{in}}} c_k \mathbf{v} \cdot d\mathbf{A} + \int_{\partial V_{\mathrm{out}}} c_k \mathbf{v} \cdot d\mathbf{A} \,. \tag{3.52}$$

Since the concentrations can be supposed to be uniform at entry and exit, we get

$$\int_{\partial V} c_k \mathbf{v} \cdot d\mathbf{A} = -\phi_{k,\mathrm{in}} \, c_{k,\mathrm{in}} + \phi_{\mathrm{out}} \, c_k \tag{3.53}$$

in terms of the ingoing flux $\phi_{k,\mathrm{in}} = \int_{\partial V_{k,\mathrm{in}}} \mathbf{v} \cdot d\mathbf{A}$ of the solution in the tube bringing species k into the reactor and the exit flux $\phi_{\mathrm{out}} = \int_{\partial V_{\mathrm{out}}} \mathbf{v} \cdot d\mathbf{A}$ of the stirred solution. These fluxes are in units of m^3 per second, and depend on the section areas of the injection and exit tubes. The volume of the solution inside the reactor being preserved, we have that $\phi_{\mathrm{out}} = \sum_k \phi_{k,\mathrm{in}}$. Since the concentrations are uniform inside the reactor, Eq. (3.51) divided by the volume V becomes

$$\frac{dc_k}{dt} = \sum_i \nu_{ki} w_i + \frac{1}{\tau}(c_{k0} - c_k) \,, \tag{3.54}$$

where $\tau \equiv V/\phi_{\mathrm{out}}$ is the mean residence time of the species inside the reactor, and

$$c_{k0} \equiv \frac{\phi_{k,\mathrm{in}}}{\phi_{\mathrm{out}}} \, c_{k,\mathrm{in}} \tag{3.55}$$

are the injected concentrations of reactants reported to the whole volume. Both the residence time τ and the injected concentrations c_{k0} are control parameters.

The evolution equations for the concentrations form a set of ordinary differential equations, which are typically nonlinear. In the limit where the residence time becomes very long, the last term of Eq. (3.54) becomes negligible and we recover the kinetic equations in a closed reactor, the so-called batch reactor, [29] in which case the concentrations will sooner or later reach their equilibrium value. In the other limit where the residence time is very short, the last term dominates so that the concentrations remain nearly equal to their value at injection: $c_k \simeq c_{k0}$. In between, the concentrations may manifest a rich variety of different stationary, oscillatory, or chaotic behaviors in some autocatalytic or cross-catalytic reaction networks[26, 27, 4, 28, 29].

3.2.2 Thermodynamics of a CSTR

A typical CSTR is functioning under atmospheric pressure and at room temperature if the reactions are not too exothermic. Under these conditions, the relevant thermodynamic potential is Gibbs' free energy G. We assume local thermodynamic equilibrium for every element of the solution and consider the free energy density:

$$g_V = \sum_k \mu_k \, c_k \,, \tag{3.56}$$

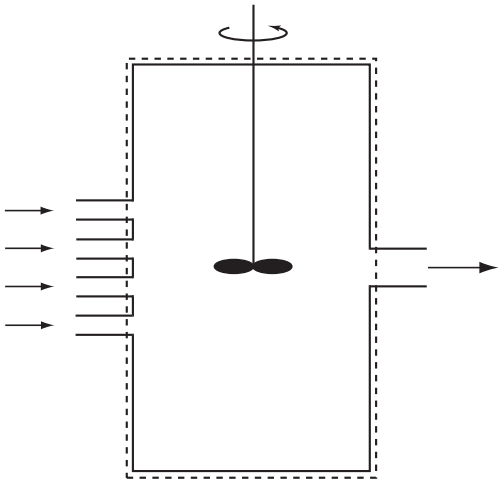


Figure 3.3: Schematic representation of a continuous-flow stirred tank reactor (CSTR). The dashed line depicts a fictitious surface delimiting the volume V of the reactor.

where μ_k is the chemical potential of species k.

Using Eq. (3.56) together with Gibbs' fundamental relation per unit volume

$$dg_V = -s_V dT + dP + \sum_k \mu_k dc_k \,, \tag{3.57}$$

where s_V is the entropy density, T the temperature, and P the pressure, one obtains the Gibbs-Duhem relation

$$s_V dT - dP + \sum_k c_k d\mu_k = 0 \,. \tag{3.58}$$

Using Eq. (3.58) under isothermal and isobaric conditions, one finds that

$$\sum_k c_k d\mu_k = 0. \tag{3.59}$$

Since the solution is well stirred, it is quasi homogeneous in the bulk of the tank, and the time evolution of the Gibbs free energy follows that of the concentrations of the various species. Using Eqs. (3.56)-(3.59), one obtains

$$\frac{dg_V}{dt} = \sum_k \mu_k \frac{dc_k}{dt}, \tag{3.60}$$

Now, using Eq. (3.54) for the concentrations, the time evolution of the free energy density then becomes

$$\frac{dg_V}{dt} = \sum_{ki} \mu_k \nu_{ki} w_i + \frac{1}{\tau} \sum_k \mu_k \left(c_{k0} - c_k\right) . \tag{3.61}$$

According to the mass action law,[28] the reaction rates are proportional to the concentrations of all the species entering in the reaction. It is convenient to make the distinction between the forward and reversed reactions so that

$$w_{\pm i} = k_{\pm i} \prod_k \left(\frac{c_k}{c^0}\right)^{\nu_{ki}^{(\pm)}} , \tag{3.62}$$

where $k_{\pm i}$ are the rate constants, $\nu_{ki}^{(\pm)}$ the numbers of molecules entering the forward or the reversed reaction, and c^0 the standard concentration of one mole per liter. The stoichiometric coefficient is thus given by $\nu_{ki} = \nu_{ki}^{(-)} - \nu_{ki}^{(+)}$, while $w_i = w_{+i} - w_{-i}$. In a dilute solution, the chemical potentials of the solute species are given by $\mu_k = \mu_k^0 + RT \ln(c_k/c^0)$ where R is the molar gas constant. Now, the ratio of the rate constants is related to the standard free energy of the reaction according to

$$\frac{k_{+i}}{k_{-i}} = \exp\Big(-\sum_k \frac{\mu_k^0 \nu_{ki}}{RT}\Big)\,. \tag{3.63}$$

The entropy production rate of the reactions is given by

$$\begin{aligned} \sigma &= -\frac{1}{T}\sum_{ki} \mu_k \nu_{ki} w_i, \\ &= R\sum_i (w_{+i} - w_{-i}) \ln \frac{w_{+i}}{w_{-i}} \geq 0\,, \end{aligned} \tag{3.64}$$

which is always non-negative.

Now, combining Eq. (3.64) with Eq. (3.61), the time evolution of the free energy density becomes

$$\frac{dg_V}{dt} = -T\sigma + \frac{1}{\tau}(\gamma_0 - g_V)\,, \tag{3.65}$$

where we have introduced the following quantity

$$\gamma_0 = \sum_k \mu_k c_{k0}. \tag{3.66}$$

In a closed reactor where τ is infinite, the free energy will decrease towards its minimal value. However, in an open reactor where τ is finite, the free energy does not need to reach its minimal value. In this regard, nonequilibrium stationary, oscillatory, or chaotic regimes can be sustained in an open reactor [30, 28].

The term $(\gamma_0 - g_V)/\tau$ in Eq. (3.65) has no definite sign, except in a stationary state where it is equal to the dissipation produced by the chemical reactions and therefore must be positive. In this case, it is sufficient to know the Gibbs free energies of incoming and outgoing chemical species in order to know the dissipation associated with chemical reactions within the reactor.

3.2.3 General properties of reaction networks in a CSTR

Let us now look at reaction networks in a CSTR, using the formalism introduced in Ch.2.

The equations (Eq.3.54) ruling the time evolution of the concentrations can be rewritten in matrix form as follows:

$$\frac{d\mathbf{c}}{dt} = \boldsymbol{\nu}\cdot\mathbf{w} + \frac{1}{\tau}(\mathbf{c}_0 - \mathbf{c})\,, \tag{3.67}$$

In terms of the s-dimensional vectors $\mathbf{c}$ and $\mathbf{c}_0$ of concentrations and injected concentrations, the $r \times s$ matrix $\boldsymbol{\nu}$ of stoichiometric coefficients, and the r-dimensional vector of reaction rates $\mathbf{w}$, where s is the number of species and r the number of reactions in the network.

In the limit $\tau \to \infty$, resupply and dilution are too slow to be noticable compared to other processes and we we recover the case of a closed reactor [31, 21].

In a stationary state, we have $\boldsymbol{\nu}\cdot\mathbf{w} = 0$, implying that $\mathbf{w}$ can be decomposed in the basis of right null eigenvectors $\mathbf{e}_\gamma$, which are called cycles: $\mathbf{w} = \sum_\gamma w_\gamma \mathbf{e}_\gamma$.

Let us now come back to the rank of the stoichiometric matrix, which for a closed reactor was shown to verify

$$\operatorname{rank}(\boldsymbol{\nu}) = r - c = s - l\,, \tag{3.68}$$

where $c = \dim\ker(\boldsymbol{\nu})$ is the number of cycles, and $l = \dim\operatorname{coker}(\boldsymbol{\nu})$ the number of conserved quantities.

In general, a conservation L law can be written as

$$L \equiv \boldsymbol{\ell}\cdot\mathbf{n}\,, \tag{3.69}$$

When there is only a single reactor with fixed volume V, we can instead write a conservation law L' in terms of concentrations $\mathbf{c}$

$$L' \equiv \boldsymbol{\ell}\cdot\mathbf{c}\,, \tag{3.70}$$

with a vector $\boldsymbol{\ell}$ such that

$$\boldsymbol{\ell}\cdot\boldsymbol{\nu} = 0\,. \tag{3.71}$$

In an open reactor where τ is finite, such quantities are no longer conserved. Instead, they converge asymptotically towards their value defined for the injected concentrations:

$$L_0 = \boldsymbol{\ell}\cdot\mathbf{c}_0\,. \tag{3.72}$$

Indeed, applying the vector $\boldsymbol{\ell}$ to Eq. (3.67), we find that

$$\frac{dL}{dt} = \frac{1}{\tau}(L_0 - L)\,, \tag{3.73}$$

the solution of which is given by

$$L(t) = L(0)\,\mathrm{e}^{-t/\tau} + L_0\left(1 - \mathrm{e}^{-t/\tau}\right). \tag{3.74}$$

It is important to emphasize that all conservation laws are broken in a CSTR. However, long-term behavior is constrained by the CSTR counterpart of a conservation law L_0 (Eq. (3.74)), provided the reactor keeps being fed with the same composition $\mathbf{c}_0$.

We can also recover this result using a full stoichiometric matrix $\boldsymbol{\nu}'$ of the CSTR. In this matrix, the reaction network also includes the reactions corresponding to influx and outflux of each of s species. Since our description does not capture what happens outside the confines of the reactor, we can write reactions of the form

$$\emptyset \longrightarrow \mathrm{A} \tag{3.75}$$

for influx, and

$$\mathrm{A} \longrightarrow \emptyset \tag{3.76}$$

for outflux. Since we are presently not concerned with the details of whether $\emptyset$ corresponds to the influx or the outflux, we will in the present approach model these reactions as[†]:

$$\emptyset \rightleftharpoons \mathrm{A}. \tag{3.77}$$

[†]For a more complete thermodynamic description one may wish to treat these separately

Since all s species can flow out of the reactor, there are s such reactions, with rates $(\mathbf{c}_0 - \mathbf{c})/\tau$.

Thus, the total number of reactions becomes $r' = r+s$. The matrix of stoichiometric coefficients can now be extended towards a $r' \times s$ matrix with $r' = r+s$. This means that the new stoichiometric matrix of the CSTR reads

$$\boldsymbol{\nu}' = (\boldsymbol{\nu}, \boldsymbol{I}), \tag{3.78}$$

where $\boldsymbol{I}$ is the identity matrix $s \times s$. Therefore Eq. (3.67) becomes

$$\frac{d\mathbf{c}}{dt} = \boldsymbol{\nu}' \cdot \mathbf{w}', \tag{3.79}$$

with the flow rate $\mathbf{w}' = (\mathbf{w}, \tilde{\mathbf{w}})^{\mathrm{T}}$ a column matrix of dimension $1 \times r'$ with $\tilde{\mathbf{w}} = (\mathbf{c}_0 - \mathbf{c})/\tau$.

In an open reactor, we also get

$$\mathrm{rank}(\boldsymbol{\nu}') = r' - \dim\ker(\boldsymbol{\nu}') = s - \dim\mathrm{coker}(\boldsymbol{\nu}'). \tag{3.80}$$

The number of conserved quantities is now equal to zero $l' = \dim\mathrm{coker}(\boldsymbol{\nu}') = 0$ so that the number of cycles is equal to the number of reactions in the original network: $c' = \dim\ker(\boldsymbol{\nu}') = r$. Therefore, there are

$$c' - c = r - c = s - l \tag{3.81}$$

cycles of the open network that were not already present in the corresponding closed network. For chemostatted systems, such cycles have been called emergent cycles [21, 31], (See also Sec. 2.5.3).

Here, we choose to call these cycles external cycles, because they involve the flow rates $\tilde{\mathbf{w}}$ which are specific to the CSTR. The other cycles are called internal. A general cycle $\mathbf{c}'$ can be split into network components and flow components as $\mathbf{c}' = (\mathbf{c}, \tilde{\mathbf{c}})^{\mathrm{T}}$. This cycle obeys $\boldsymbol{\nu}' \cdot \mathbf{c}' = \boldsymbol{\nu} \cdot \mathbf{c} + \tilde{\mathbf{c}} = 0$. Here we can make the distinction between internal cycles $\mathbf{c}_\gamma$ previously defined for the network of the closed reactor which are such that $\boldsymbol{\nu} \cdot \mathbf{c}_\gamma = 0$ and $\tilde{\mathbf{c}}_\gamma = 0$; and external cycles $\mathbf{c}_\alpha$ which are such that

$$\tilde{\mathbf{c}}_\alpha = -\boldsymbol{\nu} \cdot \mathbf{c}_\alpha \neq 0. \tag{3.82}$$

As far as the thermodynamic description of the system is concerned, Eq. (3.65) becomes

$$\frac{dg_V}{dt} = \sum_{ki} \mu_k \nu'_{ki} w'_i, \tag{3.83}$$

within the framework of the extended network. In a stationary state, the entropy production rate of Eq. (3.64) may be rewritten as:

$$\begin{aligned} \sigma &= -\frac{1}{T}\boldsymbol{\mu} \cdot \boldsymbol{\nu} \cdot \mathbf{w} = -\frac{1}{T}\sum_\lambda w_\lambda \boldsymbol{\mu} \cdot \boldsymbol{\nu} \cdot \mathbf{c}_\lambda, \\ &= -\frac{1}{T}\sum_\alpha w_\alpha \boldsymbol{\mu} \cdot \boldsymbol{\nu} \cdot \mathbf{c}_\alpha, \\ &= \frac{1}{T}\sum_\alpha w_\alpha \boldsymbol{\mu} \cdot \tilde{\mathbf{c}}_\alpha \geq 0. \end{aligned} \tag{3.84}$$

This shows that in this case the entropy production rate can be written as a sum of contributions from external cycles denoted with the index α only. A similar property was reported in the case of chemostatted systems. [21, 31]

Note that σ is the entropy production due to chemical reactions. In a CSTR, there can be other significant sources of entropy production, such as the dissipative stirring and mixing entropy between inlet and outlet.

Let us now move to two illustrative examples of the above framework. The first example is a network of small size taken from Ref.[21], and the second one is a larger network describing polymers with a mass-exchange process taken from Ref.[17].

Example with a finite network

The set of reactions in the first example (taken from [31]) are

$$\mathrm{A}+\mathrm{B} \overset{1}{\rightleftharpoons} \mathrm{C}, \qquad w_1 = k_{+1}[\mathrm{A}][\mathrm{B}] - k_{-1}[\mathrm{C}], \tag{3.85}$$

$$\mathrm{C} \overset{2}{\rightleftharpoons} \mathrm{B}+\mathrm{D}, \qquad w_2 = k_{+2}[\mathrm{C}] - k_{-2}[\mathrm{B}][\mathrm{D}], \tag{3.86}$$

$$\mathrm{B}+\mathrm{D} \overset{3}{\rightleftharpoons} \mathrm{E}, \qquad w_3 = k_{+3}[\mathrm{B}][\mathrm{D}] - k_{-3}[\mathrm{E}], \tag{3.87}$$

$$\mathrm{E} \overset{4}{\rightleftharpoons} \mathrm{A}+\mathrm{B}, \qquad w_4 = k_{+4}[\mathrm{E}] - k_{-4}[\mathrm{A}][\mathrm{B}]. \tag{3.88}$$

The stoichiometry matrix of this network is then

$$\boldsymbol{\nu} = \begin{pmatrix} -1 & 0 & 0 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \tag{3.89}$$

and the corresponding hypergraph is shown in Fig. 3.4.

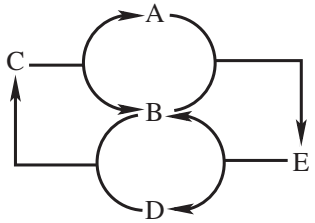


Figure 3.4: Hypergraph of the closed chemical network (3.85)-(3.88).

As shown in Ref. [21], this network has $l = 2$ conserved quantities $L_1 = [\mathrm{B}] + [\mathrm{C}] + [\mathrm{E}]$ and $L_2 = [\mathrm{A}] + [\mathrm{C}] + [\mathrm{D}] + [\mathrm{E}]$. There is only one cycle ($c = 1$), with a null right eigenvector $(1,1,1,1)^{\mathrm{T}}$.

For the open reactor network, the stoichiometric matrix $\boldsymbol{\nu}'$ is obtained from Eq. (3.78), to give

$$\boldsymbol{\nu}' = \begin{pmatrix} -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.90}$$

which has rank 5, $c' = 4$ cycles and $l' = 0$ conserved quantities. By opening the reactor, the dimension of the space of cycles has thus been increased by $s - l = 3$ consistent with Eq. (3.81).

A particular cycle decomposition we can now choose is the old cycle $\mathbf{c}_1 = (1,1,1,1,0,0,0,0,0)^{\mathrm{T}}$ and $\mathbf{c}_2 = (1,1,0,0,1,0,0,-1,0)^{\mathrm{T}}$, $\mathbf{c}_3 = (0,1,1,0,0,0,1,0,-1)^{\mathrm{T}}$, and $\mathbf{c}_4 = (0,0,1,1,-1,0,0,1,0)^{\mathrm{T}}$. The new cycles are represented in Fig. 3.5. This representation makes it clear that hypergraphs depicting the new cycles of the open network are built from the hypergraph of the closed network by removing some reactions and chemical species. Then the remaining pieces are connected together using a special symbol ϕ, which is introduced for this purpose and which describes new reaction pathways involving the exterior of the CSTR.

We note that the hypergraphs in Figs. 3.4 and 3.5 depend on the reaction network, but not on the concentration values of the involved species.

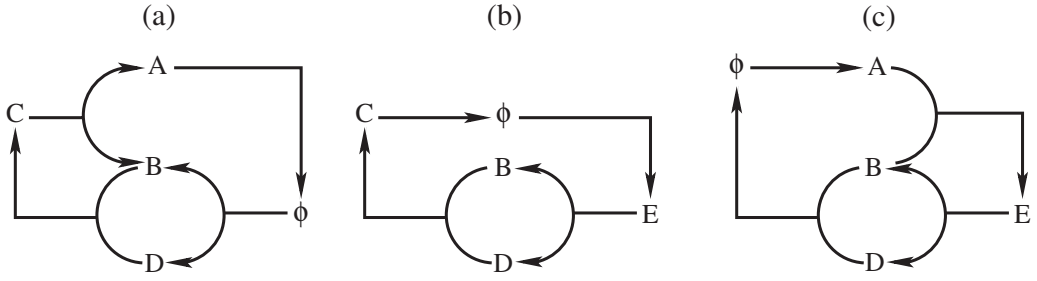


Figure 3.5: Hypergraphs of the three new cycles in the open version of the chemical network represented in Fig. 3.4. Here (a), (b) and (c) correspond to the cycles $\mathbf{c}_2$, $\mathbf{c}_3$ and $\mathbf{c}_4$ respectively. Note the appearance of the symbol ϕ which is a notation for new reactions involving the inflow and outflow of the CSTR.

Example with an infinite network

We now move to a more complex reaction network, namely the model of polymers undergoing a mass-exchange process taken from Ref. [17]. In this model, two polymers of with n and m repeating units exchange a single monomer unit, through the reaction

$$(n)+(m) \underset{\kappa}{\rightleftharpoons} (n+1)+(m-1), \qquad \text{for} \quad n \geq 1,\ m \geq 2. \tag{3.91}$$

In an open reactor, the kinetic equations can be written in the form:

$$\frac{dc_k}{dt} = \frac{1}{2} \sum_{n\geq 1, m\geq 2} \nu_{k,nm}\, w_{nm} + \frac{1}{\tau}(c_{k,0} - c_k) \qquad \text{for} \quad k \geq 1, \tag{3.92}$$

with the stoichiometric coefficients $\nu_{k,nm} = \delta_{k,n+1} + \delta_{k,m-1} - \delta_{k,n} - \delta_{k,m}$ and the rates $w_{nm} = \kappa c_n c_m - \kappa c_{n+1} c_{m-1}$ obeying the mass action law.

In the closed reactor ($\tau = \infty$), this network has two conserved quantities: the total concentration $C \equiv \sum_{k=1}^{\infty} c_k$ and the total number of monomeric units $M = \sum_{k=1}^{\infty} k\, c_k$. In the open reactor, these quantities are no longer conserved because they obey the equations

$$\frac{dC}{dt} = \frac{1}{\tau}(C_0 - C)\,, \tag{3.93}$$

$$\frac{dM}{dt} = \frac{1}{\tau}(M_0 - M)\,, \tag{3.94}$$

so that they converge asymptotically in time towards their value C_0 or M_0 fixed by the inlet concentrations.

Although the reaction network is infinite, it can be truncated by considering a finite number s of species. This is not a problem, since one can choose s to be arbitrarily large, whereas finite residence times and mass place clear constraints on the largest species that can be observed in the reaction mixture. The dynamics will then be fully captured by a finite-dimensional reaction network.

The reactions and the cycles can be enumerated using the list of all the reactions:

$$\begin{aligned} &1+2 \rightleftharpoons 2+1\,, \quad 2+3 \rightleftharpoons 3+2\,, \quad 3+4 \rightleftharpoons 4+3\,, \ldots \\ &1+3 \rightleftharpoons 2+2\,, \quad 2+4 \rightleftharpoons 3+3\,, \ldots \\ &1+4 \rightleftharpoons 2+3\,, \ldots \\ &\quad \vdots \end{aligned} \tag{3.95}$$

In the closed reactor, the number of reactions involving s species is thus equal to

$$r = \frac{1}{2}s(s-1)\,. \tag{3.96}$$

Since there are $l = 2$ conserved quantities in the closed reactor, Eq. (3.68) thus shows that the number of cycles is equal to

$$c = r - s + 2 = \frac{1}{2}(s-1)(s-2) + 1\,. \tag{3.97}$$

Accordingly, these numbers are increasing quadratically with the number s of species.

In the open reactor, the reactions include the rates $\tilde{w}_k = (c_{k0} - c_k)/\tau$ due to the flow so that the number of reactions involving s species is now given by

$$r' = r + s = \frac{1}{2}s(s+1)\,. \tag{3.98}$$

There are no conserved quantities $l' = 0$ and the number of cycles is here equal to

$$c' = c + s - 2 = r = \frac{1}{2}s(s-1)\,. \tag{3.99}$$

Therefore, opening the reactor only adds a number of new cycles $s-2$ that is increasing linearly with the number of species, while the total number of cycles of the open system is increasing quadratically with the number of species.

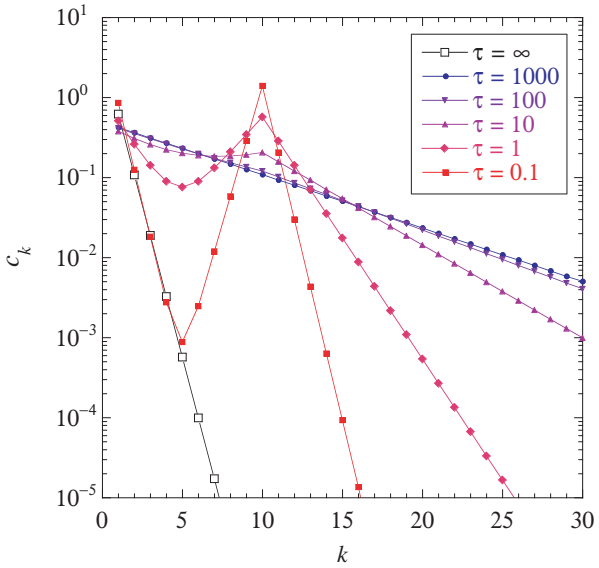


Figure 3.6: Stationary distributions of the oligomer concentrations $\{c_k\}$ for the mass-exchange process with the rate constant $\kappa = 1$ in a CSTR with the injection of monomers and 10-mers at the inlet concentrations $c_{1,0} = 1$ and $c_{10,0} = 2$ for different values of the residence time τ. If $\tau = \infty$, the reactor is closed and the stationary distribution is the equilibrium one (open squares). If τ is finite, the reactor is open and out of equilibrium (filled symbols).

In the CSTR, all the concentrations remain bounded in time. This rules out the possibility to observe an "unbalanced phase", such as the unbounded growth phase reported in Ref.[17] (see also Sec. 7.3.3) in a variant of this mass-exchange model, which was driven out-of-equilibrium by chemostats fixing the concentrations of polymers of certain lengths. In that model, the total concentration c increased linearly in time and the total number of monomers M increased quadratically. In

contrast, in a CSTR both quantities remain bounded in time, a property which follows generally from Eq. (3.73).

In Fig. 3.6, we show the stationary distribution of concentrations in a CSTR for different values of the residence time τ by injecting monomers at the concentration $c_{1,0}$ and oligomers of length $l = 10$ at the concentration $c_{10,0}$. The kinetic equations are integrated with a Runge-Kutta algorithm of orders 4 and 5 with variable steps from the initial distribution $c_k(0) = \exp(-k^2/2)$. The distribution is plotted after a time interval $t = 1000$ if $\tau = \infty, 0.1, 1, 10$, after $t = 10000$ if $\tau = 100$, and after $t = 50000$ if $\tau = 1000$, when stationarity is numerically reached. If $\tau = \infty$, the reactor is closed so that the concentrations reach their equilibrium exponential distribution

$$c_{k,\mathrm{eq}} = \frac{C(0)^2}{M(0)}\left[1 - \frac{C(0)}{M(0)}\right]^{k-1}, \tag{3.100}$$

determined by the initial values of the two invariant quantities $C(0) = 0.7533$ and $M(0) = 0.9119$, so that $c_{k,\mathrm{eq}} = 3.58 \times 0.174^k$. In contrast, under nonequilibrium conditions if τ is finite, the distribution deviates from being purely exponential and it even becomes bimodal with peaks at $k = 1$ and $k = 10$ if the open reactor is strongly out of equilibrium with a small enough residence time τ. Nevertheless, the distribution is always exponential beyond the largest injected concentration $c_{10,0}$ (see Appendix A in Ref. [20] for an extended discussion). In the open reactor, the distribution no longer depends on the initial conditions but on the values of the injected concentrations.

3.3 Serial Transfer

Now, we consider the dynamics of the reaction network in a typical serial transfer (Fig. 3.7.) experiment.[32] Serial transfer shows strong similarities with a CSTR, in that the reactor contents are diluted and refreshed with a new solution of well-defined content. Experimentally, it is easier to set up a serial transfer experiment then a CSTR since transferring a volume fraction can be done manually using only a pipette. Consequently, serial transfer is often chosen as a practical substitute for a CSTR. An important aim of this section is to see how far the analogy goes.

3.3.1 Time evolution of the concentrations

At every transfer, a fraction f of the solution volume V is transferred to another closed reactor already containing a fresh solution of volume $(1-f)V$ with reactants at the concentrations c_{k0} as illustrated in Fig. 3.7.

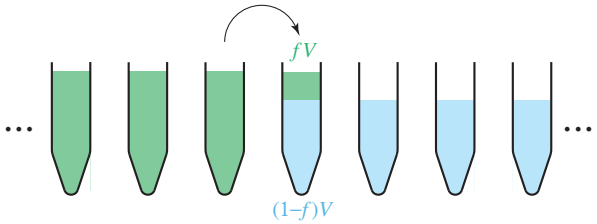


Figure 3.7: Schematic representation of a serial transfer experiment in which a volume fV of the solution of interest (green) is transferred repeatedly into a fresh solutions of volume $(1-f)V$ (blue).

Let $\mathscr{T}$ be the time interval between two transfers. During this time interval, the reactor is closed so that the concentrations evolve according to

$$\frac{d\mathbf{c}}{dt} = \boldsymbol{\nu} \cdot \mathbf{w}. \tag{3.101}$$

Let $\mathbf{c}(n\mathscr{T}-0)$ be the concentrations just before the previous transfer. The concentrations just after the transfer and stirring are thus given by

$$\mathbf{c}(n\mathscr{T}+0) = (1-f)\,\mathbf{c}_0 + f\,\mathbf{c}(n\mathscr{T}-0)\,. \tag{3.102}$$

Thereafter, the concentrations evolves according to

$$\mathbf{c}\,(t) = \mathbf{c}(n\mathscr{T}+0) + \int_{n\mathscr{T}}^{t} \boldsymbol{\nu}\cdot\mathbf{w}\left[\mathbf{c}(t')\right]\,dt' \tag{3.103}$$

with $n\mathscr{T}+0 < t < n\mathscr{T}+\mathscr{T}-0$. The concentrations just before the next transfer are thus given by

$$\mathbf{c}(n\mathscr{T}+\mathscr{T}-0) = (1-f)\,\mathbf{c}_0 + f\,\mathbf{c}(n\mathscr{T}-0) + \int_{n\mathscr{T}}^{(n+1)\mathscr{T}} \boldsymbol{\nu}\cdot\mathbf{w}\left[\mathbf{c}(t)\right]\,dt\,, \tag{3.104}$$

which defines a mapping $\mathbf{c}_{n+1} = \boldsymbol{\Phi}(\mathbf{c}_n)$ from $\mathbf{c}_n \equiv \mathbf{c}(n\mathscr{T}-0)$ to $\mathbf{c}_{n+1} \equiv \mathbf{c}(n\mathscr{T}+\mathscr{T}-0)$. A similar mapping can be obtained for the concentrations after the transfers.

Let us consider the case where transfers are repeated every time interval $\mathscr{T} = \Delta t$, which is taken to be short enough with respect to the timescale τ_r of the fastest reactions in $\boldsymbol{\nu}$. Under this condition, we can linearize the integral in Eq. (3.104) and we get the approximate ordinary differential equations:

$$\frac{\Delta\mathbf{c}}{\Delta t} \simeq \frac{1-f}{\Delta t}\,(\mathbf{c}_0 - \mathbf{c}_n) + \boldsymbol{\nu}\cdot\mathbf{w}\left[(1-f)\,\mathbf{c}_0 + f\,\mathbf{c}_n\right]\,, \tag{3.105}$$

where $\Delta\mathbf{c} = \mathbf{c}_{n+1} - \mathbf{c}_n$. Introducing the effective residence time

$$\tau \equiv \frac{\Delta t}{1-f}\,, \tag{3.106}$$

we recover in the limit $\Delta t \to 0$ the kinetic equations of the concentrations in a CSTR:

$$\frac{d\mathbf{c}}{dt} = \boldsymbol{\nu}\cdot\mathbf{w}(\mathbf{c}) + \frac{1}{\tau}\,(\mathbf{c}_0 - \mathbf{c})\,. \tag{3.107}$$

If $f = 1 - \mathscr{T}/\tau$ in the limit $\mathscr{T} \to 0$, an experiment of serial transfers between closed reactors is thus similar to an experiment in a CSTR. Therefore, similar nonequilibrium regimes are expected in both experiments under comparable conditions.

The analogy between CSTR and serial transfer works best when transfer occurs on a much faster timescale than chemistry. Since the object of interest is the reacting chemical system, experiments typically deviate from that limit, such that the system actually gets to consume the reactants that are refreshed. We will show however, that through Eq. (3.106) closely analogous behaviors are still found when reactions are not too slow with respect to transfer.

3.3.2 Thermodynamic Aspects of Serial Transfer

Let us follow Gibbs' free energy during the time evolution. Before the transfer at time $n\mathscr{T}$, the free energy density of the solution in the volume V is $g_V\left[\mathbf{c}\,(n\mathscr{T}-0)\right]$. After the transfer of the volume fV of solution into the volume $(1-f)V$ of fresh solution and the mixing of both, the free energy density becomes $g_V\left[f\mathbf{c}\,(n\mathscr{T}-0) + (1-f)\mathbf{c}_0\right]$. Thereafter, the free energy density changes in time since the concentrations evolve according to Eq. (3.101) in the closed reactor. At the end of the time interval $n\mathscr{T} < t < n\mathscr{T}+\mathscr{T}$, the free energy density has thus become

$$g_V\left[\mathbf{c}\,(n\mathscr{T}+\mathscr{T}-0)\right] = g_V\left[f\mathbf{c}\,(n\mathscr{T}-0) + (1-f)\mathbf{c}_0\right] + \int_{n\mathscr{T}}^{(n+1)\mathscr{T}} dt\,\dot{g}_V\left[\mathbf{c}(t)\right]\,, \tag{3.108}$$

where $\dot{g}_V = \boldsymbol{\mu}\cdot\boldsymbol{\nu}\cdot\mathbf{w}$ is the time derivative of the free energy in the closed reactor given by Eq. (3.61) with $\tau=\infty$. The process repeats itself at every time interval.

In the limit where $\mathscr{T}=\Delta t\to 0$ with $f=1-\Delta t/\tau$, using the same notation $\mathbf{c}_n\equiv\mathbf{c}(n\mathscr{T}-0)$ as above, Eq. (3.108) becomes

$$g_V(\mathbf{c}_{n+1}) = g_V\left[\mathbf{c}_n+\frac{\Delta t}{\tau}(\mathbf{c}_0-\mathbf{c}_n)\right]+\Delta t\,\boldsymbol{\mu}(\mathbf{c}_n)\cdot\boldsymbol{\nu}\cdot\mathbf{w}(\mathbf{c}_n)+O(\Delta t^2)\,. \tag{3.109}$$

Since $\boldsymbol{\mu}=\partial g_V/\partial\mathbf{c}$, the previous equation becomes

$$g_V(\mathbf{c}_{n+1}) = g_V(\mathbf{c}_n)+\frac{\Delta t}{\tau}\boldsymbol{\mu}(\mathbf{c}_n)\cdot(\mathbf{c}_0-\mathbf{c}_n)+\Delta t\,\boldsymbol{\mu}(\mathbf{c}_n)\cdot\boldsymbol{\nu}\cdot\mathbf{w}(\mathbf{c}_n)+O(\Delta t^2)\,. \tag{3.110}$$

In the limit $\Delta t\to 0$, we thus find the differential equation

$$\frac{dg_V}{dt} = \boldsymbol{\mu}(\mathbf{c})\cdot\boldsymbol{\nu}\cdot\mathbf{w}(\mathbf{c})+\frac{1}{\tau}\boldsymbol{\mu}(\mathbf{c})\cdot(\mathbf{c}_0-\mathbf{c})\,, \tag{3.111}$$

which is the same as Eq. (3.61) for the time evolution of the free energy in the CSTR.

In the limit $\Delta t\to 0$, there is thus equivalence between the dynamics in the CSTR and the time evolution in a serial transfer experiment.

3.3.3 General properties of the reaction network in serial transfers

The considerations of Sec. 3.2.3 extends to reaction networks in serial transfers between closed reactors. Here, a stationary state corresponds to a fixed point $\mathbf{c}_n=\mathbf{c}_*=\boldsymbol{\Phi}(\mathbf{c}_*)$ of the mapping defined by Eq. (3.104).

As in the case of the CSTR, conserved quantities of the closed network, namely quantities of the form $L=\boldsymbol{\ell}\cdot\mathbf{c}$ are no longer conserved in the open reactor. Instead, their dynamics follows a simple relaxation equation

$$L(n\mathscr{T}+\mathscr{T}-0) = (1-f)L_0+fL(n\mathscr{T}-0)\,, \tag{3.112}$$

which is the counterpart of Eq. (3.73). At the fixed point where the conserved quantity is such that $L(n\mathscr{T}+\mathscr{T}+0)=L(n\mathscr{T}+0)=L_*$, this quantity equals the quantity L_0, which is the conserved quantity of the closed network evaluated at the injected concentration and which was introduced in Eq. (3.72).

Furthermore, the fixed point $\mathbf{c}_*$ should satisfy the same condition

$$\boldsymbol{\nu}'\cdot\mathbf{w}'=0\,, \tag{3.113}$$

as in Subsec. 3.2.3 in terms of the same stoichiometric matrix (3.78), which was introduced to characterize the CSTR. Note however that now $\mathbf{w}'$ is replaced by $\mathbf{w}'=(\langle\mathbf{w}\rangle,\tilde{\mathbf{w}})^{\mathrm{T}}$ with the time average of the reaction rates over the time interval between the transfers

$$\langle\mathbf{w}\rangle = \frac{1}{\mathscr{T}}\int_{n\mathscr{T}}^{(n+1)\mathscr{T}}\mathbf{w}\left[\mathbf{c}(t)\right]dt\,, \tag{3.114}$$

which has the same value between every transfer because the process repeats itself from the point fixed $\mathbf{c}_n=\mathbf{c}_*$, and

$$\tilde{\mathbf{w}} = \frac{1-f}{\mathscr{T}}(\mathbf{c}_0-\mathbf{c}_*)\,. \tag{3.115}$$

Therefore, Eq. (3.80) applies here as well and the number of conserved quantities is equal to zero. The rates can be decomposed as $\mathbf{w}'=\sum_\lambda w'_\lambda\mathbf{e}'_\lambda$ onto the $c'=\dim\ker(\boldsymbol{\nu}')$ right null eigenvectors of the matrix $\boldsymbol{\nu}'$, which define the cycles, as in Subsec. 3.2.3.

Illustrative example

Here, we illustrate the correspondence between the serial transfers and CSTR dynamics using the mass-exchange model introduced above. The conditions of operation of the reactors are the same as in Fig. 3.6, namely monomers are injected at the concentration $c_{1,0} = 1$ and oligomers of length $l = 10$ at the concentration $c_{10,0} = 2$, and again a rate constant $\kappa = 1$ is chosen. The main difference is that now the reactor is evolving by serial transfers instead of the CSTR dynamics. The kinetic equations have been integrated using the integrator `odeint`, which is available in SciPython. The precision of this integrator is fixed to 10^{-5}, which is the same as that used in Fig. 3.6. The length distributions of the oligomers have been observed at the time $1000\mathscr{T} - 0$, at which we find that the distributions have reached stationarity. In Fig. 3.8, simulations of serial transfers have been carried out keeping the time $\mathscr{T}$ fixed while varying f. As expected in this case, the length distribution approaches the equilibrium exponential distribution in the limit $f \to 1$, since the residence time introduced in Eq. (3.106) becomes infinite.

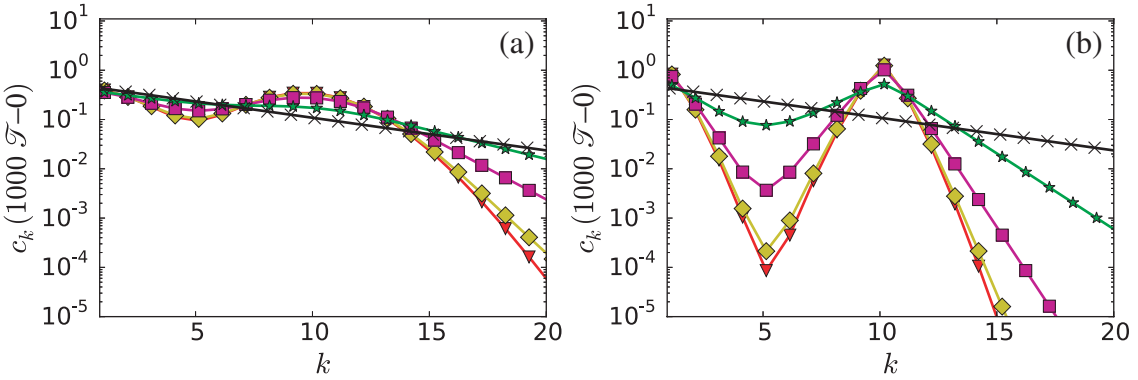


Figure 3.8: Concentrations c_k of oligomers versus their length k probed at the time $1000\mathscr{T} - 0$ after a thousand serial transfers with fixed parameters (a) $\mathscr{T} = 1$ and (b) $\mathscr{T} = 0.1$ and for various values of f. Symbols correspond to $f = 0.01$ (downward red triangles), $f = 0.1$ (yellow diamonds), $f = 0.5$ (magenta squares), $f = 0.9$ (green stars), and black crosses represent the equilibrium distribution.

In order to test more precisely the convergence towards the CSTR dynamics, we have varied in Fig. 3.9 the parameters $(f, \mathscr{T})$ while keeping the residence time $\tau_{\text{eff}} = \tau$ constant either at the value 1 or 0.1. Polymer molecules are equally reactive in this model, for which we can write a reaction timescale τ_r, corresponding to the typical waiting time to encounter any other polymer. If we consider donation of mass and reception of mass as different reactions (due to the abundance of monomers, abstraction of mass is initially slow), we can then write for the former

$$\tau_r = \frac{1}{\kappa C}. \tag{3.116}$$

Since $C = c_{1,0} + c_{10,0} = 3$, we find $\tau_r = 1/3$, which is situated between the two residence times being assessed, and explains the clear differences in overlap between Fig.3.9a and Fig.b.

The length distributions of the oligomers have been observed at the time $1000\mathscr{T} - 0$. These plots indeed confirm that, in this system, a convergence towards the CSTR is obtained when $f \to 1$, which is equivalent to $\mathscr{T} \to 0$ since the residence time τ is kept constant.

In general, the state of the reactor following serial transfers with arbitrary parameters $(f, \mathscr{T})$ can differ substantially from the predictions of the CSTR. However, if the parameters $(f, \mathscr{T})$ are chosen according to Eq. (3.106) and the time of observation is not too long, as shown in Fig. 3.9, the behavior resulting from serial transfers can be quite close to that observed by the CSTR dynamics even when the parameter f is varied in a large range from 0.001 to 0.99.

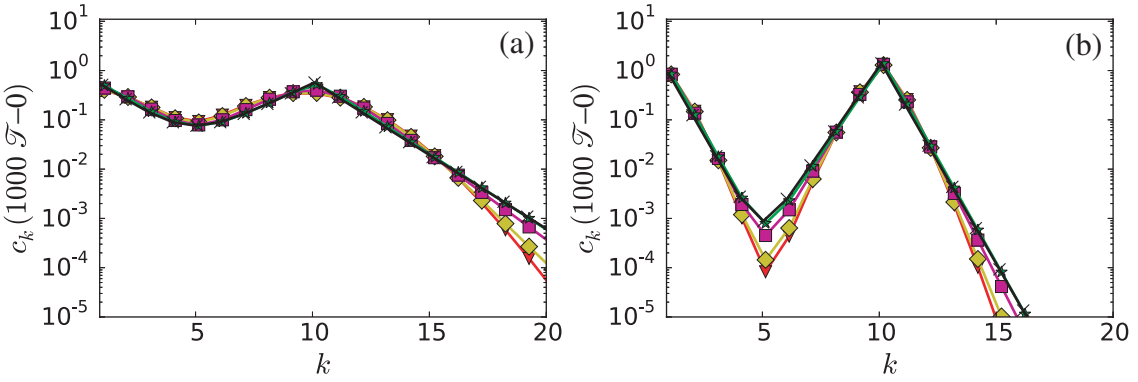


Figure 3.9: Concentrations c_k of oligomers versus their length k probed at the time $1000\mathscr{T}-0$ after a thousand serial transfers corresponding to varying $(\mathscr{T}, f)$ parameters at fixed residence time (a) $\tau_{\text{eff}} = 1$ or (b) $\tau_{\text{eff}} = 0.1$. Symbols correspond to $f = 0.01$ (downward red triangles), $f = 0.1$ (yellow diamonds), $f = 0.5$ (magenta squares), $f = 0.9$ (green stars), but now black crosses represent the length distribution predicted by the CSTR dynamics.

3.4 Osmotically coupled growing compartments

Of particular interest in origins of life are replicating compartments such as protocells. Compartments offer many advantages, such as chemically distinct environments (Ch. 5,6) and multilevel selection (Ch. 8) which make them desirable entities. However, making a copy of compartment and its contents is not trivial, especially if the ingredients one has are simple reactions and basic physical chemistry.

As part of an ongoing experimental work "Natural selection of compartmentalized autocatalytic chemical reactions"[‡] (in collaboration with Heng Lu, Cyrille Jeancolas, Philippe Pelupessy, Fabien Ferrage, Éstanislau Guilherme, Gabrielle Woronoff, Rebecca Turk Macleod, Ludovic Jullien, Eörs Szathmary, Andrew Griffths and Philippe Nghe), this section considers a replication process, where a compartment grows through osmosis and diffusive exchange with an environment, after which it is split in two (see Fig. 3.10). Experimentally, the splitting occurs through shear forces. As we will show, repeating such a process indefinitely requires autocatalysis, which is experimentally achieved through the formose reaction, which is detailed in Appendix 10.2.

3.4.1 Exchange process between compartments

Suppose we have a compartment I of volume V^{I}, and molecule numbers given by $\{n_1^{\text{I}}, n_2^{\text{I}}, ..., n_s^{\text{I}}\}$, leading to molecular fractions

$$x_i^{\text{I}} = \frac{n_i^{\text{I}}}{\sum_{j=1}^{s} n_j^{\text{I}}}. \tag{3.117}$$

During a time τ, compartment I is placed in contact with an environment, which is either i) a compartment II with a volume V^{II}, molecule numbers given by n_i^{II} and molecular fractions x_i^{II} or ii) a large reservoir acting as a chemostat, with fixed molecular fractions $\bar{x}_i$. The contact happens only through diffusion, across an oil phase or membrane separating the two compartments. On the timescale τ, this diffusion process is selective: one only observes the exchange of sufficiently mobile species (e.g. metal ions do not exchange, whereas small uncharged molecules do). At the end of the time τ, compartment I is split in two equal volumes with the same composition

$$V^{\text{I}} \to \frac{V^{\text{I}}}{2}, \quad \boldsymbol{x}^{\text{I}} \to \boldsymbol{x}^{\text{I}} \tag{3.118}$$

[‡]manuscript in preparation

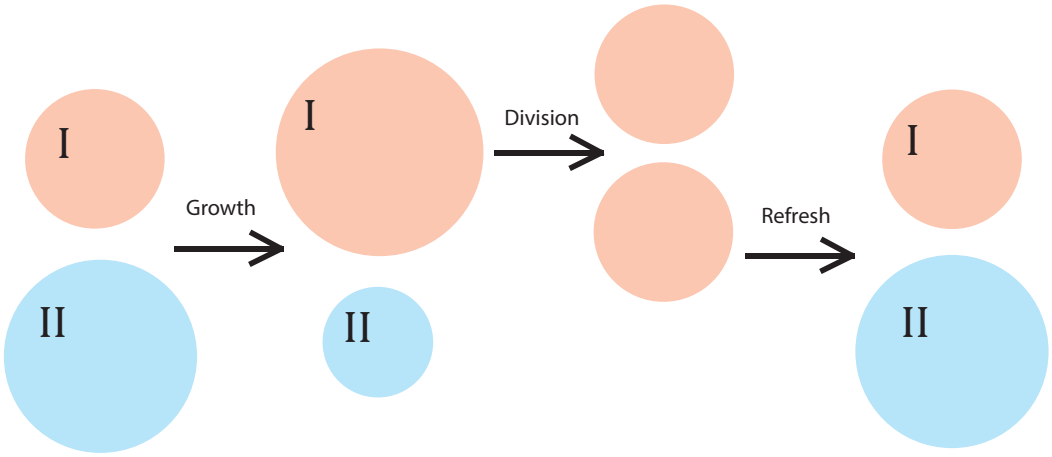


Figure 3.10: Illustration of the droplet growth-division cycle.

One of these split compartments is then taken, and assigned to be the new compartment I. This compartment is then again placed in contact with the environment. If the environment was compartment II, this compartment is replaced by a new one with the exact same initial conditions as occurred in the previous round.

The environments have the same solvent as compartment I, which we assume is always exchanged. Compartments with the same molecules have the same local chemical reaction network $\boldsymbol{\nu}_r$ and thus exhibit the same chemistry. Distinct chemistry directly implies the presence of distinct species, as elaborated upon in Sec.2.2.2 on representations in chemical networks.

By this, it is meant that pathways and catalysts are not hidden and subject to the nonambiguity condition ($\forall i,j \;\; \nu_{ij}^{+}\nu_{ij}^{-} = 0$, Eq. (2.15) in Sec. 2.2.2). For example, if the following reaction can take place in compartment I

$$\mathrm{S}^{\mathrm{I}} \rightleftharpoons \mathrm{P}^{\mathrm{I}}, \tag{3.119}$$

than this (within our framework) points to an inherent capacity of S to be converted to P, and the equivalent reaction must exist in compartment II

$$\mathrm{S}^{\mathrm{II}} \rightleftharpoons \mathrm{P}^{\mathrm{II}}. \tag{3.120}$$

However, if compartment I contains a catalyst E, a transformation

$$\mathrm{S}^{\mathrm{I}} + \mathrm{E}^{\mathrm{I}} \overset{1}{\rightleftharpoons} \mathrm{ES}^{\mathrm{I}}, \tag{3.121}$$

$$\mathrm{ES}^{\mathrm{I}} \overset{2}{\rightleftharpoons} \mathrm{EP}^{\mathrm{I}}, \tag{3.122}$$

$$\mathrm{EP}^{\mathrm{I}} \overset{3}{\rightleftharpoons} \mathrm{E}^{\mathrm{I}} + \mathrm{P}^{\mathrm{I}}, \tag{3.123}$$

can take place. If compartment II contains S but not E, this reaction pathway is unique to I. Although I and II may be highly similar (e.g. same solvent), their chemistry may be described by different chemical networks. In our treatment, we impose that any distinction between these networks must be justified through the chemical composition. In doing so, we can address the question whether these differences in local chemistry can be maintained.

3.4.2 Periodic solutions

During the time that the two compartments are in contact, they are described by reaction-diffusion equations. During this period, the combination of compartments forms a closed system, for which

we can write a Lyapunov function:

$$S = k_b \sum_i n_i^{\mathrm{I}} \ln \frac{n_i^{\mathrm{I}}}{\sum_j n_j^{\mathrm{I}}} + k_b \sum_i n_i^{\mathrm{II}} \ln \frac{n_i^{\mathrm{II}}}{\sum_j n_j^{\mathrm{II}}} - \sum_i n_i^{\mathrm{I}} \frac{\mu_i^\circ}{T} - \sum_i n_i^{\mathrm{II}} \frac{\mu_i^\circ}{T}, \tag{3.124}$$

where we have neglected the contribution of the oil phase, for which the contribution of dissolved substances at any given time is negligible. The Lyapunov function we have chosen is the entropy, for which $\dot{S} \geq 0$ and for which $\dot{S} = 0$ if and only if the total system has reached its unique fixed point. In general, systems with a Lyapunov function do not have time-periodic solutions[33].

However, at the end of time τ, we split the droplet and place it in contact with a new compartment. It is this driving that makes it possible to have periodic solutions. Let us write the vector $\boldsymbol{\zeta}$, which contains the molar fractions and volume corresponding to droplet I. During the first stage of the protocol, reactions and exchange occur during a time $\tau^- = \tau - \varepsilon$, where ε is arbitrarily small, such that

$$\boldsymbol{\zeta}(t) = \begin{bmatrix} x_1^{\mathrm{I}}(t) \\ \vdots \\ x_s^{\mathrm{I}}(t) \\ V(t) \end{bmatrix} \rightarrow \begin{bmatrix} x_1^{\mathrm{I}}(t+\tau^-) \\ \vdots \\ x_s^{\mathrm{I}}(t+\tau^-) \\ V(t+\tau^-) \end{bmatrix} \tag{3.125}$$

For the subsequent division, we then have

$$\boldsymbol{\zeta}(t+\tau) = \begin{bmatrix} x_1^{\mathrm{I}}(t+\tau) \\ \vdots \\ x_s^{\mathrm{I}}(t+\tau) \\ V(t+\tau) \end{bmatrix} = \begin{bmatrix} x_1^{\mathrm{I}}(t+\tau^-) \\ \vdots \\ x_s^{\mathrm{I}}(t+\tau^-) \\ V(t+\tau^-)/2 \end{bmatrix} = P(\boldsymbol{\zeta}(t)). \tag{3.126}$$

The function $P(\boldsymbol{\zeta}(t))$ is a Poincaré map. Here, it takes $\boldsymbol{\zeta}(t)$ as an argument and returns the vector $\boldsymbol{\zeta}(t+\tau)$ that would be obtained upon integrating the corresponding dynamical system for a time τ^-, followed by splitting. The simplest periodic orbit that can be obtained in this way, follows the map

$$\boldsymbol{\zeta}(t+\tau) = P(\boldsymbol{\zeta}(t)), \tag{3.127}$$

which implies

$$\begin{bmatrix} x_1^{\mathrm{I}}(t) \\ \vdots \\ x_s^{\mathrm{I}}(t) \\ V(t) \end{bmatrix} = \begin{bmatrix} x_1^{\mathrm{I}}(t+\tau) \\ \vdots \\ x_s^{\mathrm{I}}(t+\tau) \\ V(t+\tau^-)/2 \end{bmatrix}. \tag{3.128}$$

While the volume doubles, the chemical composition at the end is exactly that of the start (provided the division is exactly symmetric). This implies that the quantity of every species has doubled after τ^- has elapsed

$$\forall i, \quad n_i^{\mathrm{I}}(t+\tau^-) = 2n_i^{\mathrm{I}}(t). \tag{3.129}$$

Let us now consider what this means for conservation laws and autocatalysis.

3.4.3 Conservation laws under recurrence

Let us consider the stoichiometric submatrix $\bar{\boldsymbol{\nu}}$, containing the chemical species in compartment I that do not exchange. In particular, we will consider the case where left nullspace of $\bar{\boldsymbol{\nu}}$ is nonempty, such that there is at least one conservation law L. When at the end of the round compartment I splits in two, so do the abundances of conserved species: $L \to L/2$. Upon the nth repetition, $L \to L/2^n$, and soon after $L \to 0$. Since then

$$L = \sum_i a_i \, n_i^{\mathrm{I}} = 0, \tag{3.130}$$

it follows that all species obeying a mass-like conservation law must become absent from the system

$$\forall i, a_i > 0, \quad n_i^{\mathrm{I}} = 0. \tag{3.131}$$

for a conservation law that is not mass-like, e.g. of the form $L = n_A^{\mathrm{I}} - n_B^{\mathrm{I}}$, species do not need to become absent to respect Eq. (3.72). These species are still subject to dilution, however, so a persistent presence of these species will require a process that resupplies them while respecting the constraints imposed by L.

3.4.4 Growth due to one source with identical chemistry

To assess different growth regimes, it will be instructive to study our system in a situation where τ is long enough to reach equilibrium for the processes deemed fast enough to be described by $\boldsymbol{\nu}$. A compartment grows because it receives an influx of solvent and possibly building blocks:

$$\mathrm{A}^{\mathrm{I}} \rightleftharpoons \mathrm{A}^{\mathrm{II}}. \tag{3.132}$$

This exchange will equilibrate molar fractions between compartments, for those species that are

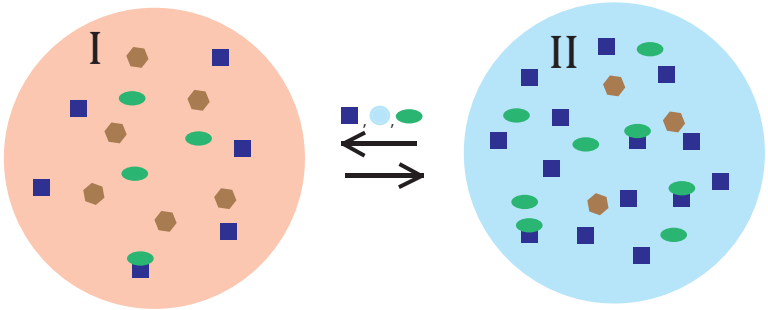


Figure 3.11: Illustration of two droplets with the same chemistry.

i) exchanged, or ii) formed only from exchanged species. Here, we explicitly suppose that both compartments have the same chemical networks for this subset (for now, we exclude the case that e.g. I has a catalyst that II does not). Denoting Ω the set of such species, the system will ultimately tend to the equilibria

$$\forall i \in \Omega, \quad \mu_i^{\mathrm{I}} = \mu_i^{\mathrm{II}}, \tag{3.133}$$

which for ideal solutions (which we will assume henceforth) corresponds to

$$\forall i \in \Omega, \quad x_i^{\mathrm{I}} = x_i^{\mathrm{II}}. \tag{3.134}$$

This follows directly from the entropy S under the constraints of the two-compartment network [34]. Species that are not in Ω still contribute to the osmotic pressure. We denote their collective fractions as $\bar{x}^{\mathrm{I}}$, $\bar{x}^{\mathrm{II}}$, and have the balance equations

$$\sum_{i=1} x_i^{\mathrm{I}} = 1 - \bar{x}^{\mathrm{I}}, \qquad \sum_{i=1} x_i^{\mathrm{II}} = 1 - \bar{x}^{\mathrm{II}}. \tag{3.135}$$

If the Eqs. (3.134) is satisfied, this balance is automatically satisfied as well: $\bar{x}^{\mathrm{I}} = \bar{x}^{\mathrm{II}}$. This is achieved by changing the compartment sizes (the number of molecules in a compartment). Since $\bar{x}^{\mathrm{I}}$ cannot be formed from exchanged species alone, it is not resupplied, and over the course of several divisions $\bar{x}^{\mathrm{I}} \to 0$. Knowing that I loses these compounds, we can now consider the following scenarios:

i) Droplet II retains a permanent fraction of nonexchanging compounds: $\bar{x}^{\mathrm{II}} \neq 0$. This may happen because II is a large reservoir, or because we constantly renew the neighbor droplet II. Since, $\bar{x}^{\mathrm{I}} \to 0$, we have $\bar{x}^{\mathrm{II}} \neq \bar{x}^{\mathrm{I}}$ and Eq. (3.134) can no longer be satisfied. The equilibrium solution is to absorb all matter in compartment II.

ii) Droplet II has no nonexchanging compounds $\bar{x}^{\mathrm{II}} = 0$, and is a large reservoir. If this is true, then Eq. (3.134) will be satisfied at all times once $\bar{x}^{\mathrm{I}} = 0$. Since no further gradients will then exist, growth is arrested, and compartment I vanishes upon subsequent divisions $V \to \ldots \to V/2^n \to \ldots \to 0$.

iii) Droplet II has no nonexchanging compounds $\bar{x}^{\mathrm{II}} = 0$ and is a neighbor droplet. In this case having two droplets or a single one is equally valid for the Lyapunov function provided.

Clearly, the final situation is counterintuitive. This is because we have so far allowed ourselves to neglect the effect of surface tension. At the scale of these droplets (100 pL), the contribution of surface tension is small with respect to other entropic contributions (such as the $\sum x_i^I \ln x_i^I$). However, when all other contributions cancel out, the growth mechanism becomes Ostwald ripening.

We conclude that two compartments undergoing the same chemistry and allowed to relax to equilibrium before compartment I is divided, will not yield persistent growth-division cycles for I. This may not come as a surprise, but making it explicit provides two directions to look further: i) choosing τ such that division happens before equilibrium and ii) having chemically distinct compartments. The interest of the former strategy is illustrated in Chapter 8 on transient compartmentalization.

3.4.5 Growth due to one source with different chemistry

Provided we have the same solvent and a set of common compounds in I and II, a distinct chemistry in compartment I acting on the common compounds implies that there are compounds in I that afford this pathway. These compounds are not present in II.

We can strictly consider three ways in which these distinct compounds afford a distinct chemistry when acting on an exchanged compound S:

i) they are solely consumed as reactants. Supposing an exchanged compound S and a unique compound X, such a reaction would e.g. be

$$\mathrm{S} + \mathrm{X} \rightleftharpoons \mathrm{Y} + \mathrm{Z}. \tag{3.136}$$

ii) they act as allocatalysts. converting species in Ω to new species. Supposing an exchanged compound S and a unique compound E, this would e.g. afford the reactions forming a unique nonexchanging compound P

$$\mathrm{S} + \mathrm{E} \rightleftharpoons \mathrm{ES} \rightleftharpoons \mathrm{EP} \rightleftharpoons \mathrm{P} + \mathrm{E}. \tag{3.137}$$

iii) they are autocatalyst, species that facilitate the formation of new species. See Ch. 5 for an elaborate discussion on autocatalysis. Let us here just say that by some combination of steps, an

autocatalytic species consuming the exchanged product is regenerated. Let us e.g. consider the Toy Formose reaction, with C_1 the exchanging species

$$C_1 + C_2 \rightleftharpoons C_3, \tag{3.138}$$
$$C_1 + C_3 \rightleftharpoons C_4, \tag{3.139}$$
$$C_4 \rightleftharpoons 2C_2. \tag{3.140}$$

If the distinct chemistry arises from i) and ii), but not iii), then the chemistry can be considered transient: the species that afford this pathway are not synthesized (if they were, we would have autocatalysis, see Ch. 5.). The reactants that afford i) and the catalyst that afford ii) do not exchange (which is the feature that makes the chemistry in I distinct). As a set of distinct species, they are neither formed (which would require autocatalysis) nor degraded (which would be reverse autocatalysis or exchange, both ruled out). Consequently, they are subject to conservation laws L (for catalysts, this is shown explicitly in Ch. 5).

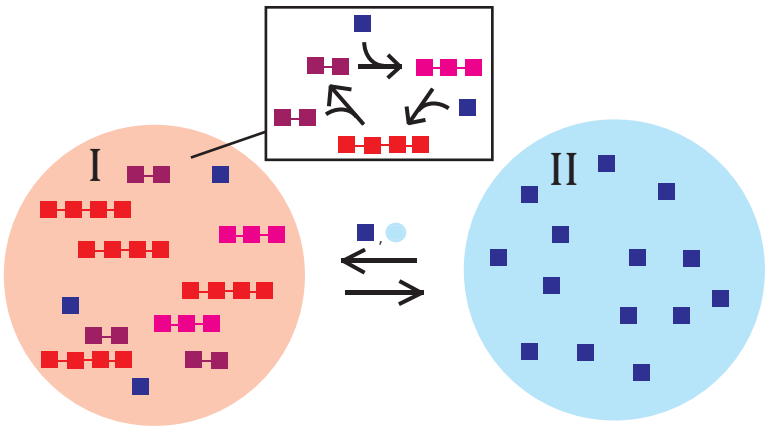


Figure 3.12: Illustration of two droplets exchanging C_1. In compartment I, the Toy Formose reaction is enabled. Dark blue squares: C_1, double purple squares: C_2, triple pink squares: C_3, quadruple red squares C_4.

Consequently, upon repeated division these species vanish, along with their associated conservation laws, after which the pathway is removed. It follows that persistent chemical distinctness in the presence of one similar neighbor compartment or reservoir is ruled out, except for autocatalysis.

Let us now consider the equilibrium question for autocatalysis. As long as I is chemically distinct, it will utilize a certain portion of exchanged species to form nonexchanging autocatalysts and their derivatives. Upon repeated division, the molar fraction $\bar{x}^{\mathrm{I}}$ will be entirely dominated by these species. Equilibrium will occur when the autocatalytic network fraction compensates the nonexchanging fraction in II: $\bar{x}^{\mathrm{I}} = \bar{x}^{\mathrm{II}}$. For an infinite reservoir, this point may never come, since $\bar{x}^{\mathrm{I}}$ can be resupplied indefinitely. For a neighbor compartment, such a point must exist due to the finite mass.

In general, we require compartment I to have a different composition, to maintain the gradients to grow it. We therefore require compartment I to be sufficiently chemically distinct from II.

3.4.6 Some illustrative examples

The following situations yield compartments capable of growth, but not persistence.

A salty compartment

A compartment containing nonexchanging compounds (e.g. a salt) and solvent is placed in contact with a reservoir or droplet containing (potentially different) nonexchanging salts and solvent. If

initially, $\bar{x}^{\mathrm{I}}(0) > \bar{x}^{\mathrm{II}}(0)$ solvent will flow to compartment I, lowering $\bar{x}^{\mathrm{I}}$. After a time τ^-, we have $\bar{x}^{\mathrm{I}}(\tau^-) \geq \bar{x}^{\mathrm{II}}(\tau^-)$, where equality corresponds to equilibrium. After division, a new compartment I,

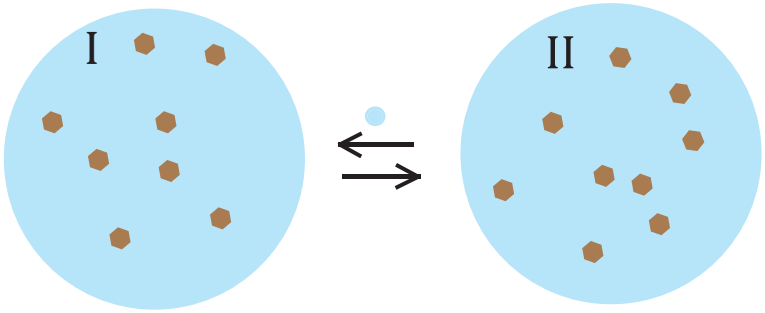


Figure 3.13: Illustration of two 'salty' droplets exchanging solvent.

with $\bar{x}^{\mathrm{I}}(\tau) < \bar{x}^{\mathrm{I}}(0)$ is generated and placed in contact with a new droplet II or the same reservoir, such that $\bar{x}^{\mathrm{II}}(\tau) = \bar{x}^{\mathrm{II}}(0)$. Upon repeated iteration, the osmotic pressure difference decreases

$$\bar{x}^{\mathrm{I}}(0) - \bar{x}^{\mathrm{II}}(0) > \bar{x}^{\mathrm{I}}(\tau) - \bar{x}^{\mathrm{II}}(\tau) > \ldots \tag{3.141}$$

Until $\bar{x}^{\mathrm{I}}(0) = \bar{x}^{\mathrm{II}}(0)$. At this point, there is no gradient and thus solvent flow. This process fails to respect the recurrence (3.128) for the volume variable V, and the compartment I will vanish upon repeated division.

A catalytically active compartment

Suppose a compartment containing a catalyst E that converts exchanging species S to immobile products P. The chemistry is given by Eqs.(3.121), (3.122) and (3.123), and has a mass-like conservation law L_1 governing catalyst abundance, and a mass-like conservation law L_2 following substrate to product

$$L_1 = n_E^{\mathrm{I}} + n_{ES}^{\mathrm{I}} + n_{EP}^{\mathrm{I}}, \tag{3.142}$$

$$L_2 = n_S^{\mathrm{I}} + n_{ES}^{\mathrm{I}} + n_{EP}^{\mathrm{I}} + n_P^{\mathrm{I}}. \tag{3.143}$$

Because we freely exchange S with the environment, our system only retains the conservation law

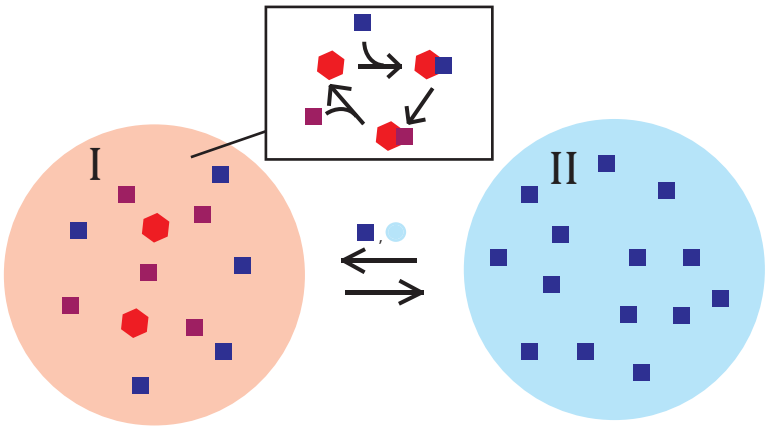


Figure 3.14: Illustration of two droplets exchanging reactant S (blue square), with compartment I having a catalyst E (red hexagon) to convert S to product P (purple squares).

L_1. Upon repeatedly dividing in two, $L_1 = 0$, and from Eq. (3.72) it follows that at that point, no catalyst will remain. The growth due to the osmotic pressure exerted by E, ES, EP and P (which exists due to the catalysis afforded by E) is therefore not persistent.

Polymer growth in a compartment

Let us now consider compartment I with N polymer strands of various lengths l, but at least $l \geq 2$. Monomers M are exchanged with the environment, and can be added to the polymer

$$[l] + \mathrm{M} \rightleftharpoons [l+1], \tag{3.144}$$

Since we have furthermore assumed a fixed number of growing polymer strands, and no way to

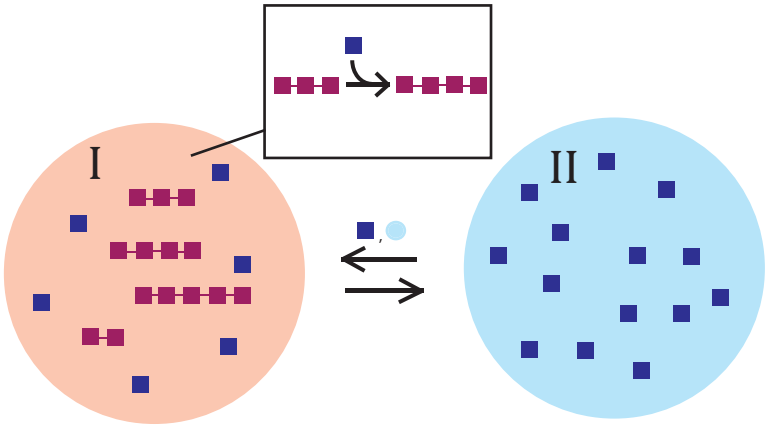


Figure 3.15: Illustration of two droplets exchanging monomer M, with compartment I having growing polymers that can incorporate M.

generate new ones, we have a conservation law

$$L = \sum_{l=2}^{\infty} n^{\mathrm{I}}_{[l]} = N. \tag{3.145}$$

The number of growing strands N will decrease upon successive divisions $N \to \ldots \to N/2^n \to \ldots \to 0$. Afterwards, there is no chemistry to make I distinct, gradients will vanish, and compartment I will disappear by further division.

3.4.7 Limit cycles

The linear stability of an orbit $\boldsymbol{\zeta}^*$ can be tested, by introducing a perturbation $\boldsymbol{\varepsilon}$ around a periodic orbit $\boldsymbol{\zeta}^*$ such that

$$\boldsymbol{\zeta}^* + \boldsymbol{\varepsilon}(t+\tau) = P(\boldsymbol{\zeta}^* + \boldsymbol{\varepsilon}(t)) \approx P(\boldsymbol{\zeta}) + [DP(\boldsymbol{\zeta})]\boldsymbol{\varepsilon}(t) \tag{3.146}$$

Here, $DP(\boldsymbol{\zeta}^*)$ is the linearized poincaré map around $\boldsymbol{\zeta}^*$. Stability of the orbit $\boldsymbol{\zeta}^*$ then requires the perturbation to dampen out

$$|DP(\boldsymbol{\zeta})| < 1. \tag{3.147}$$

Typically, the function $P(\boldsymbol{\zeta})$ can not be solved analytically and must be evaluated numerically, and similarly for the linearized Poincaré map[33].

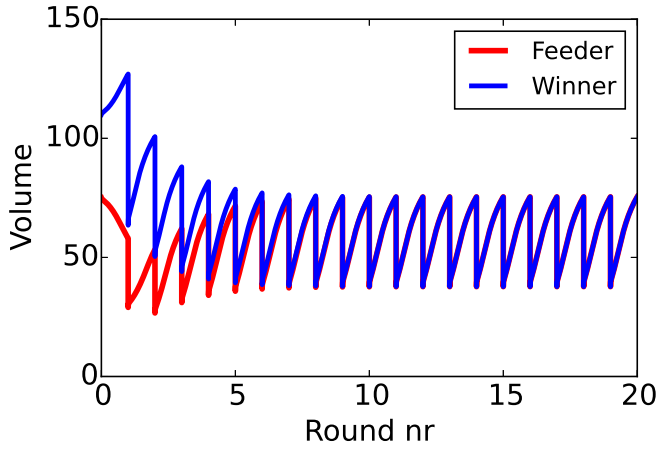


Figure 3.16: Plots for two growth scenarios. Blue: Periodic volume growth for an autocatalytic reaction droplet, which splits at the end of a round, after being fed by a 'feeder' droplet. Red: instead of propagating the 'Winner' droplet after round 1, we propagate the feeder, which acquires some of the autocatalytic compounds of the winner. Afterwards it is placed in contact with a new feeder droplet every round and it starts to converge to a winner state.

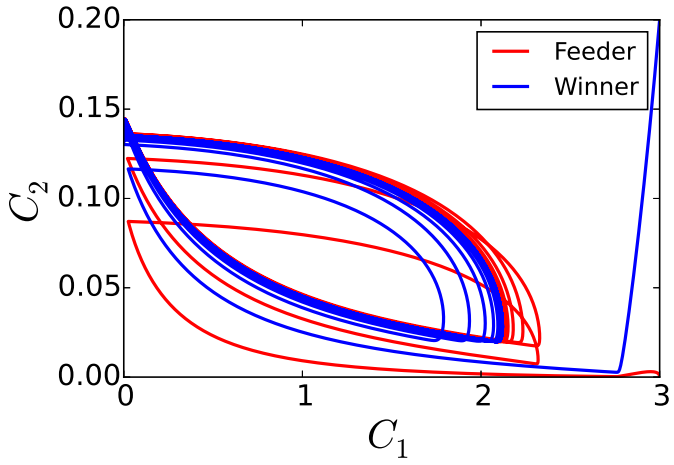


Figure 3.17: C_1 concentration as a function of C_2 concentration from $t = 0$ to 20τ, for the 'winner' compartment (blue) and a 'feeder' compartment (red). Both converge to the same limit cycle.

For the toy formose model, a stable limit cycle was obtained by simulating the protocol outlined in Appendix 10.2.

Bibliography

Articles

[1] F Schlögl. "Chemical reaction models for non-equilibrium phase transitions". In: *Z. Phys.* 253.2 (1972), pp. 147–161.

[2] P Gaspard. "Fluctuation theorem for nonequilibrium reactions". In: *J. Chem. Phys.* 120.19 (2004), pp. 8898–8905.

[3] Melissa Vellela and Hong Qian. “Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system : the Schlögl model revisited”. In: *J. R. Soc. Interface* 6 (2009), pp. 925–940.

[5] O.E. Rössler. “An equation for continuous chaos”. In: *Phys. Lett. A* 57 (1976), pp. 397–398.

[6] O.E. Rössler and K. Wegmann. “Chaos in the Zhabotinskii reaction”. In: *Nature* 271 (1978), pp. 89–90.

[8] P Dziekan et al. “Effect of a Local Source or Sink of Inhibitor on Turing Patterns Effect of a Local Source or Sink of Inhibitor on Turing Patterns”. In: *Commun. Theor. Phys.* 62 (2014), pp. 622–630.

[9] Matthieu Emond, Le Saux, and Jean-francois Allemand. “Energy Propagation Through a Protometabolism Leading to the Local Emergence of Singular Stationary Concentration Profiles”. In: *Chem. Eur. J.* 18 (2012), pp. 14375–14383.

[10] A. Lemarchand and B. Nowakowski. “Do the internal fluctuations blur or enhance axial segmentation ?” In: *EPL* 94 (2011), p. 48004.

[11] J. J. Hopfield. “Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes requiring high Specificity”. In: *Proc. Natl. Acad. Sci. U.S.A.* 71.10 (1974), pp. 4135–4139.

[12] J Ninio. “Kinetic amplification of enzyme discrimination”. In: *Biochimie* 57.5 (1975), pp. 587–595.

[13] R. Rao and L. Peliti. “Thermodynamics of accuracy in kinetic proofreading: dissipation and efficiency trade-offs”. In: *J. Stat. Mech.* 2015.6 (2015), p. 06001.

[14] Pierre Gaspard. “Kinetics and thermodynamics of DNA polymerases with exonuclease proofreading”. In: *Phys. Rev. E* 93.4 (Apr. 2016), p. 42420.

[15] D Andrieux and P Gaspard. “Nonequilibrium generation of information in copolymerization processes”. In: *Proc. Natl. Acad. Sci. U.S.A.* 105.28 (2008), pp. 9516–9521.

[16] Thomas E Ouldridge and Pieter ten Wolde. “Fundamental Costs in the Production and Destruction of Persistent Polymer Copies”. In: *Phys. Rev. Lett.* 118.15 (Apr. 2017), p. 158103.

[17] R. Rao, D. Lacoste, and M. Esposito. “Glucans monomer-exchange dynamics as an open chemical network”. In: *J. Chem. Phys.* 143 (2015), p. 244903.

[18] Christof B Mast et al. “Escalation of polymerization in a thermal gradient”. In: *Proc. Natl. Acad. Sci. U.S.A.* 110.20 (2013), pp. 8030–5.

[19] Varun Giri and Sanjay Jain. “The Origin of Large Molecules in Primordial Autocatalytic Reaction Networks”. In: *PLoS One* 7.1 (2012), e29546.

[20] Alex Blokhuis, David Lacoste, and Pierre Gaspard. “Reaction kinetics in open reactors and serial transfers between closed reactors”. In: *J. Chem. Phys.* 148 (2018), p. 144902.

[21] Matteo Polettini and Massimiliano Esposito. “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”. In: *J. Chem. Phys.* 141.2 (2014), pp. –.

[23] R Sander. “Compilation of Henry’s law constants (version 4.0) for water as solvent”. In: *Atmos. Chem. Phys.* 15 (2015), pp. 4399–4981.

[31] Riccardo Rao and Massimiliano Esposito. “Nonequilibrium Thermodynamics of Chemical Reaction Networks : Wisdom from Stochastic Thermodynamics”. In: *Phys. Rev. X* 6 (2016), p. 041064.

[32] Nilesh Vaidya et al. “Spontaneous network formation among cooperative RNA replicators”. In: *Nature* 491.7422 (Oct. 2012), pp. 72–77.

Books

[4] S.K. Scott. *Chemical chaos*. Oxford: Clarendon Press, 1991.

[22] Ron Milo and Rob Philips. *Cell Biology by the Numbers*. CRC Press, 2015.

[24] Francis A. Carey and Richard J. Sundberg. *Advanced Organic Chemistry Part A: Structure and Mechanisms*. New York: Springer, 2006.

[25] R Aris. *Elementary Chemical Reactor Analysis*. Mineola NY: Dover, 1989.

[26] P Bergé, Y Pomeau, and Ch. Vidal. *L'ordre dans le chaos*. Paris: Hermann, 1984.

[27] Ch. Vidal and H Lemarchand. *La réaction créatrice: Dynamique des systèmes chimiques*. Paris: Hermann, 1988.

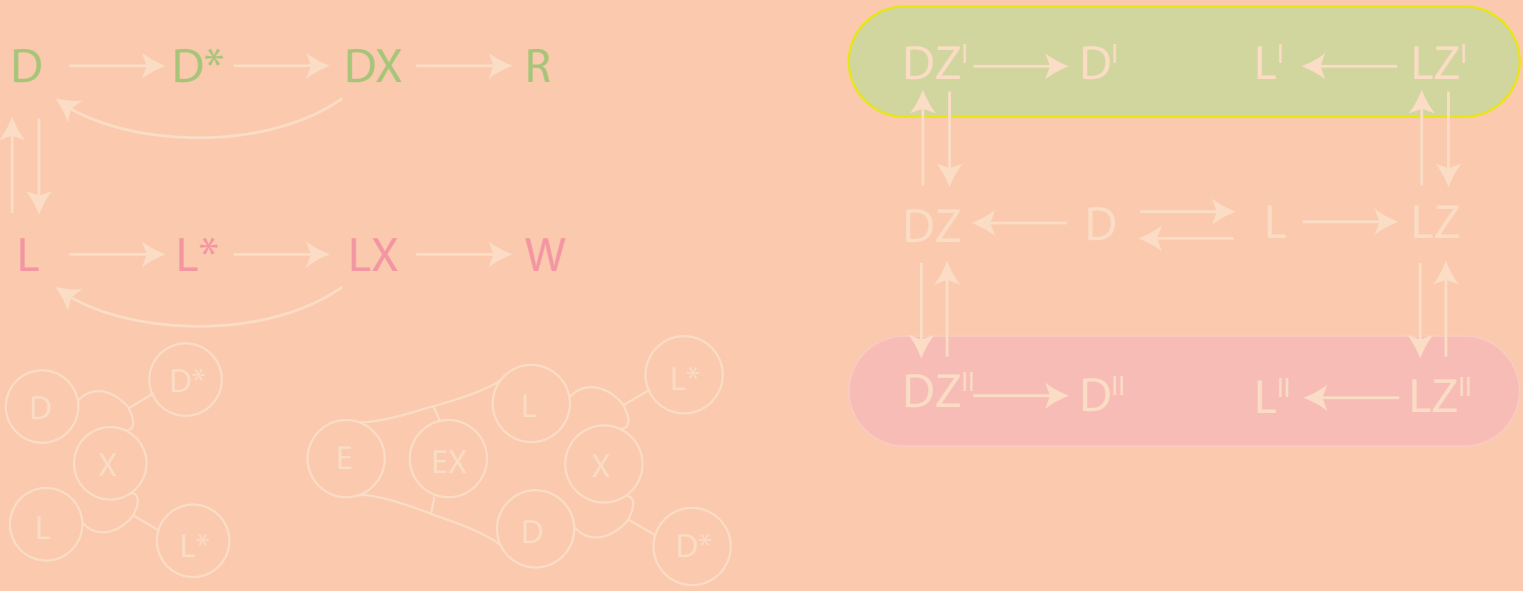
[28] G Nicolis. *Introduction to nonlinear science*. Cambridge, UK: Cambridge University Press, 1995.

[29] I.R. Epstein and J.A. Pojman. *An Introduction to Nonlinear Chemical Dynamics*. New York: Oxford University Press, 1998.

[30] I. Prigogine. *Introduction to Thermodynamics of Irreversible Processes*. New York: Wiley, 1967.

[33] Steven H Strogatz. *Nonlinear Dynamics and Chaos*. New York: Perseus Books, 1994.

[34] Dilip Kondepudi and Ilya Prigogine. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. New York: Wiley, 1998.



4. Information in chemical networks

In this section, we will illustrate the concept of information in the context of statistical physics, nonequilibrium thermodynamics and chemical networks. The notion of information has a prominent role in all Origins of Life research communities, but it is given a variety of meanings.

In chemistry, an intuitive understanding is often employed to communicate concepts and explain chemical systems. Currently, a correct 'chemical intuition' is lacking for information, nonequilibrium thermodynamics and large chemical networks. We hope to improve on this situation, by providing examples and thought experiments to connect them.

We will start (Sec. 4.1) with a short historical introduction to information, the Gibbs paradox and the role of coarse-graining in statistical physics and kinetics. The interest of reviewing these is pedagogical, a number of misconceptions concerning information in origins of life and chemistry can be readily cleared up when these concepts are made clear from the start. We deem it instructive to repeat these concepts before moving to the sections in which we present our own results.

By showing how a single-molecule information to work convertor can be scaled up to macroscopic size (Sec. 4.2), we show how molecular degrees of freedom and mixing entropy can be exploited to extract work. We also show how one can achieve dissipation-free full conversion of enantiomers to their mirror image. In doing so, these notions acquire a clear thermodynamic interpretation. Our convertors can be incorporated in a heat engine, or coupled to enantiomer reservoirs to make a two-stroke engine that continuously extracts work by racemizing enantiomers4.1.

In chemistry, we are typically interested in the reverse problem: introducing work to acquire a desired compound with high purity. In Sec. 10.3 we will first consider this problem on the level of single reactions, where we can distinguish between kinetic and thermodynamic products, which introduce fundamental limits on the purity we can achieve. We can go beyond these fundamental limits, when additional reactions come into play. We will review some interesting examples of existing error-correcting networks, and also consider some new ones. Each network has its own characteristic tradeoffs, which affect e.g. the complexity, processing time, dissipation and yield of the process.

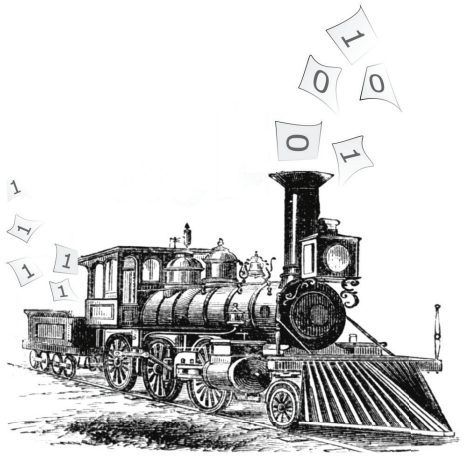


Figure 4.1: Using information as a fuel. Inspired by an original drawing by Luca Peliti

4.1 Some notions of information

Information theory was born with Claude Shannon's analysis of fundamental limits to transmission of a message through a channel in a paper titled 'A mathematical theory of communication'[1]. In the paper, a message is a sequence of symbols, encoded by bits, which is sent through a channel that can be noisy. A fundamental quantity introduced in the work was the entropy H, which averages over the logarithmic probabilities of messages

$$H = -\sum_i p_i \log_2 p_i. \tag{4.1}$$

By choosing a logarithmic base of 2, the information content of a message could be quantified, in terms of bits. In this measure, transitioning from total ignorance to full certainty that a system is in a state i, an information $-\log_2 p_i$ is acquired. On average, an information equal to Eq. (4.1) is acquired by obtaining full knowledge of a state or a message. In this sense, Information quantifies a 'surprise value'. If p_i is close to 1, little is learned, since it was close to certain anyway.

When messages are sent through a channel with finite capacity (bits/second), they are encoded by a sequence of symbols, which in turn can be encoded by a sequence of bits. For an optimal transfer of information per bit (and per symbol), messages carrying little information should be as short as possible (since they will be sent very often) and messages rich in information can be longer. Shannon's paper showed that optimal codes exist to achieve this, in noiseless channels and in noisy ones.

The average number of bits needed per symbol (for lossless transmission, independent symbols), is also quantified by an entropy H on the level of a single symbol, with p_i the probability for this symbol to be i. E.g. consider we wish to learn the content of a string of length N, composed of an alphabet of n symbols that have equal probability of occurring, s.t. $\forall i \;\; p_i = 1/n$. If we were to arbitrarily generate such a string, we would come up with the same one with a probability

$$p = \left(\frac{1}{n}\right)^N. \tag{4.2}$$

The information we acquire by learning what the content of the string is, is $N\log_2 n$ bits.

A link with thermodynamic entropy was given in a very punctual manner: *The form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics where p_i is the probability of a system being in cell i of its phase space.*

Information theory (back then: communication theory) quickly caught on, also outside of electrical engineering. Biologists, physicists, psychologists, economists and practitioners of other disciplines started using concepts from the theory. In his 1956 paper 'the bandwagon' [2], Shannon remarked that some moderation was in order *In the first place, workers in other fields should realize that the basic results of the subject are aimed in a very specific direction, a direction that is not necessarily relevant to such fields as psychology, economics, and other social sciences. Indeed, the hard core of information theory is, essentially, a branch of mathematics, a strictly deductive system. A thorough understanding of the mathematical foundation and its communication application is surely a prerequisite to other applications.*

Indeed, we should be careful not to confuse information in the entropic sense with the colloquial uses of the term information. In the latter case, one considers information as something that must be pertinent. Information theory makes no such value statements, it attempts to find fundamental limits to the transmission of encoded messages, whether they have meaning or not.

4.1.1 Information and the Gibbs Paradox

The Gibbs paradox is, like many 'paradoxes', only an apparent one. It becomes paradoxical when one takes an incomplete starting point in the derivation of thermodynamics. Gibbs showed that such an incomplete starting point could be the expression for the entropy of an ideal gas [3], which did not consider permutations between molecules. Starting from the ideal gas law, and supposing variations in energy ε proportional to temperature, Gibbs writes

$$pV = aT, \tag{4.3}$$

$$\varepsilon = cT + E \tag{4.4}$$

with a, c and E constants and V a volume. Eliminating pressure p and temperature T for the total differential of $d\varepsilon$, it is found that

$$d\varepsilon = \frac{\varepsilon - E}{c} d\eta - \frac{a}{V}\frac{\varepsilon - E}{c} dV, \tag{4.5}$$

where η is an entropy. Gibbs then rewrites the expression and integrates, to ultimately yield an expression for the entropy of the form

$$S = -c \log \frac{\varepsilon - E}{c} + \eta - a \log V. \tag{4.6}$$

Gibbs then considered a thought experiment of initially separated gases mixing by diffusion, which is often paraphrased in the following manner: There are two chambers filled with gas, I and II, maintained at identical pressure p and with respective volumes V^{I} and V^{II}. The gases have corresponding entropies S^{I} and S^{II}. An extensive total entropy S for the two chambers that follows Eq. (4.6) is then

$$S = S^{\mathrm{I}} + S^{\mathrm{II}} = m_{\mathrm{I}} a_{\mathrm{I}} \log V^{\mathrm{I}} + m_{\mathrm{II}} a_{\mathrm{II}} \log V^{\mathrm{II}}. \tag{4.7}$$

Where m is a measure of quantity (e.g. mass, but Gibbs underlines that one chooses dimensions on a case by case basis) and a an entropic prefactor as found in Eq. (4.6). Via the ideal gas law, m and a are related, here given by

$$ma = \frac{pV}{2T}. \tag{4.8}$$

Let us now open a little door connecting the chambers and let them reach equilibrium. Choosing $V^{\mathrm{II}} = V^{\mathrm{I}}$ and $V = V^{\mathrm{I}} + V^{\mathrm{II}}$, Gibbs noted that this led to an entropy change

$$\Delta S = m_{\mathrm{I}} a_{\mathrm{I}} \log V + m_{\mathrm{II}} a_{\mathrm{II}} \log V - m_{\mathrm{I}} a_{\mathrm{I}} \log \frac{V}{2} + m_{\mathrm{II}} a_{\mathrm{II}} \log \frac{V}{2} = (m_{\mathrm{I}} a_{\mathrm{I}} + m_{\mathrm{II}} a_{\mathrm{II}}) \log 2. \tag{4.9}$$

Such that

$$\Delta S = \frac{pV}{2T}. \tag{4.10}$$

Noting that this entropic increase was not accompanied by an increase in energy, Gibbs pointed out the act of undoing the mixing could be linked to energetic transformations: *When we say that when two different gases mix by diffusion, as we have supposed, the energy of the whole remains constant, and the entropy receives a certain increase, we mean that the gases could be separated and brought to the same volume and temperature which they had at first by means of certain changes in external bodies, for example, by the passage of a certain amount of heat from a warmer to a colder body.*

The apparent paradox was introduced when the gases that mixed were of the same nature, 'indistinguishable', such that the mixed state was equivalent to the unmixed one and should therefore not provide an entropy change if Eq. (4.9) was applied. It was then pointed out that the paradox was easily averted, by adding a new contribution to the entropy balance. Denoting the new expression S', this corresponded to

$$S' = S - \sum_i n_i \ln n_i. \tag{4.11}$$

The mixing paradox is first and foremost a thought experiment to find a consistent functional form for entropy in classical thermodynamics. Once this form is found, it ceases to be a paradox. As will be reviewed in the next section, this functional form has a straightforward interpretation in statistical mechanics.

The Gibbs Paradox in statistical mechanics

In statistical mechanics, the additional term in Eq. (4.11) has a combinatorial origin. For a given system state with molecular occupations $\boldsymbol{n} = \{n_1, n_2, ..., n_s\}$, and total number of molecules

$$N = \sum_i n_i, \tag{4.12}$$

we can attribute a degeneracy $z(\boldsymbol{n})$, corresponding to permutations among molecules of the same type

$$z(\boldsymbol{n}) = \frac{N!}{\Pi_i n_i!}. \tag{4.13}$$

This degeneracy is first and foremost a choice of convenience to describe macrostates. Alternatively, we could have insisted on treating every species separately, and indicate the position of the jth particle of type i by its corresponding index of a volume cell, given by $x_{i,j}$, such that a system state is characterized by the vectors $\boldsymbol{n},\boldsymbol{x}$ and the probability of a state by $p(\boldsymbol{n},\boldsymbol{x})$.

However, if our model makes no meaningful distinction between species of the same type, other than fixing labels, we automatically have a permutation symmetry $\forall i\ x_{i,j} \leftrightarrow x_{i,k}$, such that

$$p(\boldsymbol{n}, \{x_{1,1}, ..., x_{i,j}, x_{i,j+1}, ..., x_{s,m_s}\}) = p(\boldsymbol{n}, \{x_{1,1}, ..., x_{i,j+1}, x_{i,j}, ..., x_{s,m_s}\}) \tag{4.14}$$

Furthermore, no volume cell is privileged, from which it follows that $\forall i,k\ x_{i,j} \leftrightarrow x'_{i,j}$, where x' is an arbitrary different volume cell

$$p(\boldsymbol{n}, \{x_{1,1}, ..., x_{i,j}, x_{i,j+1}, ..., x_{s,m_s}\}) = p(\boldsymbol{n}, \{x_{1,1}, ..., x'_{i,j}, x_{i,j+1}, ..., x_{s,m_s}\}) \tag{4.15}$$

Let us now consider this collection of particles being partitioned in small volume cells that compose two large compartments of size V^{I} and V^{II}. If we disregard symmetries and there are no interactions or chemical transformations, then any valid configuration vector $\boldsymbol{n},\boldsymbol{x}$ is equally probable. We

describe equally probable microstates. In classical thermodynamics, maximum entropy is equated with the most likely macrostate, for which this level of detail must be considered in a more coarse-grained manner. For example, if we enquire about a macroscopic observable, such as the average number of particles of type 1 that are in compartment II, we need to consider

$$[n_1^{\mathrm{II}}] = \sum_{i=1}^{s} \sum_{s_i=1}^{m_i} \sum_{x_i \in \{V^{\mathrm{I}}, V^{\mathrm{II}}\}} \chi_{x_i}^{V^{\mathrm{II}}} p(\boldsymbol{n}, \{x_{1,1}, ..., x_{i,j}, x_{i,j+1}, ..., x_{s,m_s}\}), \tag{4.16}$$

where $\chi_{x_i}^{V^{\mathrm{II}}}$ is an indicator function that is zero for every volume cell outside of V^{II}

$$\chi_{x_i}^{V^{\mathrm{II}}} = \begin{cases} 0 & x_i \notin V^{\mathrm{II}} \\ 1 & x_i \in V^{\mathrm{II}} \end{cases} \tag{4.17}$$

Using the symmetries (4.14) and (4.15), this simplifies to

$$[n_1^{\mathrm{II}}] = n_1 \frac{V^{\mathrm{II}}}{V^{\mathrm{I}} + V^{\mathrm{II}}}. \tag{4.18}$$

Alternatively, we could have absorbed the symmetries directly in our description and opted for a more coarse-grained probability distribution, $p(\boldsymbol{n}^{\mathrm{I}}, \boldsymbol{n}^{\mathrm{II}})$,

$$p(\boldsymbol{n}^{\mathrm{I}}, \boldsymbol{n}^{\mathrm{II}}) = \frac{N!}{\Pi_i n_i^{\mathrm{I}}! n_i^{\mathrm{II}}!} \left(\frac{V^{\mathrm{I}}}{v}\right)^{n^{\mathrm{I}}} \left(\frac{V^{\mathrm{II}}}{v}\right)^{n^{\mathrm{II}}} p(\boldsymbol{n}, \boldsymbol{x}), \tag{4.19}$$

with v volume of a volume cell. Unlike our former microstates, the coarse-grained states are not equally probable. For a large system ($N \to \infty$) a small subset of coarse-grained states ($n_i^{\mathrm{I,II}} \approx [n_i^{\mathrm{I,II}}]$) will come to dominate the distribution. Naturally, a system in an arbitrary microstate will, on average, be drawn towards the occupation of these most abundant coarse-grained states. In information terms, it will move to the 'least surprising' configurations and maximize the entropy.

The operation of coarse-graining similar states thus provides us with an entropy that operates in accord with classical thermodynamics. The degeneracy $\Omega(\boldsymbol{n})$ is sometimes wrongly thought (and taught) to be a direct consequence of quantum mechanics, but more generally it can be shown that the origin is a specific choice of coarse-graining [4]. Particularly striking is the example of colloidal dispersions [5], where arguably no two colloids are the same on the atomic level. However, the level of description of colloid thermodynamics normally is not concerned with such details. In such a description, colloids judged to be sufficiently similar (e.g. occurring within a size interval of $[r, r+\delta r]$) are binned as phenomenologically equivalent species. Within the ensuing description, we then have an exchange symmetry (4.14) between these species.

Coarse graining in kinetics

An example more pertinent to chemical networks comes when considering chemical kinetics. Suppose we have a reaction

$$2\mathrm{A} \rightleftharpoons \mathrm{B} + \mathrm{C}, \tag{4.20}$$

taking the system from the microstate $\boldsymbol{n} = \{n_A, n_B, n_C\}$ to $\boldsymbol{n}' = \{n_A - 2, n_B + 1, n_C + 1\}$, by a reaction that follows mass-action kinetics. In a Master equation framework, detailed balance imposes that the transition rates W between a pair of such states verifies

$$\frac{W_{\boldsymbol{n} \to \boldsymbol{n}'}}{W_{\boldsymbol{n}' \to \boldsymbol{n}}} = \frac{z(\boldsymbol{n}')}{z(\boldsymbol{n}')} \exp(-\beta \Delta \mu^{\circ}) \tag{4.21}$$

Injecting

$$z(\boldsymbol{n}) = \frac{N!}{n_A!n_B!n_C!}, \qquad z(\boldsymbol{n}) = \frac{N!}{n_A!n_B!n_C!} \tag{4.22}$$

this becomes

$$\frac{W_{\boldsymbol{n}\to\boldsymbol{n}'}}{W_{\boldsymbol{n}'\to\boldsymbol{n}}} = \frac{n_A(n_A-1)}{(n_B+1)(n_C+1)} \exp\left(-\beta\Delta\mu^\circ\right). \tag{4.23}$$

The degeneracy ratio $z(\boldsymbol{n}')/z(\boldsymbol{n}')$ elegantly mirrors the reaction propensities used in a typical mean-field description of kinetics

$$\frac{dn_A}{dt} = k^+ n_A(n_A-1) - k^- n_B n_C, \tag{4.24}$$

where volume dependencies are absorbed in the rate constants k^+, k^-.

Coarse graining is an essential part of a thermodynamic and kinetic description. By extension, both descriptions may make similar approximations with respect to species similarity. In practice, even molecules that are considered 'the same' can vary considerably in isotopic composition: 1.07% of carbon atoms is a ^{13}C isotope, 0.368% of nitrogen atoms is ^{15}N, 0.205% of oxygen is ^{18}O, 4.29% of sulfur is ^{34}S, 24.22% of chlorine is ^{37}Cl and so forth.

It is not always appropriate to disregard the presence of such isotopes, especially when their abundance is artificially increased. Heavier isotopes can slow down reactions, which is especially pronounced for light atoms like hydrogen when being substituted by deuterium. For rate limiting steps involving the displacement of hydrogen (e.g. proton abstraction H^+, hydride shift H^-), their replacement by deuterium (D) can decrease the rate by a factor 2 to 10. When a reaction proceeds by tunneling, the retardation can amount to a factor of 50, as recently observed for the hydride shift in the formose reaction [6]. This phenomenon is known as the kinetic isotope effect and it is an important tool in the elucidation of reaction mechanisms. Isotopes can also shift equilibrium constants: the dimerization constant for hydrated formaldehyde was found to be 1.4 times higher in D_2O compared to H_2O[7].

Framed in this way, the Gibbs paradox is no paradox, but a macroscopic example of information loss in thermodynamics[4]: by mixing, particles lose their correlation with their initial half compartment. The mixing of distinct gases is irreversible and thereby yields an entropy production. This may happen to many degrees of freedom that escape our notice in a coarse-grained description (e.g. racemization of molecules, interconversion of molecular configurations).

To see how the Gibbs paradox fits in the bigger picture of heat, work and information, Sec. 4.2 describes a scalable information engine that extracts work from chiral molecules in a controlled fashion, by erasing their initial configuration. In the current literature, information processing in thermodynamics is often interpreted on the scale of single molecules or colloids[8, 9, 10, 11]. The machine we propose can equally well be treated in this regime, where it is a single-molecule engine that creates a strong correlation between particle position and a degenerate particle state, which is then exploited to extract work. By increasing the number of particles, a natural quantity to treat the problem becomes the mixing entropy. This raises a semantic issue in the field of stochastic thermodynamics, since we normally do not speak of macroscopic engines as information engines.

4.2 Macroscopic information engines

To make some notions of information more intuitive, we describe here a macroscopic system that reversibly erases the information of a chiral state **S** of an asymmetric molecule with a single stereocenter. This erasure happens through a racemization reaction: $\mathbf{S} \rightleftharpoons \mathbf{R}$, where **R** is the

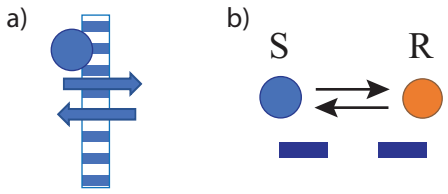


Figure 4.2: a) a membrane that only lets **S** (blue) through. b) Racemization catalyst (dark blue bars) enables the reaction $\mathbf{S} \rightleftharpoons \mathbf{R}$.

mirror image of **S**. By coupling this reaction to a membrane exchange process for **S**, we are able to generate a pressure difference, which we exploit to extract work from a heat bath. The principal components for our machine are: (i) An enantiospecific membrane, which lets only one enantiomer pass through, and (ii) a racemization catalyst, which enables the reaction $\mathbf{S} \rightleftharpoons \mathbf{R}$ (see Fig.4.2). From which we construct the system depicted in Fig. (4.4).

4.2.1 The single-molecule case

Let us first develop an intuition for a single-molecule case, illustrated in Fig. 4.3. There, we have two pistons that contain a single-molecule gas between two compartments separated by a membrane that only lets **S** through. A racemization catalyst, only present in compartment II, performs the reaction $\mathbf{S} \rightleftharpoons \mathbf{R}$. The whole system is in contact with a thermal bath (not shown) maintained at a temperature T.

The timescale of the racemization reaction is much faster than membrane exchange, which means that if the molecule is in compartment II, we will no longer know if it is in the S or R state (which are equally likely, since for enantiomers in an isotropic environment we have $\mu_{\mathrm{S}}^{\circ} = \mu_{\mathrm{R}}^{\circ}$). The molecule will on average spend the same time in either state, and we can treat the molecule as being in a 'mixed S / R state'.

Since only the S enantiomer can cross the membrane, the typical residence time of a mixed-state molecule is twice that of a pure S species. Consequently, the species will spend more time exerting pressure on the piston in chamber II than in I. As we will demonstrate more precisely in the rest of the next sections, this pressure asymmetry can be used to extract a work that exactly compensates the information loss due to racemization.

As a single-molecule information-to-energy converter, the device is reminiscent of Szilard's engine[8]. In Szilard's engine, information on the particles position is exploited by the introduction of a partition and the subsequent extraction of work. Ouldridge, Brittain and ten Wolde[12] have rephrased Szilard's original argument for the functioning of the engine in terms of nonequilibrium thermodynamics: such a process requires the introduction of a correlation between the particle position and other degrees of freedom that intervene in the process. A more mechanistic view of where this correlation comes from is often left implicit, one often treats it as a black box colloquially referred to as a 'demon'.

In our example in Fig. 4.3, the correlation processes are made explicit: when a molecule passes from compartment II to I, it must be in the S state, and thereafter remain in this state. This creates a correlation between chiral purity and location. When entering compartment II, contact with the catalyst rapidly correlates the position with the mixed state S/R. Unlike the Szilard engine, we are not dependent on one measurement followed by an extraction (or in case of a faulty measurement: a loss). Instead, we can go back and forth between these correlated states many times, to gradually extract a work of $kT \ln 2$ from the heat bath.

In principle, we can extend Szilard's engine to N molecules and have it extract work when all molecules are thought to be in one half of the room. However, observing such a state (which has a probability weight of $(1/2)^N$) in finite time becomes rapidly unfeasible with growing N.

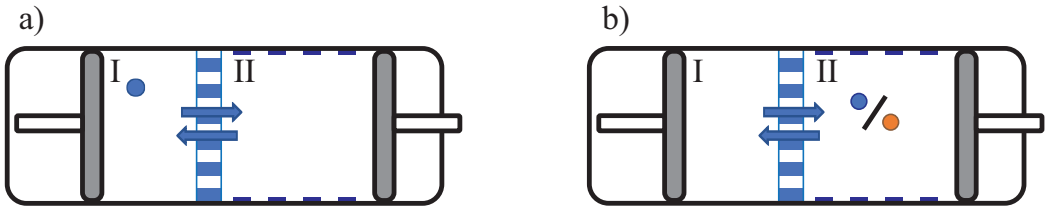


Figure 4.3: An illustration of the single-molecule case. a) a pure S state (blue) is indicative of residence in compartment I. b) a mixed S / R state (due to rapid interconversion) corresponds to residence in compartment II. The correlation between chiral state and position is mediated through the selective membrane and racemization catalyst.

Our continuous protocol in Fig. 4.3, on the other hand, is scalable: we can extract work from racemization for multiple species in parallel, as shown in Fig. 4.4. In the next section, we will make this thermodynamically explicit.

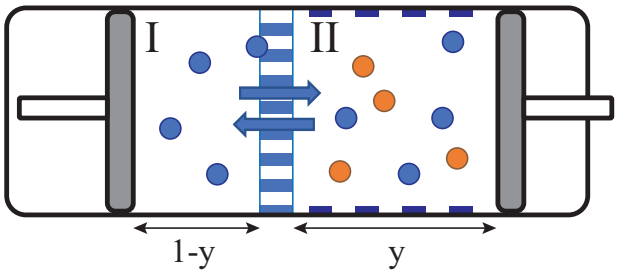


Figure 4.4: A scaled-up information-to-work converter.

4.2.2 Extraction Protocol

Let us now describe an isochoric protocol for the extraction of work from racemization, in which the pistons are moved in concerted fashion. We start with two compartments (labeled I and II in Fig. 4.4), separated by an enantiospecific membrane, which is specific to **S**. A racemization catalyst, only present in compartment II, performs the reaction S $\rightleftharpoons$ R. At the start of our protocol, II is fully compressed and I contains only **S**. During the protocol, II expands and I is compressed.

For our isochoric protocol, the total volume, V_0, is constant

$$V^{\mathrm{I}} + V^{\mathrm{II}} = V_0, \tag{4.25}$$

where $V^{\mathrm{I}}, V^{\mathrm{II}}$ are compartment volumes of compartment I and II respectively. Wherever appropriate, we will specify compartment with an upper index I, II. Similarly, we can define $p^{\mathrm{I}}, p^{\mathrm{II}}$ as the total compartment pressures which verify

$$p^{\mathrm{I}} = p^{\mathrm{I}}_{\mathbf{S}}, \tag{4.26}$$
$$p^{\mathrm{II}} = p^{\mathrm{II}}_{\mathbf{S}} + p^{\mathrm{II}}_{\mathbf{R}}. \tag{4.27}$$

where $p_{\mathbf{S}}$ denotes a partial pressure. We can now define the chemical potential of a species j as

$$\mu_j = \mu_j^\circ + \ln\left(\frac{p_j}{p^\circ}\right). \tag{4.28}$$

Here, μ_j° is a standard free energy of formation at reference pressure p°. The protocol is performed quasistatically, such that membrane exchange of **S** reaches equilibrium: $\mu^{\mathrm{I}}_{\mathbf{S}} = \mu^{\mathrm{II}}_{\mathbf{S}}$, and from (4.28) it follows that

$$p^{\mathrm{I}}_{\mathbf{S}} = p^{\mathrm{II}}_{\mathbf{S}}. \tag{4.29}$$

By definition, any pair of enantiomers has the same free energy of formation: $\mu^{\circ}_{\mathbf{S}} = \mu^{\circ}_{\mathbf{R}}$. When the racemization reaction reaches chemical equilibrium, we obtain $\mu^{\mathrm{II}}_{\mathbf{S}} = \mu^{\mathrm{II}}_{\mathbf{R}}$ and therefore

$$p^{\mathrm{II}}_{\mathbf{S}} = p^{\mathrm{II}}_{\mathbf{R}}. \tag{4.30}$$

Substituting Eqs. (4.30) and (4.29) in Eqs. (4.26) and (4.27), we obtain

$$p^{\mathrm{II}} = 2p^{\mathrm{I}}. \tag{4.31}$$

As all our processes conserve the total number of molecules N, we have

$$n^{\mathrm{I}}_{\mathbf{S}} + n^{\mathrm{II}}_{\mathbf{S}} + n^{\mathrm{II}}_{\mathbf{R}} = N, \tag{4.32}$$

where n_j denotes the number of molecules of type j. Using the ideal gas law: $pV = nk_bT$, we can write

$$\frac{n^{\mathrm{I}}_{\mathbf{S}}}{V^{\mathrm{I}}} = \frac{n^{\mathrm{II}}_{\mathbf{S}}}{V^{\mathrm{II}}} = \frac{n^{\mathrm{II}}_{\mathbf{R}}}{V^{\mathrm{II}}}. \tag{4.33}$$

We now parametrize the piston displacement using y, (with $0 \leq y \leq 1$), such that

$$V^{\mathrm{I}} = V_0(1-y), \tag{4.34}$$
$$V^{\mathrm{II}} = V_0 y. \tag{4.35}$$

Plugging (4.35) in (4.33), we obtain

$$n^{\mathrm{I}}_{\mathbf{S}} = n^{\mathrm{II}}_{\mathbf{S}} \frac{1-y}{y}. \tag{4.36}$$

If we now apply Eq. (4.32), we arrive at the expressions

$$n^{\mathrm{I}}_{\mathbf{S}} = N\frac{1-y}{1+y}, \tag{4.37}$$
$$n^{\mathrm{II}}_{\mathbf{S}} = N\frac{y}{1+y}. \tag{4.38}$$

This allows to obtain the pressure as function of y, from which we calculate work extraction during the entire protocol. For compartment I, we write:

$$p^{\mathrm{I}}(y) = p^{\mathrm{I}}_{\mathbf{S}} = \frac{k_bTn^{\mathrm{I}}_{\mathbf{S}}}{V_0(1-y)} = \frac{k_bTN}{V_0(1+y)}. \tag{4.39}$$

The work applied on the system is expressed as

$$W = -\int_{\mathrm{I,II}} pdV = -\int p^{\mathrm{II}}dV^{\mathrm{II}} - \int p^{\mathrm{I}}dV^{\mathrm{I}} = \int (p^{\mathrm{I}} - p^{\mathrm{II}})dV^{\mathrm{II}}, \tag{4.40}$$

where we have used the fact that $dV^{\mathrm{I}} = -dV^{\mathrm{II}}$, since the protocol is isochoric. The net work applied is then

$$\begin{aligned} W &= \int_0^1 (p^{\mathrm{I}}(y) - p^{\mathrm{II}}(y))V_0 dy = -\frac{k_bTN}{V_0}\int_0^1 \frac{dy}{1+y} \\ &= -k_bTN\ln(1+y)|_0^1 = -k_bTN\ln 2. \end{aligned} \tag{4.41}$$

Where we have used Eqs. (4.39) and (4.31). As the applied work is negative, there is a net extraction of $k_bTN\ln 2$ of racemization work in the reversible erasure of the configuration.

4.2.3 Information erasure

The racemization leads to a loss of information with regard to the chiral state of a given molecule, which can now be either an **S** or **R** enantiomer with equal probability. Let $P_{\mathbf{S}}(y)$ denote the probability for a particle to be in state **S** when the protocol is at y. We can then express the change of information between $y = 0$ and $y = 1$ as

$$\Delta I = k_b \left(P_{\mathbf{S}}(1) \ln \frac{P_{\mathbf{S}}(1)}{P_{\mathbf{S}}(0)} + P_{\mathbf{R}}(1) \ln \frac{P_{\mathbf{R}}(1)}{P_{\mathbf{S}}(0)} \right) = -k_b \ln 2 \tag{4.42}$$

Generally, the extractable work from information obeys

$$W_{extr} \leq T \Delta I, \tag{4.43}$$

which in our example becomes an equality due to quasi-state assumption.

4.2.4 Macroscopic engines

A heat-information engine

In order to make an engine, we can introduce a second heat bath at temperature T^*. We can then define the following cyclic protocol : I) Let y go from 0 to 1 at T (performing $\mathbf{S} \rightarrow \mathbf{S} + \mathbf{R}$). II) adiabatically expand the gas, until it reaches a temperature T^* III) Let y go from 1 to 0 at T^*. IV) adiabatically compress the gas, until it reaches a temperature T. This protocol yields a heat engine reminiscent of a Carnot engine where isothermal volume expansion has been replaced by isothermal information erasure. In step I), a work

$$W_{\mathrm{I}\rightarrow\mathrm{II}} = k_b T N \ln 2 \tag{4.44}$$

is extracted. Since the internal energy U does not change, it follows from the first law of thermodynamics that a heat

$$Q_{\mathrm{I}\rightarrow\mathrm{II}} = -k_b T N \ln 2 \tag{4.45}$$

flows into the system from the heat bath. In a Carnot engine, the corresponding operation would be isothermal expansion of a volume V to $2V$

$$W = \int_V^{2V} P dV = \int_V^{2V} k_b T N \frac{dV}{V} = k_b T N \ln 2. \tag{4.46}$$

The isothermal expansion step is analogous to the racemization step, as it increases the number of system configurations and thus increases our ignorance of the position of a particle. Doubling the volume or racemizing an asymmetric molecule are fundamentally doing the same thing.

In the subsequent adiabatic expansion, the gas is expanded until a temperature T^* is reached. No heat leaves the system, $dQ = 0$. It follows from the first law ($dU = dQ + dW$) that the energy change that accompanies this cooling is fully converted to work

$$dW = dU, \tag{4.47}$$

such that

$$dU = N c_V dT = \frac{c_V}{k_b} d(PV), \qquad W = \int_{V_0}^{V_f} P dV = N c_V (T^* - T). \tag{4.48}$$

Where c_v is the molecular heat capacity at constant volume, defined by

$$c_v = \frac{1}{N} \left(\frac{dU}{dT} \right)_{N,V}. \tag{4.49}$$

Let us denote by $\alpha = c_V/k_b$, after which we inject the results in the first law to yield

$$-(\alpha+1)PdV = \alpha V dP. \tag{4.50}$$

Dividing by PV and integrating leads to

$$\ln\left(\frac{P_f}{P_i}\right) = -\frac{\alpha+1}{\alpha}\ln\left(\frac{V_f}{V_i},\right) \tag{4.51}$$

from which it automatically follows that

$$PV^\gamma = cte \tag{4.52}$$

with $\gamma = \frac{\alpha+1}{\alpha}$. Using $P_i V_i^\gamma = P_f V_f^\gamma$, we find

$$W = \int_{V_i}^{V_f} P_i\left(\frac{V_i}{V}\right)^\gamma dV = -\alpha N k_b T_i\left(\left(\frac{V_f}{V_i}\right)^{1-\gamma} - 1\right) \tag{4.53}$$

Equating Eq. (4.53) to Eq. (4.48) then allows to extract the pressure-temperature relation for this protocol

$$T = T_i\left(\frac{P}{P_i}\right)^{\frac{\gamma-1}{\gamma}}. \tag{4.54}$$

Subsequently, in step III), we couple the system to a heat bath at temperature T^*, in which we isothermally reverse the racemization, which is followed by an injection of work that flows directly back to the heat bath

$$W_{\mathrm{III}\to\mathrm{IV}} = -k_b T^* N \ln 2, \tag{4.55}$$

$$Q_{\mathrm{III}\to\mathrm{IV}} = k_b T^* N \ln 2. \tag{4.56}$$

In the final step, the system is adiabatically compressed, until a temperature T is acquired. This is the exact reverse of step II),

$$W_{\mathrm{IV}\to\mathrm{I}} = -N c_v (T^* - T). \tag{4.57}$$

Since $W_{\mathrm{II}\to\mathrm{III}} + W_{\mathrm{IV}\to\mathrm{I}} = 0$, step II) and IV) lead to no net work extraction. From the heat $Q_{\mathrm{I}\to\mathrm{II}}$ extracted in step I), the net work extracted is $W_{\mathrm{I}\to\mathrm{II}} + W_{\mathrm{III}\to\mathrm{IV}}$. The overall energy balance for the cycle is

$$Q_{\mathrm{I}\to\mathrm{II}} = W_{\mathrm{I}\to\mathrm{II}} + W_{\mathrm{III}\to\mathrm{IV}} + Q_{\mathrm{III}\to\mathrm{IV}}. \tag{4.58}$$

The efficiency $\eta = \left(W_{\mathrm{I}\to\mathrm{II}} + W_{\mathrm{III}\to\mathrm{IV}}\right)/Q_{\mathrm{I}\to\mathrm{II}}$ then yields the form

$$\eta = 1 - \frac{T^*}{T}. \tag{4.59}$$

An autonomous information engine

Alternatively, one can employ **S** directly as a fuel. This then requires an engine that takes in **S** and releases **S + R** as exhaust (while being in contact with a heat bath). This can be done in a controlled fashion, by exchange with reservoirs of **S** and **S** + **R**. In Fig. 4.5, a constant-volume protocol is illustrated in which two converters are coupled. Provided the pressure in the S reservoir is slightly higher than in the S + R reservoir, the engine can run autonomously at constant volume and temperature.

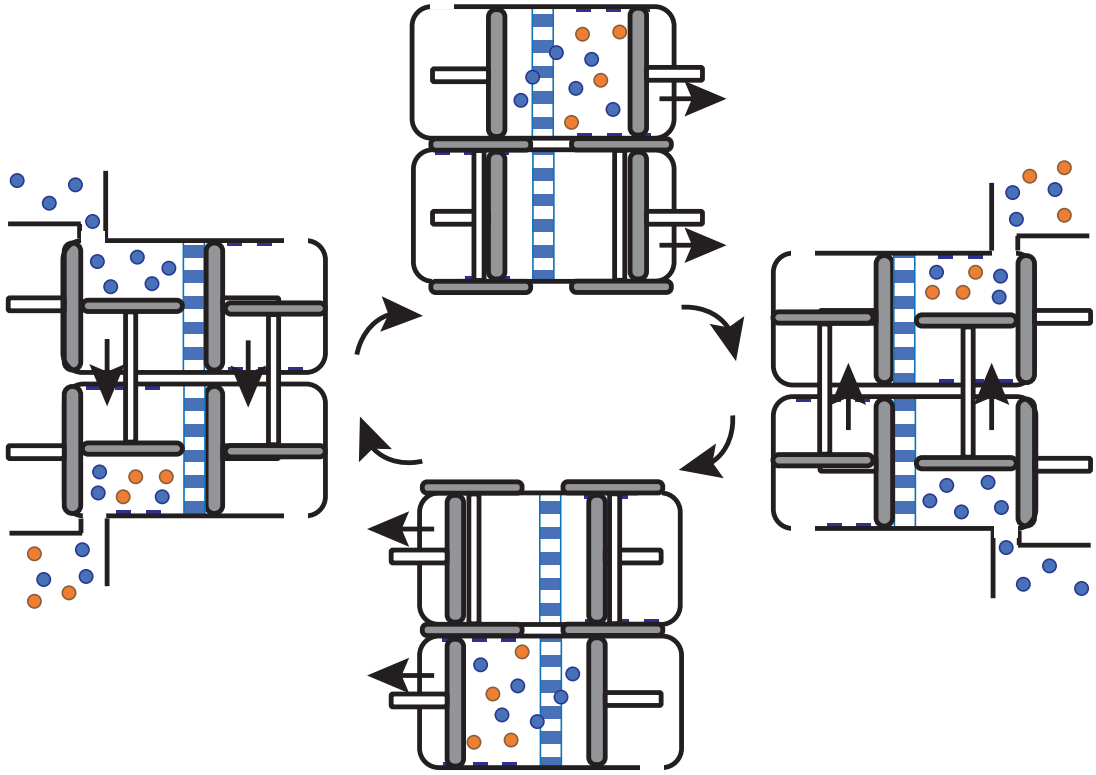


Figure 4.5: A macroscopic information engine harvesting work from reservoirs. A constant-volume intake of S from a reservoir by one converter is accompanied by the expulsion of S + P in another reservoir. Work extraction from racemization is accompanied by resetting the piston position of the other compartment. After two steps, the system finds itself in an equivalent, mirrored state and the converters exchange roles.

One should keep in mind, however, that the released energy per molecule is very small. A readily available source of chiral fuel is ibuprofen (a simple molecule with one stereocenter) (206.29 g/mol), a 200 mg tablet contains ≈ 1 mmol stereocenters, which can be converted to $W \approx 1.7$ J. The energy density per kg ibuprofen due to stereocenters is then 0.0084 MJ/kg, compared to 48 MJ/kg for diesel due to combustion*. This discrepancy is to be expected: During combustion, a large number of new, stable bonds are formed. Chemical bonds can provide energies around $100kT$ whereas stereocenters can provide at most $kT \ln 2$.

4.2.5 Thermodynamically reversible chemical reactions

In the following, we use a slightly modified version of the racemization engine to reversibly perform a chemical transformation of one pure enantiomer to another. Interestingly, this allows us to make 'thermodynamically reversible' chemical reactions. For a regular chemical reaction (constant volume, pressure) the free energy function G has a single minimum, which corresponds to chemical equilibrium. In a multicompartment system with movable walls, (see Fig.4.6) this no longer needs to be true, as the system can be set up such that mixing entropy (or equivalently: information loss) is compensated or even absent.

The setup in Fig.4.6 contains three compartments (labeled I, II and III). I and II are in contact through an enantiospecific membrane for **S**, II and III through an enantiospecific membrane for **S**.

*In the early days of the automobile, both fuel and medicine were bought at the pharmacy. The former however, came in 10 L cans.

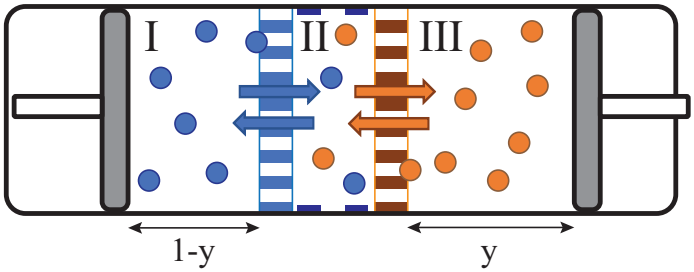


Figure 4.6: Ideal purification setup. The middle compartment (II) contains a racemization catalyst and is in contact with I through an **S**-specific membrane and with III through an **R**-specific membrane.

A pressure p is exerted by both pistons. By making the same assumptions as in the former section (equilibrium of diffusion, racemization equilibrium), we have

$$\mu_{\mathbf{S}}^{\mathrm{I}} = \mu_{\mathbf{S}}^{\mathrm{II}} = \mu_{\mathbf{R}}^{\mathrm{II}} = \mu_{\mathbf{R}}^{III}. \tag{4.60}$$

Then, from Eq. (4.28) it immediately follows that

$$p^{\mathrm{II}} = 2p^{\mathrm{I}} = 2p^{\mathrm{III}}. \tag{4.61}$$

For an isochoric protocol, upon shifting y by Δy, compartment II remains unaltered. In doing so, a net transfer between I and III takes place: $\Delta n_{\mathbf{S}}^{\mathrm{I}} = -\Delta n_{\mathbf{R}}^{\mathrm{III}}$. Note that, in this chemical transformation, no actual mixing has taken place, we have reduced the quantity of a gas of pure S (while maintaining the same pressure, (4.61)) and increased by that same quantity a gas of pure R.

The accompanying free energy change is: $\Delta G = \Delta n_{\mathbf{S}}^{\mathrm{I}} \mu_{\mathbf{S}}^{\circ} + \Delta n_{\mathbf{R}}^{\mathrm{III}} \mu_{\mathbf{R}}^{\circ} = 0$. Since we can keep the reaction contained in an arbitrarily small 'racemization zone' (here: chamber II), we can reversibly achieve any desired conversion y while expending negligible amounts of energy.

At present, the perfectly selective membranes invoked here are hardly feasible for most simple compounds, but more convoluted strategies to achieve the same (e.g. enantioselective modification due to enzymes at an interface to solubilize a desired enantiomer in a liquid membrane) have been developed. An important new dissipative technique to perform such separations is Viedma ripening[13], where autocatalytic feedback in crystal growth and redissolution (due to grinding) coupled with racemization in solution leads to the rapid accumulation of highly enantiopure crystals.

4.2.6 Liquid-phase racemization

The racemization engine can also be realized with liquid solutions instead of gases. Such a device is shown in Fig. 4.7. We will now proceed with a general proof, showing that such a setup can be used to reversibly extract the full work corresponding to the transformation. Suppose we have $s^{\mathrm{I}}+1$ different species in the leftmost compartment, whose number of molecules we will denote by $\boldsymbol{n}^{\mathrm{I}}$: $\{n_0^{\mathrm{I}}, n_1^{\mathrm{I}}, n_2^{\mathrm{I}}, .., n_k^{\mathrm{I}}, .., n_{s^{\mathrm{I}}}^{\mathrm{I}}\}$, where the solvent has the label 0. These molecules can all cross the membrane. We will furthermore suppose there is a catalytic species in compartment II, which allows for $s^{\mathrm{II}} - s^{\mathrm{I}}$ other species to form through reactions of some or all of the former $s^{\mathrm{I}}+1$ species. These s^{II} new species cannot pass through the membrane and thus only exist in compartment II.

The stoichiometry of the chemistry is encoded in the stoichiometry matrix $\boldsymbol{\nu}$, which was introduced in chapter 2 (See also Refs. [14, 15, 16, 17]). An entry ν_{ik} in the matrix provides a sign and an integer, which corresponds to the number of molecules of type k converted (negative sign) or produced (positive sign), when reaction i is performed once. Molecules that are not involved in reaction i have a zero entry. We will denote the number of molecules in compartment II by: $\{n_0^{\mathrm{II}}, n_1^{\mathrm{II}}, n_2^{\mathrm{II}}, .., n_k^{\mathrm{II}}, .., n_{s^{\mathrm{II}}}^{\mathrm{II}}\}$. During our quasistatic process, the exchange process equilibrates

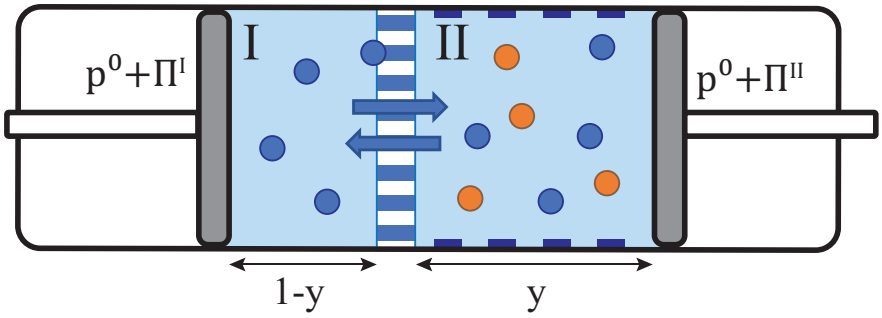


Figure 4.7: Setup for a liquid-phase racemization engine. Applied pressures are indicated next to the corresponding pistons.

permeable species $k = \{0,..,s^{\mathrm{I}}\}$ between compartments, the chemistry equilibrates species within compartment II:

$$\mu_k^{\mathrm{I}} = \mu_k^{\mathrm{II}}, \qquad 0 \le k \le s^{\mathrm{I}}, \tag{4.62}$$

$$\sum_k \nu_{ik}\mu_k^{\mathrm{II}} = 0, \qquad \forall i. \tag{4.63}$$

We let the pressures $\Pi^{\mathrm{I}}, \Pi^{\mathrm{II}}$ (seef Fig. 4.7) be equal to the osmotic pressures of ideal mixtures in compartment I and II, respectively

$$\Pi^{\mathrm{I}} = \frac{k_b T}{v_0} \ln\left(1 - \sum_{k\neq 0} x_k^{\mathrm{I}}\right) = \frac{k_b T}{v_0} \ln(x_0^{\mathrm{I}}), \tag{4.64}$$

$$\Pi^{\mathrm{II}} = \frac{k_b T}{v_0} \ln\left(1 - \sum_{k\neq 0} x_k^{\mathrm{II}}\right) = \frac{k_b T}{v_0} \ln(x_0^{\mathrm{II}}). \tag{4.65}$$

Where v_0 is the volume of a solvent molecule. The protocol consists of starting with a filled compartment I, which occupies a volume V_0, and an empty compartment II. By applying an infinitesimal overpressure dp on top of $p^\circ + \Pi^{\mathrm{I}}$, we displace a very small volume dV from I to II, after which we adjust Π^{I} and Π^{II} to the new compositions according to

$$\Pi^{\mathrm{II}}(V^{\mathrm{II}} + dV) = \Pi^{\mathrm{II}}(V^{\mathrm{II}}) + \left(\frac{d\Pi^{\mathrm{II}}}{dV}\right)_T dV = \Pi^{\mathrm{II}}(V^{\mathrm{II}}) + \frac{k_b T}{v_s}\left(\frac{d\ln(x_s^{\mathrm{II}})}{dV}\right)_T dV. \tag{4.66}$$

Let $F^{\mathrm{I}}, F^{\mathrm{II}}$ be the Helmholtz free energy of compartment I and II, respectively. Their differentials are of the following form

$$dF^j = -(p^0 + \Pi^j)dV - S^j dT + \sum_k \mu_k^j dN_k^j \tag{4.67}$$

Where $j \in \{\mathrm{I}, \mathrm{II}\}$. For isothermal operation, the change in Helmholtz free energy for a small displacement dV can be written as

$$dF = \left(\frac{d(F^{\mathrm{II}} - F^{\mathrm{I}})}{dV}\right)_T dV. \tag{4.68}$$

The negative sign comes from the fact that compartment I is compressed, whereas II is expanded. The differential can be written as

$$\left(\frac{dF^{\mathrm{I}}}{dV}\right)_T = -(p^0 + \Pi^{\mathrm{I}}) + \sum_k \mu_k^{\mathrm{I}} \left(\frac{dN_k^{\mathrm{I}}}{dV}\right)_T, \tag{4.69}$$

which directly leads to

$$dF = (\Pi^{\mathrm{I}} - \Pi^{\mathrm{II}})dV - \sum_k \mu_k^{\mathrm{I}} \left(\frac{dN_k^{\mathrm{I}}}{dV}\right)_T dV + \sum_k \mu_k^{\mathrm{II}} \left(\frac{dN_k^{\mathrm{II}}}{dV}\right)_T dV. \tag{4.70}$$

We will now inspect the chemical potential terms in more detail, and demonstrate that they vanish.

Let us define $\xi_{y'}^i$ as the net number of times we performed reaction i (forward minus backward) since the beginning of the protocol ($y = 0$) up to the point $y = y'$. For a given species k, the summation $\sum_i \nu_{ik} \xi_{y'}^i$ provides the net number of species k produced by all reactions throughout this protocol.

Where $N^{\mathrm{I}}, N^{\mathrm{II}}$ are the number of particles in compartment I and II respectively. For any given species k in compartment II, we can write a mass balance of the form [accumulation] = [influx] + [production/consumption by reaction], such that

$$\left(\frac{dN_k^{\mathrm{II}}}{dV}\right)_T dV = -\left(\frac{dN_k^{\mathrm{I}}}{dV}\right)_T dV + \sum_i \nu_{ik} d\xi_y^i. \tag{4.71}$$

Which allows us to decompose the final term of Eq. (4.70)

$$\sum_k \mu_k^{\mathrm{II}} \left(\frac{dN_k^{II}}{dV}\right)_T dV = \sum_k -\mu_k^{\mathrm{II}} \left(\frac{dN_k^{\mathrm{I}}}{dV}\right)_T dV + \sum_k \mu_k^{\mathrm{II}} \sum_i \nu_{ki} d\xi_y^i. \tag{4.72}$$

We can rewrite the latter contribution to find

$$\sum_i d\xi_y^i \sum_k \nu_{ik} \mu_k^{\mathrm{II}} = 0, \tag{4.73}$$

which vanishes due to Chemical Equilibrium, as stated in Eq. (4.62). This reduces Eq. (4.72) to

$$\sum_k \mu_k^{\mathrm{II}} \left(\frac{dN_k^{\mathrm{II}}}{dV}\right)_T dV = -\sum_k \mu_k^{\mathrm{II}} \left(\frac{dN_k^{I}}{dV}\right)_T dV. \tag{4.74}$$

For the m species not in I, no exchange takes place, so

$$\left(\frac{dN_k^{\mathrm{I}}}{dV}\right)_T dV = 0, \;\; k > l, \tag{4.75}$$

while for the other species we can apply Eq. (4.62). All chemical contributions cancel and we are left with

$$dF = \left(\Pi^{\mathrm{I}}(V_0 - V) - \Pi^{\mathrm{II}}(V)\right) dV, \tag{4.76}$$

which leads to:

$$\Delta F = \int_0^{V_0} \left(\frac{dF}{dV}\right)_T dV = \int_0^{V_0} \left(\Pi^{\mathrm{I}}(V_0 - V) - \Pi^{\mathrm{II}}(V)\right) dV = W. \tag{4.77}$$

Therefore, the work corresponds to the total Helmholtz free energy change and the protocol is reversible. The maximum extractable work can also be identified with the Gibbs free energy change for the process, which is:

$$\Delta G = T\Delta I + \sum_k \mu_k^{\circ} \Delta n_k + V_0 \left(\Pi^{\mathrm{II}}(V_0) - \Pi^{\mathrm{I}}(V_0)\right) = -T\Delta S_{tot}. \tag{4.78}$$

If the total number of solute and solvent molecules remains unchanged, x_0 is unaltered. Then from Eq. (4.65) we find that the pressure term cancels. For the racemization reaction, $\Delta\mu^{\circ} = 0$, and so the only remaining term is related to the first term ΔI.

4.2.7 Final remarks

We have demonstrated that mixing entropy ('information') and energy can be harvested reversibly, in a general sense. An interesting consequence is that chemical processes could be made considerably more efficient: neither the heat production (or consumption) nor the mixing entropy intrinsic to chemistry need to induce losses, provided we use a multicompartment protocol such that $\Delta S = 0$. Of course, the fact that we wish to perform the protocol with a certain speed and in a certain (forward) direction is accompanied by its own production of entropy (A contribution that is typically not considered in classical equilibrium thermodynamics). This raises provocative questions: can one define a thermodynamic efficiency for a chemical transformation? And how efficient can chemistry realistically be made?

The setup is reminiscent of the one used in pressure-retarded osmosis [18], which is used to harvest mixing entropy generated by the salinity gradient of fresh water and sea water. Our setup, however, is focused on exploiting molecular degrees of freedom, not concentration differences. Technologically, it encounters very similar limitations: membranes and other transport barriers may often not be selective enough to efficiently extract work from information and vice versa. Other sources of dissipation, such as fluid flows, may lead to further losses.

Such limitations underline the need for an understanding in terms of nonequilibrium thermodynamics. In particular when the goal is the purification or production of a desired compound, rather than energy production, dissipative strategies become of great interest, as is showcased in Ref[19]. In biology, dissipative proofreading strategies are ubiquitous. What is especially striking about these strategies, is that they can be understood in terms of relatively simple chemical networks. Indeed, some of the major recent advances in chemical purification techniques rely on moving beyond the limits imposed by a single chemical reaction, where even a passive second reaction may already lead to dramatic improvements. In Appendix 10.3, we illustrate this point further.

Bibliography

Articles

[1] C.E. Shannon. "A Mathematical Theory of Communication". In: *Bell Syst. Tech. J.* 27.April 1928 (1948), pp. 379–423, 623–656.

[2] C.E. Shannon. "The Bandwagon". In: *IRE Trans. - Inf. Theory* 2.1 (1956), p. 3.

[5] Daan Frenkel. "Why colloidal systems can be described by statistical mechanics: Some not very original comments on the Gibbs paradox". In: *Mol. Phys.* 112.17 (2014), pp. 2325–2329.

[6] Liang Cheng, Charles Doubleday, and Ronald Breslow. "Evidence for tunneling in base-catalyzed isomerization of glyceraldehyde to dihydroxyacetone by hydride shift under formose conditions". In: *Proc. Natl. Acad. Sci. U. S. A.* 112.14 (2015), pp. 4218–4220.

[7] Michal Rivlin, Uzi Eliav, and Gil Navon. "NMR studies of the equilibria and reaction rates in aqueous solutions of formaldehyde". In: *J. Phys. Chem. B* 119.12 (2015), pp. 4479–4487.

[8] Leo Szilard. "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen". In: *Zeitschrift für Phys.* 53 (1929), pp. 11–12.

[9] Takahiro Sagawa and Masahito Ueda. "Minimal energy cost for thermodynamic information processing: measurement and information erasure". In: *Phys. Rev. Lett.* 102.25 (2009), p. 250602.

[10] S. Toyabe et al. "Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality". In: *Nat. Phys.* 6 (2010), pp. 988–992.

[11] S Ciliberto. “Experiments in Stochastic Thermodynamics : Short History and Perspectives”. In: *Phys. Rev. X* 7 (2017), p. 021051.

[13] C. Viedma. “Chiral Symmetry Breaking During Crystallization: Complete Chiral Purity Induced by Nonlinear Autocatalysis and Recycling”. In: *Phys. Rev. Lett.* 94.6 (2005), p. 065504.

[14] Rutherford Aris. “Prolegomena to the Rational Analysis of Systems of Chemical Reactions”. In: *Arch. Ration. Mech. Anal.* 19 (1965), pp. 81–98.

[15] Bruce L Clarke. “Stoichiometric network analysis”. In: *Cell Biophys.* 12 (1988), pp. 237–253.

[16] Matteo Polettini and Massimiliano Esposito. “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”. In: *J. Chem. Phys.* 141.2 (2014), pp. –.

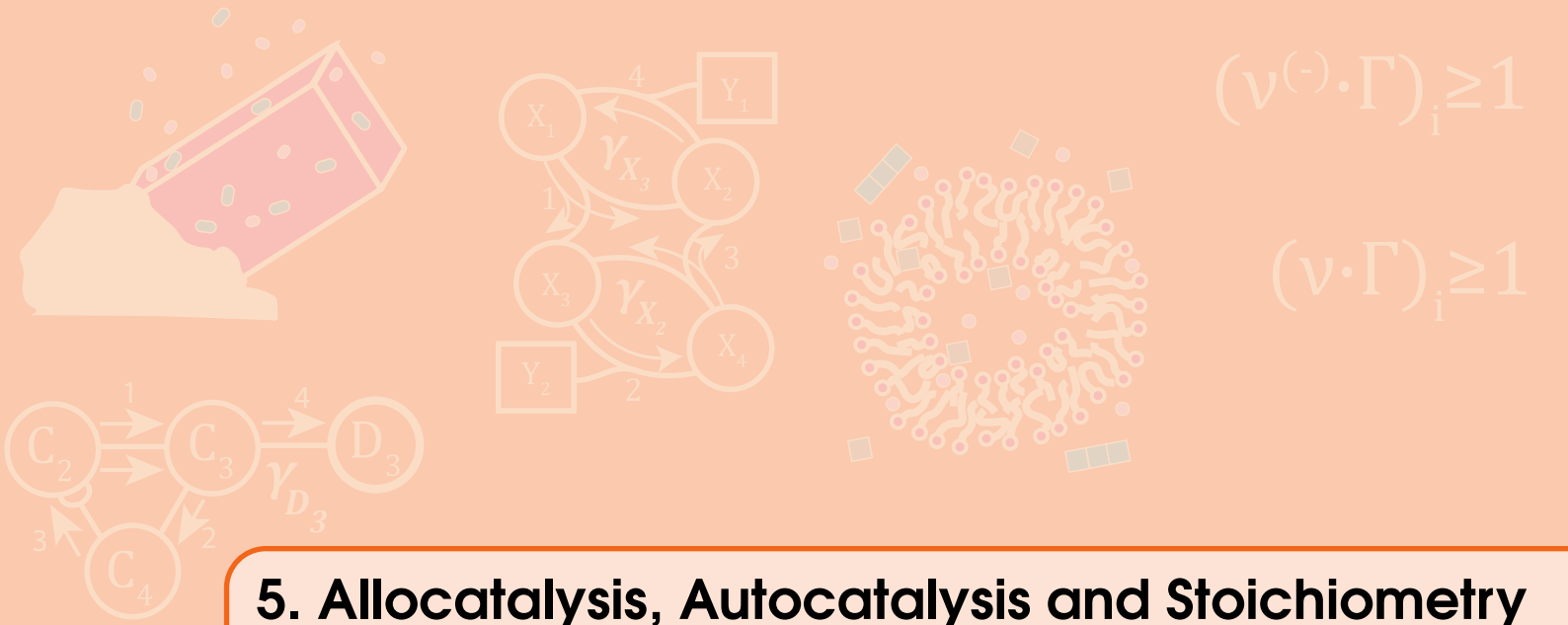
[18] Ngai Yin Yip and Menachem Elimelech. “Thermodynamic and Energy Efficiency Analysis of Power Generation from Natural Salinity Gradients by Pressure Retarded Osmosis”. In: *Environ. Sci. Technol.* 46 (2012), pp. 5230–5239.

[19] Cato Sandford, Daniel Seeto, and Alexander Y Grosberg. “Active sorting of particles as an illustration of the Gibbs mixing paradox”. In: (2018). arXiv: 1705.05537v2.

Books

[3] J. Willard Gibbs. *On the equilibrium of heterogeneous substances*. New Haven: The Academy, 1874.

[17] Martin Feinberg. *Foundations of Chemical Reaction Network Theory*. New York: Springer, 2019.



5. Allocatalysis, Autocatalysis and Stoichiometry

In this chapter, we will formally introduce the familiar concepts of catalysis and autocatalysis in their colloquial use, their strict chemical use and subsequently show their connection with elementary reactions and cycles. Our focus here will be on the structural network features, and not kinetics. For a complementary perspective on kinetic aspects of autocatalysis, see Ref. [1].

In Sec.5.1, we describe the properties of autocatalysis and catalysis in terms of elementary reactions that are unambiguous: the whole network can be inferred from the stoichiometric matrix alone. In characterizing these properties in reaction networks, we will introduce the notion of stoichiometric allocatalysis and stoichiometric autocatalysis, which we make as consistent as possible with IUPAC definitions. At the end of the chapter we provide more rigorous proof for these findings.

The framework also admits other stoichiometric processes, such as molecular exchange processes between different phases and compartments, e.g. due to partitioning, diffusion or evaporation. This gives rise to new emergent types of multicompartment autocatalysis (Sec. 5.4), which are reminiscent of ecological phenomena (e.g. mutualism, syntrophy, parasitism). In Chapter 6, we will extend these notions to chemical evolution.

In Sec. 5.6 we discuss some thermodynamic aspects of autocatalysis. Autocatalytic reactions that can be nucleated by single molecules almost surely imply a spontaneous process. Autocatalytic reactions with a threshold (bistabilities) are considerably less spontaneous.

Finally, in Sec. 5.8 we consider a case where self-replication is frustrated, due to the use of a composite chemostat.

5.1 Catalysis and Generalized Catalysis in chemistry

A short definition often employed for a catalyst, is ‘a chemical compound which accelerates a chemical reaction while not being consumed itself’. To be more precise, let us here provide the exact entry for ‘catalyst’ from the IUPAC Gold book [2]:

Definition 5.1.1 — Catalyst. "A substance that increases the rate of a reaction without modifying the overall standard Gibbs energy change in the reaction; the process is called catalysis. The catalyst is both a reactant and product of the reaction. The words catalyst and catalysis should not be used when the added substance reduces the rate of reaction (see inhibitor). Catalysis can be classified as homogeneous catalysis, in which only one phase is involved, and heterogeneous catalysis, in which the reaction occurs at or near an interface between phases. Catalysis brought about by one of the products of a reaction is called autocatalysis. Catalysis brought about by a group on a reactant molecule itself is called intramolecular catalysis. The term catalysis is also often used when the substance is consumed in the reaction (for example: base-catalysed hydrolysis of esters). Strictly, such a substance should be called an activator."

The distinction between catalysts that produce themselves and those that do not will turn out to be essential. We will distinguish between them by introducing the following terms

Definition 5.1.2 — Autocatalyst. A catalyst that performs autocatalysis, i.e. it is a catalyst that can be produced in excess in an autocatalytic cycle.

Definition 5.1.3 — Allocatalyst. A catalyst that does not perform autocatalysis (nor 'reverse autocatalysis'), performing an allocatalytic cycle does not lead to an overall change in the corresponding allocatalyst population.

5.1.1 Typical allocatalysis

If there is a single catalyst E, substrate S and product P, standard shorthand notations for the catalyzed reaction are:

$$\mathrm{S}+\mathrm{E} \rightleftharpoons \mathrm{E}+\mathrm{P}, \tag{5.1}$$

$$\mathrm{S} \overset{\mathrm{E}}{\rightleftharpoons} \mathrm{P}. \tag{5.2}$$

Of these two, only Eq. (5.1) makes explicit that E is a product and a reactant, which is required. Both of these shorthand reactions are ambiguous on the level of the stoichiometric matrix, since Eq. (2.15) (nonambiguity condition) is not satisfied. In order to remove the ambiguity, we turn to a more detailed description with a catalytic cycle described by steps 1 to 3 and, to compare, an uncatalyzed pathway given by the fourth reaction.

$$\begin{array}{c} \\ \mathrm{S} \\ \mathrm{P} \\ \boldsymbol{\nu} = \mathrm{E} \\ \mathrm{ES} \\ \mathrm{EP} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \end{array} \begin{array}{l} \\ \mathrm{S}+\mathrm{E} \overset{1}{\rightleftharpoons} \mathrm{ES} \\ \mathrm{ES} \overset{2}{\rightleftharpoons} \mathrm{EP} \\ \mathrm{EP} \overset{3}{\rightleftharpoons} \mathrm{E}+\mathrm{P} \\ \mathrm{S} \overset{4}{\rightleftharpoons} \mathrm{P} \end{array} \tag{5.3}$$

This stoichiometric matrix verifies nonambiguity (Eq. (2.15)).

A convenient reduced description can be obtained by removing the chemical species S and P from $\boldsymbol{\nu}$ to obtain $\boldsymbol{\nu}_X$:

$$\begin{array}{c} \\ \mathrm{E} \\ \boldsymbol{\nu}_X = \mathrm{ES} \\ \mathrm{EP} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \end{array}. \tag{5.4}$$

In the present case, this leads to the emergent cycle $\boldsymbol{c}^* = (1,1,1,0)^T$ as defined in Eq. (2.116). This cycle corresponds to the following net reaction

$$\mathrm{S}+\mathrm{E}+\mathrm{ES}+\mathrm{EP} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{E}+\mathrm{ES}+\mathrm{EP}+\mathrm{P}, \tag{5.5}$$

which is built from the stoichiometric matrix $\boldsymbol{\nu}$.

In terms of the stoichiometric matrix $\boldsymbol{\nu}_X$, (resp. $\boldsymbol{\nu}_Y$), we obtain the net reactions

$$\mathrm{E}+\mathrm{ES}+\mathrm{EP} \quad \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{E}+\mathrm{ES}+\mathrm{EP}, \tag{5.6}$$

$$\mathrm{S} \quad \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{P}. \tag{5.7}$$

We can now distinguish between the elementary reaction 4 in Eq. (5.4) and the nonelementary net reaction in Eq. (5.7). For an arbitrary catalytic cycle, the net reaction should contain catalytic species (E, ES, EP in our example) with the same coefficients on the reactant and product side.

Catalysis, inhibition and temperature cycling

In treating the stereotypical case of allocatalysis in (5.3), we implicitly suppose, by the use of the word catalysis, that the (catalytic) cycle $\boldsymbol{c}^*$ leads to a rate enhancement for $\mathrm{S} \rightleftharpoons \mathrm{P}$ with respect to a system performing only the uncatalyzed reaction.

In general, this is not true for any choice of rate constants in (5.3). For example, if $k_2^+ < k_4^+$ or $k_3^+ < k_4^+$, with the index denoting the reaction number in (5.3), the species E will slow down the reaction. If $k_2^+ < k_4^+$, this happens by trapping reagent S in a less reactive state. If $k_3^+ < k_4^+$, the product is trapped in the EP complex. The latter situation is very common in nonenzymatic template-assisted RNA chemistry: when binding to a template is too strong, detachment becomes a rate-limiting step.

To escape more rapidly from such a bound state, the complex EP can be heated (For nucleotide polymers we wish to increase the temperature above the 'melting temperature' [3]). By applying temperature-cycling protocols with a cycling time shorter than $1/k_1^+$, a net increase in the reaction rate can be achieved. It is not clear from definition 5.1.1. whether we can extend the notion of a catalyst to such active protocols, but doing so may be problematic*.

We see that the notion of 'catalysis' is somewhat subtle: there are ways to accelerate reactions that we may not call catalysis and a species that is a catalyst for one set of conditions, may be an inhibitor in another. We will now move our discussion to another subtle situation: chain reactions.

Chain reactions

A chain reaction is a reaction that maintains itself, by consuming and resupplying its active compounds. The first active compounds of such a reaction need to be supplied through an initiation reaction, such as

$$\mathrm{Br}_2 \overset{1}{\rightleftharpoons} 2\mathrm{Br}^\bullet, \tag{5.8}$$

for (2.125), where such an initiation can be stimulated by radiation.

In that same example, we converted $\mathrm{Br}_2 + \mathrm{H}_2$ to 2HBr

$$\mathrm{Br}^\bullet + \mathrm{H}_2 \overset{2}{\rightleftharpoons} \mathrm{HBr} + \mathrm{H}^\bullet, \quad \mathrm{H}^\bullet + \mathrm{Br}_2 \overset{3}{\rightleftharpoons} \mathrm{HBr} + \mathrm{Br}^\bullet. \tag{5.9}$$

For which the overall reaction becomes

$$\mathrm{Br}_2 + \mathrm{H}_2 + \mathrm{Br}^\bullet + \mathrm{H}^\bullet \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} 2\mathrm{HBr} + \mathrm{Br}^\bullet + \mathrm{H}^\bullet. \tag{5.10}$$

By following the atoms, it is clear that the 'active species' is replenished, but we do not recover the original 'active species'. We are often assuming that at the end of a catalytic cycle, we are

*EP does not change the overall $\Delta^\circ G$, but changing the temperature does. 5.1.1 only discusses the action of the catalyst on $\Delta^\circ G$. A temperature-cycling protocol can convert more S molecules to P than a constant-temperature equilibration, due to kinetic trapping in either the EP complex or dissociated species. This pumping action is distinct from the passive accelerating a reaction $\mathrm{S} \rightleftharpoons \mathrm{P}$, which will halt at equilibrium.

recovering the original catalyst species we started with. The IUPAC definition makes no mention of this property, it only requires the catalyst to be a reactant and a product. In this sense, chain reactions are a form of catalysis (A point well-understood in e.g. atmospheric chemistry). In our example, $Br^\bullet$, $H^\bullet$ are allocatalysts.

Topologically, this is more intuitive: by looking at a hypergraph or stoichiometric matrix, we can not tell whether we are recovering the original catalyst or making a new one (for this, atom mapping is typically required, e.g. SMILES notation[4]).

Let us now finally move to a final subtle example of catalysis: a composition of a forward and a backward autocatalytic reaction.

Interlinked autocatalytic reactions

The following minimal example obeys the definition of a catalyst given above, but that a chemist might reluctantly name as such. Starting with a catalyst E, a substrate S is, in the first two steps, converted to another copy of E. Subsequently, E forms a dimer, after which a product P and a single E is released:

$$\mathrm{S}+\mathrm{E} \overset{1}{\rightleftharpoons} \mathrm{ES} \tag{5.11}$$

$$\mathrm{ES} \overset{2}{\rightleftharpoons} 2\mathrm{E} \tag{5.12}$$

$$2\mathrm{E} \overset{3}{\rightleftharpoons} \mathrm{E}_2 \tag{5.13}$$

$$\mathrm{E}_2 \overset{4}{\rightleftharpoons} \mathrm{E}+\mathrm{P} \tag{5.14}$$

We can write the stoichiometric matrix $\boldsymbol{\nu}$ for this system

$$\begin{array}{c} \\ \mathrm{S} \\ \mathrm{P} \\ \boldsymbol{\nu} = \mathrm{E} \\ \mathrm{ES} \\ \mathrm{E}_2 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \mathrm{S}+\mathrm{E} \overset{1}{\rightleftharpoons} \mathrm{ES} \\ \mathrm{ES} \overset{2}{\rightleftharpoons} 2\mathrm{E} \\ 2\mathrm{E} \overset{3}{\rightleftharpoons} \mathrm{E}_2 \\ \mathrm{E}_2 \overset{4}{\rightleftharpoons} \mathrm{E}+\mathrm{P} \end{array} \tag{5.15}$$

For which we can now take a submatrix $\boldsymbol{\nu}^*$ by removing substrate S and product P

$$\begin{array}{c} \\ \mathrm{E} \\ \boldsymbol{\nu}^* = \mathrm{ES} \\ \mathrm{E}_2 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \mathrm{E} \overset{1}{\rightleftharpoons} \mathrm{ES} \\ \mathrm{ES} \overset{2}{\rightleftharpoons} 2\mathrm{E} \\ 2\mathrm{E} \overset{3}{\rightleftharpoons} \mathrm{E}_2 \\ \mathrm{E}_2 \overset{4}{\rightleftharpoons} \mathrm{E} \end{array} \tag{5.16}$$

The submatrix $\boldsymbol{\nu}^*$ which admits a right nullvector $\boldsymbol{c}^* = (1,1,1,1)^T$, corresponding to the net reaction

$$\mathrm{S} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \mathrm{P}. \tag{5.17}$$

Note that reactions like 1 and 4 allow to arbitrarily increase or reduce the catalyst population. What is striking is that we cannot write a catalytic cycle for this system without such reactions. This has the following cause: although we use E to convert S to P, E is not an allocatalyst.

As will be shown with more rigor in upcoming sections, this system is composed of two autocatalytic reactions (1+2 and 3+4), which admit no mass-like conservation laws ($L^{(+)}$) for catalytic species (here, E, E_2 and ES). The lack of such a conservation law allows autocatalysts to accumulate or degrade through their own catalysis.

Allocatalysts on the other hand, are characterized by an unchanging abundance. The linear constraint that forbids such accumulation is exactly the mass-like conservation law, which in example Eq. (5.3) is the sum

$$L^{(+)} = n_{\mathrm{E}} + n_{\mathrm{ES}} + n_{\mathrm{EP}}, \tag{5.18}$$

corresponding to a left nullvector $l^{(+)} = (0,0,1,1,1)$.

The cycle given by 5.11-5.14 is an example of an allocatalytic cycle, which is mediated by autocatalysts.

We will now turn to the implementation of catalysis in the stoichiometric matrix framework.

5.1.2 Autonomy and Siphons

Let us start by introducing an important property, which we will call autonomy:

Definition 5.1.4 — Autonomy. In a reversible chemical network, a stoichiometric (sub)matrix $\boldsymbol{\nu}$ is autonomous, if it contains no $\emptyset$-reactions, which are of the form

$$\sum_i \nu_{i,j} X_i \rightleftharpoons \emptyset \tag{5.19}$$

This property can easily be verified in a stoichiometric matrix $\boldsymbol{\nu}$: $\boldsymbol{\nu}_*$ has no zero columns in $\boldsymbol{\nu}_*^{(+)}$ and $\boldsymbol{\nu}_*^{(-)}$. If nonambiguity is respected ($\forall k,i\ \nu_{ki}^{(-)}\nu_{ki}^{(+)} = 0$, Eq. (2.15)), it also means that columns in $\boldsymbol{\nu}$ must have positive and negative entries.

A fully described reaction network $\boldsymbol{\nu}$ in a closed system conserves elements and isotopes and thus cannot have $\emptyset$-reactions. Such a network is always autonomous. $\emptyset$-reactions arise from a description of a subsystem that exchanges matter, either with an outside environment (e.g. external chemostats or CSTR fluxes), or by forming chemical species not described by $\boldsymbol{\nu}$ (e.g. homogeneous chemostats, buffers, species not taken in consideration for the description).

The concept of autonomy is closely related to the concept a siphon in Chemical reaction networks (CRNs).

Definition 5.1.5 — Siphon. A Siphon Σ is a subset $\Sigma \subset \mathscr{S}$ of all species S, which for each reaction that has a species in Σ as a product, has at least one of its reactants in Σ[5]

In this definition, a reaction can be irreversible in a mathematical sense: the reverse reaction does not exist. For a reversible CRN, a reverse reaction does exist, and the siphon definition must apply to the forward and backward direction. A reaction must then imply both one or more products and one or more reactants from Σ (siphon reactions), or none at all (external reactions).

Consider a reversible chemical system, with a particular subset of interest $\Sigma \subset \mathscr{S}$ which is a siphon. Letting $x_\Sigma \subset [s]$ be the species indices of siphon species and $r_\Sigma \subset [r]$ the reaction indices of siphon reactions, the submatrix $\boldsymbol{\nu}[x_\Sigma|r_\Sigma]$ contains only the siphon Σ with its siphon reactions, which thus excludes $\emptyset$-reactions. $\boldsymbol{\nu}[x_\Sigma|r_\Sigma]$ is therefore autonomous. The siphon property was proven to be a defining feature of various interpretations of catalysis and autocatalysis [5, 6, 7]. While these definitions differ from our definitions and those maintained by IUPAC, we find that this property is a robust feature, whose analogue is autonomy in our framework.

There are some intuitive arguments for why autonomy is such a central property. The capacity to perform allo- and autocatalysis is conditional on the presence of our species of interest. We cannot catalyze a reaction pathway if the corresponding allocatalyst is absent. We cannot perform autocatalysis if there is no autocatalyst to begin with. An $\emptyset$-reaction like $\emptyset \to \mathrm{A}$ takes place regardless of our system contents. Within our framework, we can treat multiple compartments and environments, and the symbol $\emptyset$ is reserved for unaltered components that are left out of the description. Since no feedback operates on such components, their absence is a natural tennet of autonomy.

In the next section, we will introduce the concept of 'stoichiometric allocatalysis', which uses as its input only the structure of a nonambiguous reaction network (stoichiometry). The concept bundles various nuances that one may or may not wish to call catalysis under a common name (hence our insistence on consistency with the IUPAC definition). The interest of doing so is that the stoichiometric framework treats these different nuances as equivalent. In addition, for most of our purposes (in particular, prebiotic scenarios) we should avoid limiting our scope without proper justification.

In a number of practical situations, however, we do have a proper justification: we may know our system rather well and wish to model some of the fluxes. Our requirements for stoichiometric allocatalysis requires a rather detailed description, which is the price we have to pay for the nonambiguity condition. For many purposes this degree of detail is unwanted. In Ref. [8] a thermodynamically consistent coarse-graining procedure is outlined that complements our approach, allowing to make the step back to a pragmatic level of description for modelling.

5.1.3 Stoichiometric Allocatalysis

Let us now formalize what has been discussed in the former sections, by introducing a concept we will call 'stoichiometric allocatalysis'. In short, a stoichiometric allocatalyst is a substance that affords a pathway to convert other substances, without the net formation or consumption of itself. Let us now provide a slightly more detailed definition:

Definition 5.1.6 — Stoichiometric Allocatalysts. A set of substances $\{\mathrm{X_k}\}$ that afford a transformation of other species, without the net consumption or production these substances. The transformation occurs through a cycle with reaction vector $\boldsymbol{c}^*$, that leads to net conversion of external Y-species ($\forall\, i,j\ \ \mathrm{X_i} \neq \mathrm{Y_j}$). This leads to an overall reaction

$$\sum_k \bar{n}_k^{(+)}\mathrm{X_k} + \sum_l n_l^{(+)}\mathrm{Y_l} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k \bar{n}_k^{(-)}\mathrm{X_k} + \sum_l n_l^{(-)}\mathrm{Y_l}, \tag{5.20}$$

Constrained by

$$\sum_k n_k^{(+)}\mathrm{Y_k} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k n_k^{(-)}\mathrm{Y_k}, \quad n_j^{(+)} \neq n_j^{(-)}, \tag{5.21}$$

$$\sum_k \bar{n}_k^{(+)}\mathrm{X_k} \underset{-\boldsymbol{c}^*}{\overset{\boldsymbol{c}^*}{\rightleftharpoons}} \sum_k \bar{n}_k^{(-)}\mathrm{X_k}, \quad \bar{n}_j^{(+)} = \bar{n}_j^{(-)}. \tag{5.22}$$

The reactions used in $\boldsymbol{c}^*$ respect a mass-like conservation law L^* for the allocatalyst population:

$$L^* = \sum_k a_k \mathrm{X_k}, \ \ \forall k\ a_k \geq 1. \tag{5.23}$$

Due to the mass-like conservation law L^*, no composition of reactions used in $\boldsymbol{c}^*$ allow for the net degradation or accumulation of allocatalysts.

Note that a species is either internal or external: $\forall\, i,j\ \ \mathrm{X_i} \neq \mathrm{Y_j}$. This clear separation between internal catalytic species and external reactants and products is stressed by the 'allo' (other) prefix, and it is this feature that separates allocatalysis from autocatalysis.

As pointed out before, if reactions in the catalytic cycle do not permit to accumulate or degrade more of the species in the ensemble of catalytic internal species, the concentration space is constrained by a conservation law L^*. We have chosen to use this as a defining feature of stoichiometric allocatalysis, thus making an explicit choice to exclude compositions of autocatalytic reactions.

Formally, we can bundle these requirements to describe a submatrix containing only allocatalysts and a set of reactions that compose one catalytic cycle $\boldsymbol{c}^*$. We will refer to such a submatrix as

an allocatalytic submatrix

Definition 5.1.7 — Allocatalytic submatrix. A submatrix $\boldsymbol{\nu}^* = \boldsymbol{\nu}([x_{\boldsymbol{c}^*}|r_{\boldsymbol{c}^*}])$ that is i) autonomous, ii) admits an emergent cycle $\boldsymbol{c}^*$ involving all species and reactions in $\boldsymbol{\nu}^*$, and iii) admits a mass-like conservation law L^* containing all internal species, is an allocatalytic submatrix. $\boldsymbol{c}^*$ is an allocatalytic cycle, and all species $x_{\boldsymbol{c}^*}$ are catalysts.

In our examples, such as Eq. (5.4), an allocatalytic submatrix was obtained by removing substrates and products from the description.

Note that we have lifted an important requirement for catalysis in our definition: acceleration is no longer required, only a new path for the reaction. This is a pragmatic choice: to know if a reaction network performs stoichiometric allocatalysis, a nonambiguous $\boldsymbol{\nu}$ provides all necessary information. Whether it accelerates the reaction under particular reaction conditions and whether it is a chain-reaction or not requires more detailed information. Stoichiometric allocatalysis is a purely stoichiometric criterion.

In practice, it is specified when a system is autocatalytic. When one speaks of catalysis it concerns allocatalysis. We have adopted this same practice throughout this manuscript: if a process is referred to as catalysis, it is implied that it concerns generalized allocatalysis unless clearly stated otherwise.

5.1.4 Example: catalysis of intercompartment exchange

Amino acids bear polar and charged endgroups and have great difficulty traversing phospholipid membranes. Stillwell [9] added an aldehyde (RCO), with which the amino group (RNH_2) was thought to form an imine (RNCR) through the reaction

$$RNH_2 + RCO \rightleftharpoons RNCR + H_2O, \tag{5.24}$$

which traverses the membrane more easily. A model for such a is given by the stoichiometric matrix $\boldsymbol{\nu}$

$$\boldsymbol{\nu} = \begin{array}{c} \\ RCO^{I} \\ RNH_2{}^{I} \\ RNCR^{I} \\ H_2O^{I} \\ RCO^{II} \\ RNH_2{}^{II} \\ RNCR^{II} \\ H_2O^{II} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \tag{5.25}$$

When we remove $RNH_2{}^{I}$ and $RNH_2{}^{II}$, the resulting matrix $\tilde{\boldsymbol{\nu}} = \boldsymbol{\nu}(2,6|\emptyset)$ is autonomous (Fig. 5.1), revealing an emergent cycle $\boldsymbol{c}_1^* = (1,1,1,1,1)^T$. From Eq. (2.54) we find the corresponding net transport reaction

$$RNH_2{}^{I} \underset{-\boldsymbol{c}_1^*}{\overset{\boldsymbol{c}_1^*}{\rightleftharpoons}} RNH_2{}^{II}. \tag{5.26}$$

We find two conservation laws, which we write as a differential conservation law L_1 and a mass-like conservation law L_2

$$L_1 = n_{H_2O^{I}} - n_{RNCR^{I}} + n_{H_2O^{II}} - n_{RNCR^{II}}, \tag{5.27}$$

$$L_2 = n_{RCO^{I}} + n_{RNCR^{I}} + n_{RCO^{II}} + n_{RNCR^{II}}. \tag{5.28}$$

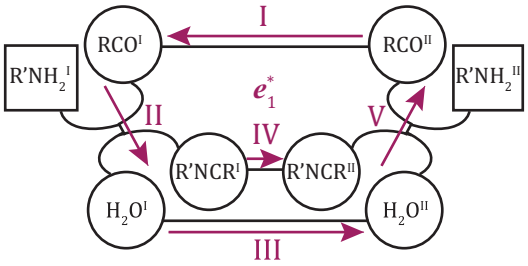


Figure 5.1: Hypergraph for Stillwell's experiment, reaction labels correspond to order of appearance in $\boldsymbol{\nu}$. A forward catalytic cycle $\boldsymbol{c}_1^*$ corresponds to a net transport reaction $\mathrm{RNH_2}^{\mathrm{I}} \rightarrow \mathrm{RNH_2}^{\mathrm{II}}$. Circular nodes correspond to catalysts, square nodes to external species.

L_1 corresponds to the fact that $\mathrm{H_2O}$ and RNCR are formed and consumed in pairs. L_2 corresponds to RCO being a catalyst preserved in the cycle. Note that we could also have chosen a mass-like conservation law for all internal species. Introducing

$$L^* = 2n_{\mathrm{RCO^I}} + n_{\mathrm{RNCR^I}} + n_{\mathrm{H_2O^I}} + 2n_{\mathrm{RCO^{II}}} + n_{\mathrm{RNCR^{II}}} + n_{\mathrm{H_2O^{II}}}.$$

We find that $L^* = 2L_2 - L_1$.

Note that $\bar{\boldsymbol{\nu}}$ is autonomous, has an emergent cycle ($\boldsymbol{c}_1^*$) that uses all reactions and internal species, and a mass-like conservation law L^* containing all internal species. It is therefore an allocatalytic submatrix.

In addition, we can also remove water ($\mathrm{H_2O^I}, \mathrm{H_2O^{II}}$), and its associated reaction 3, to obtain $\boldsymbol{\nu}^* = \bar{\boldsymbol{\nu}}(3,6|3)$

$$\boldsymbol{\nu}^* = \begin{array}{c} \\ \mathrm{RCO^I} \\ \mathrm{RNCR^I} \\ \mathrm{RCO^{II}} \\ \mathrm{RNCR^{II}} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \end{array} \tag{5.29}$$

We have $\ell = 1, c = 1$. The remaining conservation law corresponds to L_2 in Eq. (5.28), the cycle is a new emergent cycle $\boldsymbol{c}_2^* = (1,1,1,1)^T$, with Eq. (2.54) yielding the net reaction

$$\mathrm{H_2NR^I} + \mathrm{H_2O^{II}} \underset{-\boldsymbol{c}_2^*}{\overset{\boldsymbol{c}_2^*}{\rightleftharpoons}} \mathrm{H_2NR^{II}} + \mathrm{H_2O^I}. \tag{5.30}$$

The aldehyde then functions as an antiporter. Like $\bar{\boldsymbol{\nu}}$, $\boldsymbol{\nu}^*$ is an allocatalytic submatrix, but for corresponding to a different catalyzed reaction ((5.30)).

If we now remove $\mathrm{RCO^I}$, the resulting subnetwork is no longer autonomous, as reactions I and II become $\emptyset$-reactions

$$\mathrm{RCO^{II}} \rightleftharpoons \emptyset, \qquad \emptyset \rightleftharpoons \mathrm{RNCR^I}. \tag{5.31}$$

This system has no more conservation laws. According to Eq. (2.54) the cycle $\boldsymbol{c}_2^* = (1,1,1,1)^T$ now verifies

$$\mathrm{H_2NR^I} + \mathrm{H_2O^{II}} + \mathrm{RCO^I} \underset{-\boldsymbol{c}_2^*}{\overset{\boldsymbol{c}_2^*}{\rightleftharpoons}} \mathrm{RCO^I} + \mathrm{H_2NR^{II}} + \mathrm{H_2O^I} \tag{5.32}$$

which highlights that $\mathrm{RCO^I}$ functions as a catalyst in the pathway. The same applies to all other species in $\bar{\boldsymbol{\nu}}^*$, e.g. if we also remove $\mathrm{RNCR^{II}}$ we have

$$\mathrm{H_2NR^I} + \mathrm{H_2O^{II}} + \mathrm{RCO^I} + \mathrm{RNCR^{II}} \underset{-\boldsymbol{c}_2^*}{\overset{\boldsymbol{c}_2^*}{\rightleftharpoons}} \mathrm{RCO^I} + \mathrm{RNCR^{II}} + \mathrm{H_2NR^{II}} + \mathrm{H_2O^I}. \tag{5.33}$$

5.2 Stoichiometric Autocatalysis in chemistry

In chemistry, an autocatalytic reaction refers to a net reaction that forms one or more product(s) that accelerate the reaction. In the IUPAC definition of catalysis, autocatalysis is specified as: "*Catalysis brought about by one of the products of a reaction is called autocatalysis.*" To explain what autocatalysis is, one often uses the most minimal example, which is a cooperative isomerization reaction of the form

$$\mathrm{A}+\mathrm{B} \rightleftharpoons 2\mathrm{B}. \tag{5.34}$$

In Sec. 2.2.2 we showed that this example can be rewritten in a form that verifies nonambiguity, by considering two dissolved species that initially approach each other, forming an 'encounter complex' [AB], with the encounter complex subsequently undergoing the reaction:

$$\mathrm{A}+\mathrm{B} \rightleftharpoons [\mathrm{AB}], \tag{5.35}$$

$$[\mathrm{AB}] \rightleftharpoons 2\mathrm{B}. \tag{5.36}$$

More generally, we can write a reaction balance for an autocatalytic reaction vector $\mathbf{\Gamma}$, which is of the form

$$Food + n_{\mathrm{A}}\mathrm{A} + n_{\mathrm{B}}\mathrm{B} + \ldots \underset{-\mathbf{\Gamma}}{\overset{+\mathbf{\Gamma}}{\rightleftharpoons}} Waste + m_{\mathrm{A}}\mathrm{A} + m_{\mathrm{B}}\mathrm{B} + \ldots\,, \quad m_{\mathrm{X_i}} > n_{\mathrm{X_i}} \geq 1, \tag{5.37}$$

where 'Food' and 'Waste' are unspecified other compounds that are consumed or produced. For autocatalysts, stoichiometric coefficients are given. In reaction (5.34), A is the 'Food', and there is no waste.

For a species to be considered an autocatalyst: it must participate in the reaction ($n_{\mathrm{X_i}} \geq 1$) and yield more copies of itself ($m_{\mathrm{X_i}} > n_{\mathrm{X_i}}$). We can be a bit more precise, by stipulating that there may also be allocatalysts $\{E_k\}$ involved in the overall reaction[†], leading to a reaction balance of the form

$$Food + \sum_k n_k^{(+)} X_k + \sum_j \bar{n}_j^{(+)} E_j \underset{-\mathbf{\Gamma}}{\overset{+\mathbf{\Gamma}}{\rightleftharpoons}} Waste + \sum_k n_k^{(-)} X_k + \sum_j \bar{n}_j^{(-)} E_j, \quad \bar{n}_k^{(+)} = \bar{n}_k^{(-)}, \quad n_k^{(+)} > n_k^{(-)} \geq 1. \tag{5.38}$$

In our upcoming discussions, such involvement of allocatalysts is often left implicit, since our methodology will involve removing species that are not autocatalysts from the description. In doing so, allocatalysts are treated on the same footing as food and waste.

5.3 Self-Replication and autocatalysis

In this section we provide stoichiometric criteria for a network to display self-replication and autocatalysis. Whether such behavior will manifest itself depends on kinetic parameters and concentrations, which will be treated in Ch.6.

5.3.1 Stoichiometrically Feasible Reproduction (SFR)

Let us start with the property of self-reproduction, which loosely means: a set of species reproduce themself, and possibly some other species. A more technical, stoichiometric property that we will be using quite often is the following:

[†] An experimentally important example is Mg^{2+}, which is used in many RNA networks and the formose reaction

Definition 5.3.1 — Stoichiometrically Feasible Reproduction (SFR). A stoichiometric submatrix $\bar{\boldsymbol{\nu}}$ admits Stoichiometrically Feasible Reproduction (SFR), if and only if i) $\bar{\boldsymbol{\nu}}$ is autonomous, ii) single-block and iii) there exists a reaction vector $\boldsymbol{\Gamma} = (\Gamma_1, \Gamma_2, .., \Gamma_r) \in \mathbb{Z}^r$, such that

$$\forall i \in \{1, .., s\} \;\; \Delta n_i = (\bar{\boldsymbol{\nu}} \cdot \boldsymbol{\Gamma})_i \geq 1. \tag{5.39}$$

In other words, there is a combination of reactions that simultaneously yields at least one copy of each species in $\bar{\boldsymbol{\nu}}$ in a nontrivial manner (without $\emptyset$-reactions).

Let us consider the example of an autonomous, single-block submatrix $\bar{\boldsymbol{\nu}}$ that is invertible. We denote $\bar{\boldsymbol{\nu}}^{-1} = (\boldsymbol{g}^{(1)}, .., \boldsymbol{g}^{(s)})$, such that

$$\bar{\boldsymbol{\nu}}\bar{\boldsymbol{\nu}}^{-1} = \boldsymbol{I}, \quad \bar{\boldsymbol{\nu}} \cdot \boldsymbol{g}^{(i)} = \hat{\boldsymbol{e}}_i \tag{5.40}$$

where $\boldsymbol{I}$ is the identity matrix and $\hat{\boldsymbol{e}}_i$ is the ith unit vector. A reaction vector $\boldsymbol{g}^{(i)}$ verifies

$$\Delta n_j = (\bar{\boldsymbol{\nu}} \cdot \boldsymbol{g}^{(i)})_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{5.41}$$

with δ_{ij} the Kronecker delta symbol. $\bar{\boldsymbol{\nu}}$ is full-rank: its linearly independent replication cycles $\boldsymbol{g}^{(i)}$ span all s dimensions of concentration space.

In principle, $\boldsymbol{g}^{(i)}$ may contain fractional reactions, which leads us to define the scaled reaction vector $\bar{\boldsymbol{g}}^{(i)} = m_i \boldsymbol{g}^{(i)}$, where $m_i > 0$ is some constant such that $\bar{\boldsymbol{g}}^{(i)} \in \mathbb{Z}^r$. It follows that $\boldsymbol{\Gamma} = \sum_i \bar{\boldsymbol{g}}^{(i)}$ obeys Eq. (5.39), meaning an autonomous, single-block submatrix that is invertible is guaranteed to admit SFR.

For a noninvertible matrix that admits SFR, conservation laws are constrained: the replication vector $\boldsymbol{\Gamma}$ requires a conservation law $\boldsymbol{\ell}$ to obey

$$\boldsymbol{\ell} \cdot (\bar{\boldsymbol{\nu}} \cdot \boldsymbol{\Gamma}) = \sum_i \ell_i \Delta n_i = 0. \tag{5.42}$$

Eq. (5.42) does not admit mass-like conservation laws $\boldsymbol{\ell}^+$, since neither ℓ_i^+ nor Δn_i can have negative entries. It follows that SFR can only allow differential conservation laws $\boldsymbol{\ell}^\pm$. Since the full stoichiometric matrix $\boldsymbol{\nu}$ always has mass-like conservation laws, it cannot be SFR. From the perspective of a marginal observer seeing only $\bar{\boldsymbol{\nu}}$, an SFR breaks mass conservation. This happens by converting outside matter ('food') to copies of species within the subnetwork $\bar{\boldsymbol{\nu}}$.

5.3.2 Stoichiometrically Feasible Autocatalysis (SFA)

Definition 5.3.2 — Stoichiometrically Feasible Autocatalysis (SFA). A stoichiometric submatrix $\boldsymbol{\nu}_\diamond$ admits Stoichiometrically Feasible Autocatalysis (SFA), if, in addition to admitting SFR, it also verifies

$$\forall i \in \{1, .., s\} \quad (\boldsymbol{\nu}_\diamond^{(-)} \cdot \boldsymbol{\Gamma}_\diamond)_i \geq 1, \tag{5.43}$$

where $\boldsymbol{\nu}_\diamond^{(-)}, \boldsymbol{\Gamma}_\diamond$ are constructed such that $\forall i \; \Gamma_{\diamond i} \geq 0$.

In this definition, $\boldsymbol{\nu}_\diamond^{(-)}, \boldsymbol{\Gamma}_\diamond$ are constructed according to Eq. (2.56), such that $\forall i \; \Gamma_{\diamond i} \geq 0$.

Writing the net chemical equation for $\boldsymbol{\Gamma}_\diamond$ allows us to recover the standard chemical interpretation of autocatalysis. Limiting our scope to species in $\boldsymbol{\nu}_\diamond$, Eq. (5.38) provides a net reaction

$$\sum_i \left(\boldsymbol{\nu}_\diamond^{(-)} \cdot \boldsymbol{\Gamma}_\diamond\right)_i \mathrm{X_i} \longrightarrow \sum_i \left(\boldsymbol{\nu}_\diamond^{(+)} \cdot \boldsymbol{\Gamma}_\diamond\right)_i . \mathrm{X_i} \tag{5.44}$$

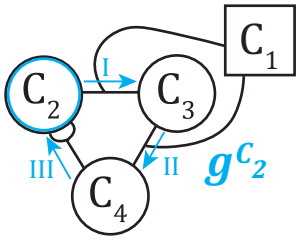


Figure 5.2: The minimal formose reaction. Upon removing C_1, we obtain replication cycles. Arrows illustrate the replication cycle $\boldsymbol{g}^{C_2}$. Circular nodes contain autocatalysts. Square nodes contain external (food) species.

In which each species increases in number (SFR)

$$\forall i \left(\boldsymbol{\nu}_\diamond^{(+)} \cdot \boldsymbol{\Gamma}_\diamond - \boldsymbol{\nu}_\diamond^{(-)} \cdot \boldsymbol{\Gamma}_\diamond\right)_i \geq 1, \tag{5.45}$$

and each species participates: $\forall i\ (\boldsymbol{\nu}_\diamond^{(-)} \cdot \boldsymbol{\Gamma}_\diamond)_i \geq 1$. The simplest example of autocatalysis according to Eq. (5.44) is the net reaction

$$\mathrm{X} \longrightarrow 2\mathrm{X}. \tag{5.46}$$

In the upcoming sections, examples will be provided to demonstrate these concepts.

5.3.3 Example: formose

As an example, we first consider the Toy Formose reaction, described by

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{C_1} \\ \mathrm{C_2} \\ \mathrm{C_3} \\ \mathrm{C_4} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & -1 & 0 \\ -1 & 0 & 2 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \mathrm{C_1} + \mathrm{C_2} \overset{1}{\rightleftharpoons} \mathrm{C_3} \\ \mathrm{C_1} + \mathrm{C_3} \overset{2}{\rightleftharpoons} \mathrm{C_4} \\ \mathrm{C_4} \overset{3}{\rightleftharpoons} 2\mathrm{C_2} \end{array} \tag{5.47}$$

Upon chemostatting C_1 (see Fig. 5.3), we obtain an autonomous submatrix $\bar{\boldsymbol{\nu}}$

$$\bar{\boldsymbol{\nu}} = \begin{array}{c} \\ \mathrm{C_2} \\ \mathrm{C_3} \\ \mathrm{C_4} \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \mathrm{C_2} \overset{1}{\rightleftharpoons} \mathrm{C_3} \\ \mathrm{C_3} \overset{2}{\rightleftharpoons} \mathrm{C_4} \\ \mathrm{C_4} \overset{3}{\rightleftharpoons} 2\mathrm{C_2} \end{array} \tag{5.48}$$

Which is single-block and admits an inverse

$$\bar{\boldsymbol{\nu}}^{-1} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 2 \end{array} \\ \begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \end{array} = (\boldsymbol{g}^{\mathrm{C_2}}, \boldsymbol{g}^{\mathrm{C_3}}, \boldsymbol{g}^{\mathrm{C_4}}). \tag{5.49}$$

The columns of $\bar{\boldsymbol{\nu}}^{-1}$ correspond to reaction vectors, whose application yields one net copy of the corresponding molecule, e.g. $\Delta n_i = (\bar{\boldsymbol{\nu}} \cdot \boldsymbol{g}^{(i)})_i = 1$. This is illustrated for C_2 in Fig. 5.3, $\bar{\boldsymbol{\nu}} \cdot \boldsymbol{g}^{\mathrm{C_2}}$.

Since all vectors in $\bar{\boldsymbol{\nu}}^{-1} = (\boldsymbol{g}^{\mathrm{C_2}}, \boldsymbol{g}^{\mathrm{C_3}}, \boldsymbol{g}^{\mathrm{C_4}})$, are positive, we can directly use expression Eq. (2.47) for $\bar{\boldsymbol{\nu}}$ to write overall reactions:

$$\begin{aligned} \boldsymbol{g}^{\mathrm{C_2}} &: \mathrm{C_2} + \mathrm{C_3} + \mathrm{C_4} \longrightarrow 2\mathrm{C_2} + \mathrm{C_3} + \mathrm{C_4} \\ \boldsymbol{g}^{\mathrm{C_3}} &: 2\mathrm{C_2} + \mathrm{C_3} + \mathrm{C_4} \longrightarrow 2\mathrm{C_2} + 2\mathrm{C_3} + \mathrm{C_4} \\ \boldsymbol{g}^{\mathrm{C_4}} &: 2\mathrm{C_2} + 2\mathrm{C_3} + \mathrm{C_4} \longrightarrow 2\mathrm{C_2} + 2\mathrm{C_3} + 2\mathrm{C_4} \end{aligned} \tag{5.50}$$

A particular feature of formose is that every species intervenes in its own replication cycle $\boldsymbol{g}^{(i)}$. In upcoming examples, we show that this is generally not the case. We can construct $\boldsymbol{\Gamma} = \boldsymbol{g}^{\mathrm{C}_2} + \boldsymbol{g}^{\mathrm{C}_3} + \boldsymbol{g}^{\mathrm{C}_4}$, which leads to the overall reaction

$$\boldsymbol{\Gamma} : 5\mathrm{C}_2 + 4\mathrm{C}_3 + 3\mathrm{C}_4 \longrightarrow 6\mathrm{C}_2 + 5\mathrm{C}_3 + 4\mathrm{C}_4. \tag{5.51}$$

Since this reaction is autocatalytic for each species, $\bar{\boldsymbol{\nu}}$ admits SFA. Furthermore, $\bar{\boldsymbol{\nu}}$ admits no smaller SFA submatrices, which means it is also an autocatalytic core.

If we remove C_2, such that $\mathbf{X} = (\mathrm{C}_3, \mathrm{C}_4)^{\mathbf{T}}$, $\mathbf{Y} = (\mathrm{C}_1, \mathrm{C}_2)^{\mathbf{T}}$, the replication cycle $\boldsymbol{g}^{\mathrm{C}_2}$ becomes an emergent cycle

$$\boldsymbol{\nu} = \begin{pmatrix} \boldsymbol{\nu}_Y \\ \boldsymbol{\nu}_X \end{pmatrix}, \quad \boldsymbol{\nu}_X \cdot \boldsymbol{g}^{\mathrm{C}_2} = \mathbf{0}. \tag{5.52}$$

Using $\boldsymbol{\nu}_Y$ in Eq. (2.54) yields a commonly used representation for this autocatalytic cycle

$$2\mathrm{C}_1 + \mathrm{C}_2 \longrightarrow 2\,\mathrm{C}_2. \tag{5.53}$$

Note that $\boldsymbol{\nu}_X$ is neither autonomous (2 $\emptyset$-reactions) nor invertible ($o = 1$).

5.3.4 Example: decorated formose

Let us now consider a decorated formose reaction, in which an extra reaction $\mathrm{C}_3 \rightleftharpoons \mathrm{D}_3$ has been introduced

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{C}_1 \\ \mathrm{C}_2 \\ \mathrm{C}_3 \\ \mathrm{D}_3 \\ \mathrm{C}_4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & -1 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \end{array}. \tag{5.54}$$

$\boldsymbol{\nu}$ now admits two autonomous and invertible submatrices: $\bar{\boldsymbol{\nu}}$ from Eq. (5.48), and $\boldsymbol{\nu}_*$, obtained by removing C_1

$$\boldsymbol{\nu}_* = \begin{array}{c} \\ \mathrm{C}_2 \\ \mathrm{C}_3 \\ \mathrm{D}_3 \\ \mathrm{C}_4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 2 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \end{array}. \tag{5.55}$$

The inverse now contains a new replication cycle

$$\boldsymbol{\nu}_*^{-1} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 2 & 2 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array} = (\boldsymbol{g}^{\mathrm{C}_2}, \boldsymbol{g}^{\mathrm{C}_3}, \boldsymbol{g}^{\mathrm{D}_3}, \boldsymbol{g}^{\mathrm{C}_4}). \tag{5.56}$$

$\boldsymbol{g}^{\mathrm{D}_3} = (2,1,1,1)^T$ is the sole replication cycle which uses the added reaction, and D_3 only ever occurs as a product

$$\boldsymbol{g}^{\mathrm{D}_3} : 2\mathrm{C}_2 + 2\mathrm{C}_3 + \mathrm{C}_4 \longrightarrow 2\mathrm{C}_2 + 2\mathrm{C}_3 + \mathrm{C}_4 + \mathrm{D}_3. \tag{5.57}$$

If we now consider $\boldsymbol{\Gamma} = \boldsymbol{g}^{\mathrm{C}_2} + \boldsymbol{g}^{\mathrm{C}_3} + \boldsymbol{g}^{\mathrm{D}_3} + \boldsymbol{g}^{\mathrm{C}_4}$, we no longer obtain an autocatalytic form

$$\boldsymbol{\Gamma} : 7\mathrm{C}_2 + 6\mathrm{C}_3 + 4\mathrm{C}_4 \longrightarrow 8\mathrm{C}_2 + 7\mathrm{C}_3 + 5\mathrm{C}_4 + \mathrm{D}_3. \tag{5.58}$$

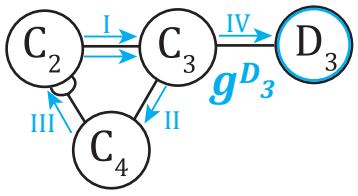


Figure 5.3: A decorated Toy Formose reaction given by the submatrix $\boldsymbol{\nu}_*$, obtained by removing C_1. The replication cycle $\boldsymbol{g}^{D_3}$ is illustrated in blue. In this network, only species C_2, C_3 and C_4 are autocatalysts.

It follows that $\bar{\boldsymbol{\nu}}_*$ admits SFR, but not SFA.

Let us now remove D_3 to turn $\boldsymbol{g}^{D_3}$ into an emergent cycle. Let us then construct ν^Y from the rows of C_1 and D_3. From Eq. (2.54), the net reaction is then

$$3C_1 \longrightarrow D_3, \tag{5.59}$$

in which D_3 is only a product, which reaffirms that $\boldsymbol{g}^{D_3}$ is not part of the autocatalytic core.

5.3.5 Example: Cross-catalytic autocatalysis

When autocatalysis is achieved through cross-catalysis or catalysis, some column vectors of $\bar{\boldsymbol{\nu}}^{-1}$ will correspond to simpler emergent cycles. The smallest autocatalytic cycles will be composite cycles. As an example, let us consider the network in Fig. 5.4. Upon removing Y_1 and Y_2, the internal network is described by

$$\bar{\boldsymbol{\nu}} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \end{array} \tag{5.60}$$

The submatrix $\bar{\boldsymbol{\nu}}$ is invertible, autonomous and single-block with

$$\bar{\boldsymbol{\nu}}^{-1} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} = (\boldsymbol{g}^{X_1}, \boldsymbol{g}^{X_2}, \boldsymbol{g}^{X_3}, \boldsymbol{g}^{X_4}), \tag{5.61}$$

Contrary to formose, these replication cycles do not involve their own species:

$$\begin{aligned} \boldsymbol{g}^{X_1} &: X_2 + X_3 + X_4 \longrightarrow X_1 + X_2 + X_3 + X_4 \\ \boldsymbol{g}^{X_2} &: X_3 + X_4 \longrightarrow X_2 + X_3 + X_4 \\ \boldsymbol{g}^{X_3} &: X_1 + X_2 \longrightarrow X_1 + X_2 + X_3 \\ \boldsymbol{g}^{X_4} &: X_1 + X_2 + X_3 \longrightarrow X_1 + X_2 + X_3 + X_4. \end{aligned} \tag{5.62}$$

If we take $\boldsymbol{\Gamma} = \boldsymbol{g}^{X_1} + \boldsymbol{g}^{X_2} + \boldsymbol{g}^{X_3} + \boldsymbol{g}^{X_4}$, Eq. (2.54) yields a collectively autocatalytic reaction

$$2X_1 + 3X_2 + 3X_3 + 2X_4 \longrightarrow 3X_1 + 4X_2 + 4X_3 + 3X_4. \tag{5.63}$$

We can construct two catalytic submatrices

$$\boldsymbol{\nu}_{12} = \begin{array}{c} \\ X_1 \\ X_2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 4 \end{array} \\ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}, \qquad \boldsymbol{\nu}_{34} = \begin{array}{c} \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cc} 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array} \tag{5.64}$$

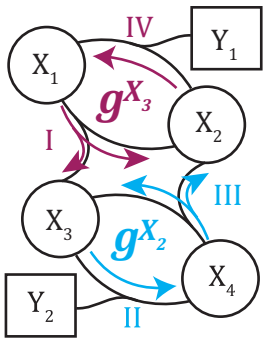


Figure 5.4: An autocatalytic network, through two crossed allocatalytic cycles. Circular nodes contain autocatalysts, square nodes contain external (food) species.

which are autonomous and admit a mass-like conservation law $\boldsymbol{\ell} = (1,1)$. The emergent cycle $\boldsymbol{e}^* = (1,1)^T$ is a catalytic cycle, composed of the same reactions as $\boldsymbol{g}^{\mathrm{X}_2}$ (resp. $\boldsymbol{g}^{\mathrm{X}_3}$) for $\boldsymbol{\nu}_{12}$ (resp. $\boldsymbol{\nu}_{34}$). Catalyzed reactions are then found to be

$$\begin{aligned} \mathrm{Y}_2 &\longrightarrow \mathrm{X}_2 \\ \mathrm{Y}_1 &\longrightarrow \mathrm{X}_3 \end{aligned} \tag{5.65}$$

Within the context of their allocatalytic submatrices $\boldsymbol{\nu}_{12}$ and $\boldsymbol{\nu}_{34}$, the internal species are allocatalysts. In the autocatalytic matrix $\bar{\boldsymbol{\nu}}$, however, they cease to be allocatalysts and are autocatalysts instead.

If we make a composition $\boldsymbol{g}^* = \boldsymbol{g}^{\mathrm{X}_2} + \boldsymbol{g}^{\mathrm{X}_3}$, both X_2 and X_3 take an autocatalytic role in the internal reaction as expressed by Eq. (2.57) for $\bar{\boldsymbol{\nu}}$

$$\boldsymbol{g}^* : \mathrm{X}_1 + \mathrm{X}_2 + \mathrm{X}_3 + \mathrm{X}_4 \longrightarrow \mathrm{X}_1 + 2\mathrm{X}_2 + 2\mathrm{X}_3 + \mathrm{X}_4 \tag{5.66}$$

If we only consider reproduced autocatalytic species and consumed reactants, we obtain a more conventional representation:

$$\boldsymbol{g}^* : \mathrm{Y}_1 + \mathrm{Y}_2 + \mathrm{X}_2 + \mathrm{X}_3 \longrightarrow 2\mathrm{X}_2 + 2\mathrm{X}_3. \tag{5.67}$$

5.4 Multicompartment autocatalysis

In practice, a system can be composed of many different environments with different local chemistries and compositions, performing selective exchange among each other. In terms of stoichiometry and reaction network topology, catalysis of a reaction or catalysis of transport obey the same criteria. Since we have now established what such criteria are for autocatalysis, we can swiftly conclude that there exists an entire new class of autocatalytic processes that rely on physical chemical processes other than pure chemistry, such as evaporation, permeation, adsorption, absorption, partitioning etc. In Fig. 5.5, some environments are depicted that lead to specific exchange. In this section, this novel type of autocatalysis will be illustrated by examples.

Transport between environments allow for new types of autocatalytic cycles. For example, the system in Fig. 5.6 can display autocatalysis, because the transport catalyst B imports its own

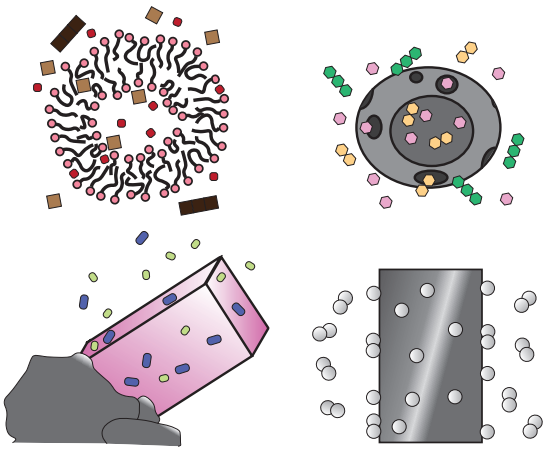


Figure 5.5: Examples of environments: i) a vesicle with a semipermeable membrane, ii) a hollow solid with small pores, iii) a mineral surface displaying selective adsorption, iv) a metal incorporating single atoms from gas molecules.

starting material BC. Its full stoichiometric matrix $\boldsymbol{\nu}$ is given by

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{BC^I} \\ \mathrm{B^I} \\ \mathrm{[BCB]^I} \\ \mathrm{D^I} \\ \mathrm{DC^I} \\ \mathrm{BC^{II}} \\ \mathrm{B^{II}} \\ \mathrm{[BCB]^{II}} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix} \end{array} \tag{5.68}$$

By removing $\mathrm{BC^{II}}$, $\mathrm{D^I}$ and $\mathrm{DC^I}$, we obtain an autonomous network, described by the invertible matrix $\tilde{\boldsymbol{\nu}}$

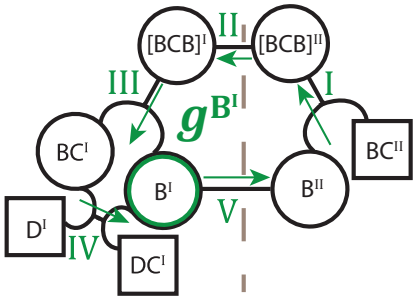


Figure 5.6: A reaction network exhibiting autocatalytic exchange. A transport barrier (dashed brown line) divides two environments. An autocatalytic submatrix is obtained by removing $\mathrm{BC^{II}}$, $\mathrm{D^I}$, $\mathrm{DC^I}$. Arrows illustrate replication cycle $\boldsymbol{g}^{\mathrm{B^I}}$.

$$\bar{\boldsymbol{\nu}} = \begin{array}{c} \\ \mathrm{BC^I} \\ \mathrm{B^I} \\ \mathrm{[BCB]^I} \\ \mathrm{B^{II}} \\ \mathrm{[BCB]^{II}} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 \end{pmatrix} \end{array}. \tag{5.69}$$

The resulting $\bar{\boldsymbol{\nu}}$ is invertible, autonomous and single-block. Since it admits no further submatrix with these properties, it is also the autocatalytic core, with replication cycles given by

$$\bar{\boldsymbol{\nu}}^{-1} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 1 & 2 & 1 & 2 \end{array} \\ \begin{pmatrix} 1 & 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 & 2 \end{pmatrix} \end{array} \tag{5.70}$$

Where $\bar{\boldsymbol{\nu}}^{-1} = (\boldsymbol{g}^{\mathrm{BC^I}}, \boldsymbol{g}^{\mathrm{B^I}}, \boldsymbol{g}^{\mathrm{[BCB]^I}}, \boldsymbol{g}^{\mathrm{B^{II}}}, \boldsymbol{g}^{\mathrm{[BCB]^{II}}})$. $\boldsymbol{g}^{\mathrm{BC^I}}$ corresponds to a catalytic cycle, the other replication cycles are autocatalytic in themselves.

5.4.1 Autocatalysis as an emergent phenomenon

Let us now solely consider the chemistry in the left compartment in Fig. 5.6, whose full matrix $\boldsymbol{\nu}_\mathrm{I}$ is given by

$$\boldsymbol{\nu}_\mathrm{I} = \begin{array}{c} \\ \mathrm{BC^I} \\ \mathrm{B^I} \\ \mathrm{[BCB]^I} \\ \mathrm{D^I} \\ \mathrm{DC^I} \end{array} \begin{array}{c} \begin{array}{cc} 3 & 4 \end{array} \\ \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \end{array} \tag{5.71}$$

A network with $r = 2$ can only admit an SFR if there is at least one stoichiometric coefficient $\nu_{ki} = 2$ (see proof in Appendix). It follows that $\boldsymbol{\nu}_\mathrm{I}$ cannot perform autocatalysis by itself.

This demonstrates that single-pot chemistry that is not autocatalytic may well become autocatalytic in the presence of exchange. This form of autocatalysis is precluded in other approaches (discussed in detail Ch. 6) such as autocatalytic sets [10, 11], where all nutrients (the 'food set') are locally chemostatted and ipso facto abundant.

As considered in Sec. 3.1 and our example of facilitated transport 5.1.4, it may be useful to extend these notions beyond the single compartment: a system may acquire its nutrients through chemically mediated exchange processes.

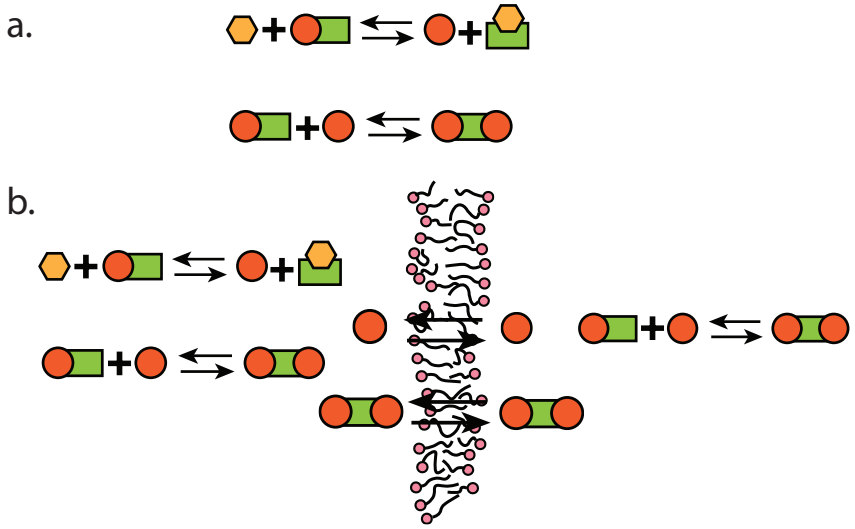


Figure 5.7: a) Illustration of the chemical network from 5.6 in a closed system. b) Illustration of the same chemical system, but with a selective exchange process through a membrane. Only the latter system can display autocatalysis.

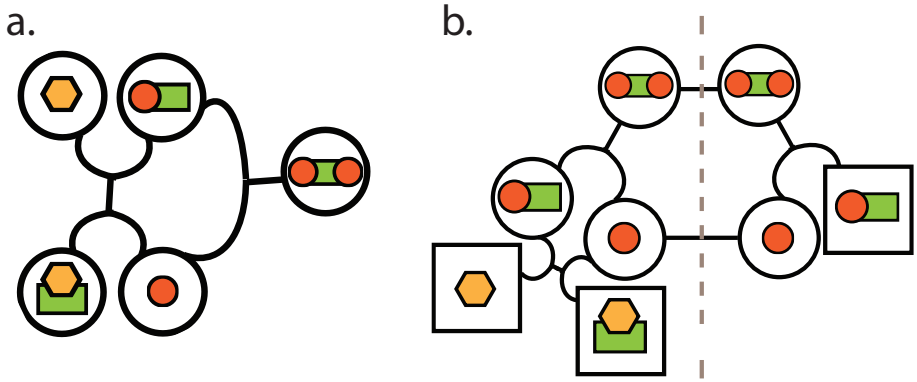


Figure 5.8: a) Graph for the chemical network from 5.6 in the closed system. b) Graph of the same chemical system, but with exchange reactions compartments (separated by dashed line).

In our example in Fig. 5.6, autocatalysis is an emergent property of chemical separation and cooperative transport. The transport barrier allows BC to play the role of chemostatted food in the form of BC^{II} and the role of a rare internal autocatalyst in the form of BC^{I}.

Topologically speaking, there is strictly more autocatalysis when a system is open to an environment, since adding more reactions and exchange processes cannot alter the fact that existing ones admit SFA or not, this is a property of a submatrix. Whether more or less autocatalysis will manifest itself is a question of kinetics.

5.4.2 Multicompartment autocatalysis and its analogues

Many environments in contact can share common autocatalytic components. E.g. we can have a pair of vesicles (see Fig. 5.5) in a solution of BC, which yields the network in Fig. 5.9. Having more compartments accelerates the transport reactions and compartment chemistries necessary for an autocatalytic cycle, improving autocatalysis overall. Fig. 5.9 can be seen as an autocatalytic

analogue of mutualism to exploit a nutrient BC. In microbial ecology, analogues modes of cooperation are found, e.g. the secretion of degradative enzymes and carrier molecules (e.g. siderophores, chalkophores) by different organisms liberates nutrients that diffuse to many different beneficiaries.

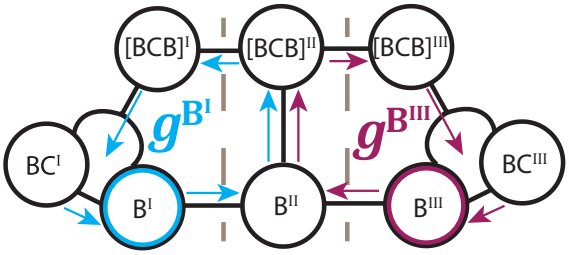


Figure 5.9: Autocatalytic transport for multiple compartments. Transport barriers (brown dotted lines) separate compartments I and III. The species $D^{I}, DC^{I}, BC^{II}, D^{III}, DC^{III}$ have been removed.

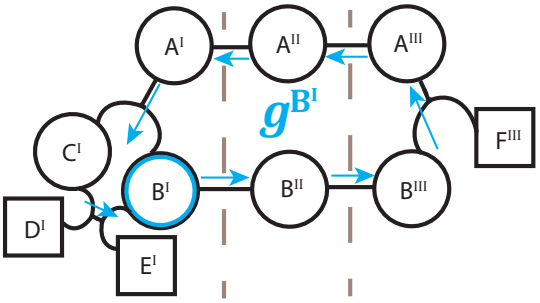


Figure 5.10: Autocatalytic transport with two different compartments I and III.

A more elaborate multicompartment cycle is given in Fig. 5.10. Here, environment I and III can e.g. be vesicles with different internal chemistries, allowing for overall autocatalysis. This is the autocatalytic analogue of (obligate) syntrophy, which in biology represents nutritional interdependence, e.g. between microbia crossfeeding essential metabolites.

5.5 Solitary and joint autocatalysis

So far, our examples have focused on autocatalytic cycles in which autocatalysts only react with species in the environment, and not with other autocatalysts. A number of interesting phenomena take place when we consider autocatalytic cycles that involve autocatalysts that react with other autocatalysts. In this section, we will introduce a formal distinction between these fundamentally different network structures, referring to the former as solitary autocatalysis and the latter as joint autocatalysis. We will then illustrate how this distinction provides an interpretation for some properties of a well-known example of joint autocatalysis.

Let us start by defining our two categories of autocatalysis

Definition 5.5.1 — Solitary Autocatalysis. A form of autocatalysis, in which all forward reaction steps involve a single autocatalyst.

A characteristic example is provided by Toy formose (Appendix 10.2), where upon removal of the environment (food) species C_1 from our description, we obtain

$$C_2 \longrightarrow C_3 \longrightarrow C_4 \longrightarrow 2C_2. \tag{5.72}$$

Every forward reaction in Eq. (5.72) involves a single autocatalyst. Note that this cannot be true in the reverse sense: the production of more autocatalysts (through autocatalysis) requires at least one reaction that produces two autocatalysts.

We are now in a position to define joint autocatalysis

Definition 5.5.2 — joint Autocatalysis. A form of autocatalysis, in which at least one forward reaction step involves more than one autocatalyst.

Using our convention that reactions are at most bimolecular, let us consider a population of autocatalysts $\{X_1,..,X_s\}$. Joint autocatalysis then means that in their autocatalytic cycle, there will be at least one forward reaction step of the form

$$X_k + X_l \rightarrow .. \tag{5.73}$$

where k can be equal to l. If the network contains one step of the form Eq. (5.73), there must be at least two reactions that produce two autocatalysts to ensure net accumulation. An example of this is seen in Fig. 5.11.

A minimal example of joint autocatalysis was given by Schlögl[12], through a termolecular reaction

$$2X + B \overset{1}{\rightleftharpoons} 3X, \tag{5.74}$$

with B a species provided by a reservoir. We will now develop a version of this reaction in terms of bimolecular reactions that verifies nonambiguity.

Nonambiguity for Schlögls termolecular reaction

Oftentimes, reaction(5.74) is referred to as a termolecular reaction, which has led some to consider it as a purely hypothetical example. As discussed in Sec. 2.2, termolecular reactions can generally be seen as a succession of smaller, bimolecular steps in which intermediate complexes are formed. In Refs. [13, 14] it was shown how the termolecular reaction could be written and treated in terms of two bimolecular reactions. This system did not verify nonambiguity, however. To make such reactions amenable to study within the framework introduced in this chapter, we will need a different decomposition.

For reaction (5.74), we can construct a decomposition with at most bimolecular reactions which also verifies nonambiguity (Eq. (2.15)), by writing

$$X + B \overset{1\alpha}{\rightleftharpoons} BX, \tag{5.75}$$

$$BX + X \overset{1\beta}{\rightleftharpoons} X_3, \tag{5.76}$$

$$X_3 \overset{1\gamma}{\rightleftharpoons} X_2 + X, \tag{5.77}$$

$$X_2 \overset{1\delta}{\rightleftharpoons} 2X. \tag{5.78}$$

if reaction 1β is a sufficiently slow rate-limiting step, 1α, 1γ, 1δ will rapidly establish local equilibrium, characterized by equilibrium constants

$$K_{1\alpha} = \frac{[BX]}{[X][B]}, \quad K_{1\gamma} = \frac{[X_3]}{[X_2][X]}, \quad K_{1\delta} = \frac{[X_2]}{[X]^2}. \tag{5.79}$$

For the rate of reaction 1β we then have

$$R_{1\beta} = k^+_{1\beta}[BX][X] - k^-_{1\beta}[X_3] = k^+_{1\beta}K_{1\alpha}[B][X]^2 - k^-_{1\beta}K_{1\gamma}K_{1\delta}[X]^3. \tag{5.80}$$

Provided $[X] \gg [X_2],[X_3],[BX]$, Eq. (5.74) provides a satisfactory simplification of the more detailed network given by reactions (5.75) to (5.78).

The more detailed reaction network is now fit for analysis within our framework, since it now satisfies nonambiguity (Eq. (2.15)) and all reactions are at most bimolecular (Eq. (2.9)). The stoichiometric matrix for this network is

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{X} \\ \mathrm{BX} \\ \mathrm{X_3} \\ \mathrm{X_2} \\ \mathrm{B} \end{array} \begin{array}{c} \begin{array}{ccccc} 1\alpha & 1\beta & 1\gamma & 1\delta & \end{array} \\ \begin{pmatrix} -1 & -1 & 1 & 2 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}. \tag{5.81}$$

Let us now proceed by taking a submatrix where the reservoir species B is removed, $s_r = \{5\}$. This generates an autonomous submatrix $\boldsymbol{\nu}_* = \boldsymbol{\nu}(5|5)$

$$\boldsymbol{\nu}_* = \begin{array}{c} \\ \mathrm{X} \\ \mathrm{BX} \\ \mathrm{X_3} \\ \mathrm{X_2} \end{array} \begin{array}{c} \begin{array}{cccc} 1\alpha & 1\beta & 1\gamma & 1\delta \end{array} \\ \begin{pmatrix} -1 & -1 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \end{array}. \tag{5.82}$$

Which is a square, full-rank matrix. Since it is also autonomous, it immediately follows that it admits SFR and by extension (see Sec. 5.7) SFA and that it is capable of autocatalysis.

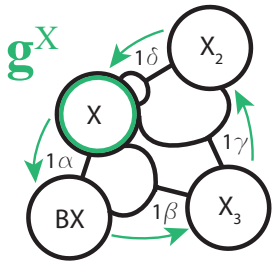


Figure 5.11: Hypergraph corresponding to $\boldsymbol{\nu}_*$. The replication cycle $\boldsymbol{g}^X$ is illustrated in green. The forward reaction 1β makes this network an example of joint autocatalysis. The consumption of two autocatalysts in 1β is compensated by the production of two autocatalysts in 1γ and 1δ.

Since $\boldsymbol{\nu}$ is full-rank, replication cycles $\boldsymbol{g}$ can be found for each component individually by the inverse $\bar{\boldsymbol{\nu}}^{-1} = (\boldsymbol{g}^{\mathrm{X}}, \boldsymbol{g}^{\mathrm{BX}}, \boldsymbol{g}^{\mathrm{X_3}}, \boldsymbol{g}^{\mathrm{X_2}})$. Here $\bar{\boldsymbol{\nu}}^{-1}$ is

$$\bar{\boldsymbol{\nu}}^{-1} = \begin{array}{c} 1\alpha \\ 1\beta \\ 1\gamma \\ 1\delta \end{array} \begin{pmatrix} 1 & 2 & 3 & 2 \\ 1 & 1 & 3 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 1 \end{pmatrix} \tag{5.83}$$

Interestingly, $\boldsymbol{\nu}_*$ contains two submatrices $\boldsymbol{\nu}_\bullet, \boldsymbol{\nu}_\circ$ that also admit SFA, but with autocatalytic cycles performed in opposite directions. The first one is obtained by removing $\mathrm{X_2}$ and reaction 1δ $\boldsymbol{\nu}_\bullet = \boldsymbol{\nu}_*(4|4)$, the second one is obtained by removing BX and reaction 1α, $\boldsymbol{\nu}_\circ = \boldsymbol{\nu}_*(2|1)$

$$\boldsymbol{\nu}_\bullet = \begin{array}{c} \mathrm{X} \\ \mathrm{BX} \\ \mathrm{X_3} \end{array} \begin{pmatrix} -1 & -1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad \boldsymbol{\nu}_\circ = \begin{array}{c} \mathrm{X} \\ \mathrm{X_3} \\ \mathrm{X_2} \end{array} \begin{pmatrix} -1 & 1 & 2 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \tag{5.84}$$

which have inverses

$$\bar{\boldsymbol{\nu}}_{\bullet}^{-1} = \begin{matrix} 1\alpha \\ 1\beta \\ 1\gamma \end{matrix} \begin{pmatrix} -1 & 0 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}, \quad \bar{\boldsymbol{\nu}}_{\circ}^{-1} = \begin{matrix} 1\beta \\ 1\gamma \\ 1\delta \end{matrix} \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}. \tag{5.85}$$

From the sign change between the inverses, it is directly clear that the replication cycles in $\bar{\boldsymbol{\nu}}_{\bullet}^{-1}$ are performed in the opposite direction (see Fig. 5.12)

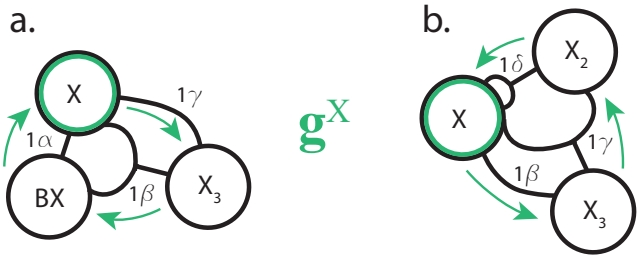


Figure 5.12: Hypergraphs corresponding to a) $\boldsymbol{\nu}_{\bullet}$ and b) $\boldsymbol{\nu}_{\circ}$. The replication cycles $\boldsymbol{g}^X$ (illustrated in green) run in opposite directions.

Now that we have established how joint autocatalysis can be treated within the framework, let us look at an interesting property that can come with bimolecular reactions between autocatalysts.

Thresholds and bistability

In solitary autocatalysis, a reaction can start from a single autocatalyst molecule. For joint autocatalysis, this ceases to be true: there need to be enough autocatalysts such that they can plausibly meet. Let us now consider adding an irreversible degradation reaction

$$2\mathrm{X} + \mathrm{B} \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} 3\mathrm{X}, \tag{5.86}$$

$$\mathrm{X} \xrightarrow{k_2^+} \emptyset, \tag{5.87}$$

where k_1^+, k_1^-, k_2^+ are rate constants. The reservoir species B is maintained at a fixed concentration [B]. Let us now examine the kinetic equation for [X]

$$\frac{d[\mathrm{X}]}{dt} = k_1^+[\mathrm{X}]^2[\mathrm{B}] - k_1^-[\mathrm{X}]^3 - k_2^+[\mathrm{X}]. \tag{5.88}$$

At steady-state, [X] is found from the third-order polynomial equation, which admits the solutions

$$[\mathrm{X}] = 0 \ \vee \ [\mathrm{X}]^- = \frac{k_1^+[\mathrm{B}] - \sqrt{(k_1^+[\mathrm{B}])^2 - 4k_1^-k_2^+}}{2k_1^-} \ \vee [\mathrm{X}]^+ = \frac{k_1^+[\mathrm{B}] + \sqrt{(k_1^+[\mathrm{B}])^2 - 4k_1^-k_2^+}}{2k_1^-} \tag{5.89}$$

To have any chance of accumulating in the face of degradation, [X] must be sufficiently abundant, we thus expect the fixed point $[\mathrm{X}] = 0$ to be inherently stable. This can be confirmed from its linear stability to a small perturbation ε

$$\frac{d[\mathrm{X}]}{dt} = k_1^+\varepsilon^2[\mathrm{B}] - k_1^-\varepsilon^3 - k_2^+\varepsilon. \tag{5.90}$$

The leading order contribution $-k_2^+\varepsilon < 0$ confirms that such a reaction cannot be started with a small amount of autocatalysts. For $(k_1^+[B])^2 > 4k_1^-k_2^+$, $[\mathrm{X}]^-$ and $[\mathrm{X}]^+$ are physical solutions. Since

$[X]^+ > [X]^-$ and because 0 is a stable fixed point, $[X]^-$ is an unstable fixed point and $[X]^+$ is again a stable fixed point. When $(k_1^+[B])^2 < 4k_1^-k_2^+$, the only fixed point is 0.

From physical considerations, we may expect such behavior to be typical for joint autocatalysis in the presence of degradation. The characteristic encounter of two autocatalysts (5.73) stabilizes a state without autocatalysts. For the right combination of parameters and concentrations, autocatalysis is no longer outpaced by degradation and the autocatalytic process can be maintained.

Schlögl's termolecular reaction is typically combined with a second reservoir, to fix the concentration of a molecule A that interconverts to X. This leads to the reaction network:

$$2X + B \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} 3X, \tag{5.91}$$

$$X \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} A. \tag{5.92}$$

Our choice for a degradation reaction can be recast in this tradition, if we set $[A] = 0$. If we consider the quantity corresponding to equilibrium of: $X \rightleftharpoons A$, i.e.

$$[X]^* = \frac{k_2^-}{k_2^+}[A]. \tag{5.93}$$

The kinetic term due to formation of X from a reservoir species can be incorporated in a weakened degradation rate for concentrations above $[X]^*$

$$\frac{d[X]}{dt} = k_1^+[X]^2[B] - k_1^-[X]^3 - k_2^+([X] - [X]^*). \tag{5.94}$$

Note that when $[A] > 0$, $[X] = 0$ ceases to be a fixed point. There exist parameter ranges for rate constants and $[A], [B]$ for which a state with abundant autocatalysts is the only fixed point[12, 15].

Schlögl's reaction network[12] has been studied as a minimal model for bistability [15, 16] and pattern formation[17, 18, 19] in chemistry. Upon adding a third reaction, the Brusselator network is obtained, which can exhibit oscillations[20, 21].

5.6 Thermodynamic spontaneity and autocatalysis

We will now consider the spontaneity for autocatalytic reactions, i.e. we will consider whether $\Delta G < 0$. A recent analysis discussed the question of spontaneity[22], but considered only the role of the standard free energy of formation. Here, we will show that the disorder term can provide an important contribution to ΔG as well.

We will ask the question of spontaneity for 'triggering' autocatalysis, i.e. perturbing a system such that a i) previously inactive or i) poorly active autocatalytic cycle reinforces itself. For this analysis, we will distinguish between i) solitary autocatalysis and ii) joint autocatalysis. The critical difference between these two is that the former starts from a single autocatalyst, whereas the latter involves the presence of a critical mass of autocatalysts.

5.6.1 Spontaneity of solitary autocatalysis

As an example of the first case, we will consider the Toy Formose reaction network in a CSTR with a steady influx of C_1 and H_2O. The reaction $C_1 + C_1 \rightarrow C_2$ does not take place. However, when a first molecule of C_2, C_3 or C_4 enters, the autocatalytic cycle can directly be triggered. In particular at the start, such a cycle should be extremely favorable, since abundant molecules are turned to rare ones.

Consider our first C_2 molecule, which performs the net reaction $2C_1 + C_2 \longrightarrow 2C_2$. The associated free energy change of the system is then

$$\Delta G_{cycle} = \mu^\circ_{C_2} - 2\mu^\circ_{C_1} + kT \ln \frac{(N_{C_2}+1)N_T}{N_{C_1}(N_{C_1}-1)} \tag{5.95}$$

Where $N_T = N_{C_1} + N_{C_2} + N_{C_3} + N_{C_4} + N_{H_2O}$. The latter term corresponds to disorder, and is reminiscent of a mixing entropy. If we have e.g. 1 Moles of C_1 and 55 Moles of H_2O, the latter term would yield $-50kT$ for the first cycle. Consequently, the nucleation of autocatalytic cycles is favored by thermodynamics, even if $\Delta\mu^\circ$ is not favorable to such a reaction.

Various authors have claimed that autocatalysis requires $\Delta\mu^\circ < 0$, and some have claimed that this must even be true for every single reaction step. From Eq. (5.95) it follows that this is too strict a requirement, the disorder contribution can compensate an unfavorable $\Delta\mu^\circ$. The step $C_4 \longrightarrow 2C_2$ in formose is a case in point where $\Delta\mu^\circ > 0$. Due to the disorder contribution, triggering solitary autocatalysis is quite generally expected to be thermodynamically spontaneous. Whether triggering will be successful (see Ch.6) will largely be a question of kinetics[‡].

5.6.2 Spontaneity of joint autocatalysis

For joint autocatalysis, the argument is a bit different, since we cannot trigger the reaction with a single molecule: multiple autocatalysts must be capable of meeting each other appreciably to finish the autocatalytic cycle. The rate at which this must occur is set by the timescales of competing processes such as degradation.

As an example, let us take Schlögl's termolecular reaction,

$$2X + B \underset{-\boldsymbol{g}^X}{\overset{\boldsymbol{g}^X}{\rightleftharpoons}} 3X. \tag{5.96}$$

with B a chemostatted species. Let N^*_X be the abundance of X where encounters are numerous enough to compete with other processes like degradation.

If we now perform an autocatalytic cycle, such that $N'_X = N^*_X + 1$, the associated free energy change is

$$\Delta G_{cycle} = \mu^\circ_X - \mu^\circ_B + kT \ln \frac{N^*_X + 1}{N_B}. \tag{5.97}$$

For this reaction to be competitive, the forward reaction must at least be as fast as degradation (but this is not enough, as explained in Sec. 5.5). When the rate of forward reaction and degradation (which we assume unimolecular) are equal, we find $k^+ N_X^2 N_B = k^\emptyset N_X$, which means that $N^*_X \propto \frac{1}{\tau}$.

At this point, the abundance of X will far exceed the single-molecule limit $N^*_X \gg 1$, making its disorder contribution weaker than for solitary autocatalysis. The contribution of $\Delta\mu^\circ$ will then become more important in determining the spontaneity of autocatalysis.

5.7 SFR, SFA and autocatalysis

Let us start by reiterating the conventions that a stoichiometric matrix $\boldsymbol{\nu}$ must respect in our framework. First, reactions are decomposed in steps that are at most bimolecular

$$\forall i \quad \sum_k \nu^{(+)}_{ki} \leq 2, \sum_k \nu^{(-)}_{ki} \leq 2. \tag{5.98}$$

[‡]Of course, if $\Delta\mu^\circ$ becomes too large, the kinetics will in turn be prohibitively slow.

Second, a species involved in a reaction step is either a reactant or a product, but never both

$$\forall k,i\ \nu_{ki}^{(-)}\nu_{ki}^{(+)}=0, \tag{5.99}$$

which we refer to as nonambiguity.

Let $\boldsymbol{\nu}$ be a stoichiometric matrix that performs Stoichiometrically Feasible Reproduction (SFR). SFR implies that $\boldsymbol{\nu}$ is autonomous: every reaction has one or more reactants and one or more products. Furthermore, there exists a reaction vector $\boldsymbol{\Gamma}$, s.t. $\forall i,(\boldsymbol{\nu}\cdot\boldsymbol{\Gamma})_i\geq 1$, which is equivalent to stating that there exists an overall reaction

$$n_{\mathrm{A}}\mathrm{A}+n_{\mathrm{B}}\mathrm{B}+...\underset{-\boldsymbol{\Gamma}}{\overset{+\boldsymbol{\Gamma}}{\rightleftharpoons}}m_{\mathrm{A}}\mathrm{A}+m_{\mathrm{B}}\mathrm{B}+...\ ,\quad \forall\ i\ \ m_{\mathrm{X_i}}>n_{\mathrm{X_i}}\geq 0. \tag{5.100}$$

A network that is SFA, comes with a slightly stronger requirement:

$$n_{\mathrm{A}}\mathrm{A}+n_{\mathrm{B}}\mathrm{B}+...\underset{-\boldsymbol{\Gamma}}{\overset{+\boldsymbol{\Gamma}}{\rightleftharpoons}}m_{\mathrm{A}}\mathrm{A}+m_{\mathrm{B}}\mathrm{B}+...\ ,\quad \forall\ i\ \ m_{\mathrm{X_i}}>n_{\mathrm{X_i}}\geq 1. \tag{5.101}$$

Throughout these notes, we will maintain the convention that $\boldsymbol{\Gamma}_i\geq 1$, which can always be achieved by the appropriate choice of sign for reactions in $\boldsymbol{\nu}$. Since those reactions are reversible, the choice of their sign is arbitrary. By construction, a positive entry on the left of eqs. (5.100) and (5.101) corresponds to consumption, with the consumption and production for X_i being

$$n_i=(\boldsymbol{\nu}^{+}\cdot\boldsymbol{\Gamma})_i,\quad m_i=(\boldsymbol{\nu}^{-}\cdot\boldsymbol{\Gamma})_i. \tag{5.102}$$

It follows directly that SFA implies every species is both consumed and produced. Eq. (2.15) shows that a reaction cannot both consume and produce the same species. In SFA, any species X_i is consumed by at least one reaction, and at least produced by one different reaction.

All cycles can be removed

Let us recall that the rank of $\boldsymbol{\nu}$ corresponds to the number of reactions minus cycles:

$$\mathrm{rank}(\nu)=\mathrm{r}-\mathrm{c}. \tag{5.103}$$

A reaction network that has cycles ($c>0$) has, by definition, more reactions than necessary to span its concentration space whose dimension is $\mathrm{rank}(\nu)$. As noted in Sec. 2.6, the removal of a reaction has one of the following two consequences: i) removal of a cycle $c'=c-1$. ii) generation of a conservation law $l'=l+1$. As long as $c>0$, one can always remove a reaction that is part of the reaction vector $\boldsymbol{c}$ of a cycle, to make that cycle impossible. It follows that we can remove any excess cycles by removing the appropriate surplus reactions, to yield a network with $c=0$. In the rest of this section, we will always consider that the network has been pruned in this fashion, so that we can rapidly find the necessary irreducible submatrices for the rest of our demonstration.

SFR not SFA implies reducibility

Let us now consider a network that performs SFR but not SFA. Eqs. (5.100) and (5.101) allow for this situation, if a choice of $\boldsymbol{\Gamma}$ yields $n_{X_k}=0$ for at least one species X_k. The species X_k is then only produced. For any reaction r, we distinguish between two ways in which this can happen: O1) X_k is the only product. O2) X_k is produced along with another molecule. in Fig. 5.13 these reactions are illustrated.

Let us now consider removing X_k from our description. This operation transforms an O1-reaction to an $\emptyset$-reaction. An O2-reaction will still have another product. Letting r_{O1} be the set of indices for $O1$ reactions in which X_k engages, it follows that the submatrix from which X_k and r_{O1} are removed, $\boldsymbol{\nu}^{*}=\boldsymbol{\nu}(k|r_{O1})$, is autonomous.

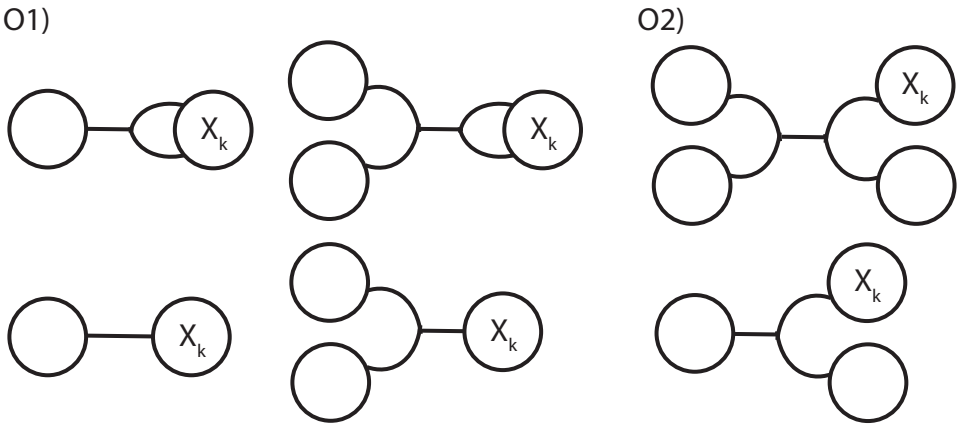


Figure 5.13: Illustration of all possible O1 reactions and O2 reactions for X_k allowed by Eq. (5.98).

Upon removal of r_{O1}, the overall reaction balance (eq. (5.100)) is modified. Since we have only touched on reactions that produce X_k, all other species are still produced upon the removal of r_{O1} from $\mathbf{\Gamma}$

$$m^*_{X_j} = m_{X_j} \geq n_{X_j} \geq 0, j \neq k. \tag{5.104}$$

Which guarantees that the submatrix $\boldsymbol{\nu}^* = \boldsymbol{\nu}(k|r_{O1})$ admits SFR. It follows that a network that is SFR but not SFA is reducible to a smaller network that is SFR.

Autonomy requires that O1-reactions have reactants and products. If r_{O1} is nonempty, the removal of reactions lead to a corresponding decrease of consumed reactants: $\exists j \neq k, n^*_{X_j} < n_{X_j}$.

The smallest SFR networks

In an SFR network, all species are produced, but not all species are consumed. Autonomy requires that there are no $\emptyset$-reactions, which means that at least some species must also be consumed. Eq. (5.99) then implies that an SFR network must have at least two reactions. Let us now consider these smallest networks for which $r = 2$.

Eq. (5.98) only allows a species to be consumed/produced with stoichiometry 1 or 2. We require all species to be produced and at least one to be consumed as well. $r = 2$ then implies there is a species X_k s.t. $n_k = 1$, $m_k = 2$, which is achieved by two separate reactions. Two reactions can produce at most 4 species, so at most two other species than X_k can be produced for $r = 2$. An SFR with $r = 2$ must have $2 \leq s \leq 3$.

For $s = 3$, there are two distinct SFRs to consider:

$$\boldsymbol{\nu}_1 = \begin{pmatrix} -1 & 2 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}, \quad \boldsymbol{\nu}_2 = \begin{pmatrix} -1 & 2 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} \tag{5.105}$$

Where $\boldsymbol{\nu}_1$ admits SFA, and $\boldsymbol{\nu}_2$ does not. However, the submatrix afforded by removing the third row, $\boldsymbol{\nu}_2(3|\emptyset)$, does.

For $s = 2$, all SFRs are SFA:

$$\boldsymbol{\nu}_3 = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}, \quad \boldsymbol{\nu}_4 = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}, \quad \boldsymbol{\nu}_5 = \begin{pmatrix} -1 & 2 \\ 2 & -2 \end{pmatrix} \tag{5.106}$$

SFR implies SFA

By our reduction operation, a network that performs SFR but not SFA becomes a smaller subnetwork that again admits SFR, and may or may not admit SFA. This procedure removes one species at

a time, while guaranteeing that this operation yields a network that is autonomous and admits SFR. When this procedure encounters SFA, the protocol can no longer be applied. The reduction operation is guaranteed to find an SFA at some point, since sufficiently reduced networks can only be SFA ($s = 2, r = 2$).

SFA and reducibility

A stoichiometric matrix $\boldsymbol{\nu}$ that admits SFA may in turn contain smaller submatrices that are also SFA. In particular, SFA allows a network to have mixed conservation laws $\boldsymbol{\ell}^{\square}$, s.t. $\boldsymbol{\ell}^{\square} \cdot \boldsymbol{\nu} = 0$. In an SFA, such a conservation law requires there to be at least two bimolecular steps, such that molecules are formed and consumed together. We will now demonstrate the following: a matrix $\boldsymbol{\nu}$ that admits SFA and with a mixed conservation law $\boldsymbol{\ell}^{\square}$ can be reduced to a submatrix that admits an SFR. This is achieved by removing a species consumed in a bimolecular reaction.

Having removed all cycles ($c = 0$), it follows that $s - \ell = r$. Let us now consider the bimolecular consumption reactions involving two distinct species.If there are r' of such bimolecular consumption reactions, the other $r - r'$ consume only one type of species. Since we require that every species must be consumed at least once, we require at least $s - r'$ reactions.

Upon removing a species X_k, any $O1$ reactions that form X_k must be removed to preserve autonomy, any $O2$ reactions that form X_k remain untouched. In principle, X_k can also be consumed by $O1$ reactions in the forward sense. We will now show that for SFA with bimolecular consumption, there must exist a species X_k that is not consumed by such reactions.

Let us first consider the converse: every species in bimolecular consumption reactions is also consumed by forward O1 reactions. Then every species is involved in at least one O1 reaction, meaning $r \geq s + r'$. Equating this to $s - \ell$, with $\ell > 0$, we run into a contradiction. In the presence of a mixed conservation law, there must therefore be a species engaged in a forward O2 reaction that does not engage in forward O1 reactions, as before, let us construct a submatrix $\boldsymbol{\nu}^* = \boldsymbol{\nu}(k|r_{O1})$. Upon the removal of r_{O1} from $\boldsymbol{\Gamma}$, all non-X_k species are still produced in equal amount

$$m^*_{\mathrm{X}_j} = m_{\mathrm{X}_j} \geq n_{\mathrm{X}_j} \geq 0, j \neq k. \tag{5.107}$$

Which guarantees that the submatrix $\boldsymbol{\nu}^* = \boldsymbol{\nu}(\mathrm{X}_k|r_{O1})$ admits SFR. SFA with a mixed conservation law thus implies reducibility to a smaller SFR.

Further reductions

SFR implies reducibility to a smaller network that performs SFA. SFA with mixed conservation laws implies reducibility to a smaller network that performs SFR. Upon repetition of the derived procedures, we must then end up with a network that i) performs SFA and ii) has no mixed conservation laws.

By definition, SFR and SFA exclude mass-like conservation laws $\boldsymbol{\ell}^{+}$. Furthermore, we can trivially remove the appropriate reactions to get rid of cycles, to span the same concentration space. Then, $\ell = 0$ and $c = 0$. Applying the fundamental theorem of linear algebra

$$s - \ell = r - c, \tag{5.108}$$

we are left with $s = r$, a full-rank square matrix. It follows that a network that performs SFR must contain an autonomous and invertible submatrix that only performs SFA.

SFR not SFA implies reducibility. For an invertible SFA, reducibility can still occur. As evidenced by the Schlögl network (Sec. 5.5), an invertible SFA network can still have submatrices that are themselves SFA and invertible. These submatrices were obtained by removing a redundant reactant and a corresponding $O1$ reaction that this generates.

5.8 Autocatalysis and frustrated amplification: composite chemostats

Autocatalysis tends to be portrayed as the action of a set of actively accumulating autocatalysts, feeding on an abundant food supply, which often takes the form of a chemostat. In our characterization of SFR and SFA, we have implicitly supposed that if food comes from chemostats, such chemostats are 'simple'. For composite chemostats, some particular behavior can occur, on which we will elaborate in this section with an example that can perform autocatalysis through simple chemostats and 'frustrated amplification' through composite chemostats.

As discussed in Sec. 3.1, a chemostat can either be simple

$$\emptyset \rightleftharpoons \mathrm{A}, \tag{5.109}$$

thereby fixing the chemical potential (and thereby concentration) of A, or it can be composite:

$$\emptyset \rightleftharpoons \mathrm{A}+\mathrm{B}, \tag{5.110}$$

which fixes the combined chemical potential, and thereby a product of concentrations (e.g. a solubility product) $x_\mathrm{A}x_\mathrm{B}$. When such a nonlinearity is in place, feedbacks can occur on the level of the food: the consumption of A will lead to an accumulation of B.

Let us illustrate this by a literature example using molecular cages [23, 24], which reversibly traps a carbodiimide (DCC) in a cage C, according to a reaction

$$[\mathrm{DCC}:\mathrm{C}] \rightleftharpoons \mathrm{C}+\mathrm{DCC}. \tag{5.111}$$

The reactant DCC reacts with a carboxylic acid moiety $\mathrm{RCO_2H}$, to form a reactive intermediate X

$$\mathrm{DCC}+\mathrm{RCO_2H} \rightleftharpoons \mathrm{X}. \tag{5.112}$$

Subsequently, this intermediate X can react with an amine moiety $\mathrm{RNH_2}$ to yield an amide product P and a urea moiety DCU:

$$\mathrm{X}+\mathrm{RNH_2} \rightleftharpoons \mathrm{P}+\mathrm{DCU}. \tag{5.113}$$

These species form more stable complexes with the cage

$$\mathrm{C}+\mathrm{DCU} \rightleftharpoons [\mathrm{DCU}:\mathrm{C}], \tag{5.114}$$

$$\mathrm{C}+\mathrm{P} \rightleftharpoons [\mathrm{P}:\mathrm{C}]. \tag{5.115}$$

In Ref. [24, 23], a displacement process was considered to be the dominant contribution to the exchange of caged species (the lifetime of the release by the cage being estimated at $10^7 s$):

$$[\mathrm{DCC}:\mathrm{C}]+\mathrm{DCU} \rightleftharpoons [\mathrm{DCU}:\mathrm{C}]+\mathrm{DCC}, \tag{5.116}$$

$$[\mathrm{DCC}:\mathrm{C}]+\mathrm{P} \rightleftharpoons [\mathrm{P}:\mathrm{C}]+\mathrm{DCC}. \tag{5.117}$$

Overall, a caged species $[\mathrm{DCC}:\mathrm{C}]$ liberates C, which reacts to form two better guest species DCU and P, which subsequently liberate 2DCC for their cages, leading to an overall balance consistent with stoichiometric autocatalysis for the species DCC

$$\mathrm{DCC}+\mathrm{RCO_2H}+\mathrm{RNH_2}+2[\mathrm{DCC}:\mathrm{C}] \rightleftharpoons 2\mathrm{DCC}+[\mathrm{P}:\mathrm{C}]+[\mathrm{DCU}:\mathrm{C}] \tag{5.118}$$

Notice that, with the appropriate rates and separations of timescales (which may not be the case here), such a process can also be afforded without a displacement reaction, but rather a composition of exchanges with the cage.

In the next sections, we will show that this alternative situation requires a different approach, as we need to consider the feedback on a composite chemostat of empty cages and reactant.

With displacement reaction

Let us first consider the displacement pathway, by constructing a stoichiometric matrix $\boldsymbol{\nu}$

$$\boldsymbol{\nu} = \begin{array}{c} \\ [\text{DCC:C}] \\ \text{C} \\ \text{DCC} \\ \text{RCO}_2\text{H} \\ \text{X} \\ \text{RNH}_2 \\ \text{P} \\ \text{DCU} \\ [\text{P:C}] \\ [\text{DCU:C}] \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{aligned} [\text{DCC} : \text{C}] &\rightleftharpoons \text{C} + \text{DCC} \\ \text{DCC} + \text{RCO}_2\text{H} &\rightleftharpoons \text{X} \\ \text{X} + \text{RNH}_2 &\rightleftharpoons \text{P} + \text{DCU} \\ [\text{DCC} : \text{C}] + \text{P} &\rightleftharpoons [\text{P} : \text{C}] + \text{DCC} \\ [\text{DCC} : \text{C}] + \text{DCU} &\rightleftharpoons [\text{DCU} : \text{C}] + \text{DCC} \end{aligned} \tag{5.119}$$

We can construct an autocatalytic submatrix $\boldsymbol{\nu}^*$, by removing $\text{DCC} : \text{C}], \text{C}, \text{RCO}_2\text{H}, \text{RNH}_2, [\text{P} : \text{C}], [\text{DCU} : \text{C}$, and the initial release reaction, to yield the considerably smaller submatrix

$$\boldsymbol{\nu}^* = \begin{array}{c} \\ \text{DCC} \\ \text{X} \\ \text{P} \\ \text{DCU} \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \qquad \begin{aligned} \text{DCC} &\rightleftharpoons \text{X} \\ \text{X} &\rightleftharpoons \text{P} + \text{DCU} \\ \text{P} &\rightleftharpoons \text{DCC} \\ \text{DCU} &\rightleftharpoons \text{DCC} \end{aligned} \tag{5.120}$$

Note that $\boldsymbol{\nu}^*$ is invertible and autonomous. From the hypergraph of the subnetwork (Fig. 5.14) it also becomes directly apparent that it is an irreducible example of an SFA.

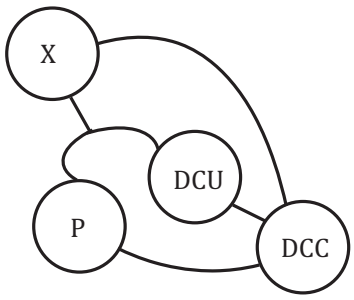


Figure 5.14: Molecular cages performing autocatalysis.

Passive exchange

Let us now consider a situation where we can stoichiometrically still obtain (5.123), but where the feedback occurs on the level of a composite chemostat.

$$\boldsymbol{\nu} = \begin{array}{c} \\ [\mathrm{DCC{:}C}] \\ \mathrm{C} \\ \mathrm{DCC} \\ \mathrm{RCO_2H} \\ \mathrm{X} \\ \mathrm{RNH_2} \\ \mathrm{P} \\ \mathrm{DCU} \\ [\mathrm{P{:}C}] \\ [\mathrm{DCU{:}C}] \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \end{array} \quad \begin{array}{c} [\mathrm{DCC:C}] \rightleftharpoons \mathrm{C} + \mathrm{DCC} \\ \mathrm{DCC} + \mathrm{RCO_2H} \rightleftharpoons \mathrm{X} \\ \mathrm{X} + \mathrm{RNH_2} \rightleftharpoons \mathrm{P} + \mathrm{DCU} \\ [\mathrm{P:C}] \rightleftharpoons \mathrm{C} + \mathrm{P} \\ [\mathrm{DCU:C}] \rightleftharpoons \mathrm{C} + \mathrm{DCU} \end{array} \tag{5.121}$$

Contrary to stoichiometric autocatalysis derived before, $\boldsymbol{\nu}$ admits no SFA submatrix. What happens in this network is fundamentally different: a dynamic equilibrium

$$[\mathrm{DCC:C}] \rightleftharpoons \mathrm{C} + \mathrm{DCC} \tag{5.122}$$

is shifted, by the consumption of C by P and DCU. If this dynamic equilibrium is the fastest reaction and $[\mathrm{DCC:C}]$ is abundant, then the product $x_C x_{DCC}$ remains fixed. One x_{DCC} is consumed for every two C in the transformation $\boldsymbol{g} = (1,1,1,-1,-1)$, leading to a participation balance

$$\begin{aligned} 2[\mathrm{DCC:C}] + \mathrm{DCC} + 2\mathrm{C} \quad &+ \quad \mathrm{X} + \mathrm{RCO_2H} + \mathrm{RNH_2} + \mathrm{DCU} + \mathrm{P} \overset{\boldsymbol{g}}{\underset{-\boldsymbol{g}}{\rightleftharpoons}} \\ 2\mathrm{DCC} \quad &+ \quad 2\mathrm{C} + \mathrm{X} + \mathrm{DCU} + \mathrm{P} + [\mathrm{P:C}] + [\mathrm{DCU:C}]. \end{aligned} \tag{5.123}$$

Coupled with rapid equilibration of (5.122), C will decrease in abundance, and DCC will increase. If the consumption of DCC or X is rate-limiting, such a feedback has an accelerating effect. Upon this acceleration, other reactions become limiting, which is either the mass injection by the chemost (Eq. (5.122)) or reaction with the cages. These reactions are not accelerated by the feedback in place and a linear regime is entered.

An instructive alternative perspective, is consider what would happen if DCC would not form a complex with cage C, and instead both species were abundantly present, or resupplied by chemostats (e.g. $\mathrm{DCC}^I \rightleftharpoons \mathrm{DCC}^{II}$). We can then remove C, DCC and reaction (5.122) from our description, to yield a subnetwork:

$$\boldsymbol{\nu} = \begin{array}{c} \\ \mathrm{RCO_2H} \\ \mathrm{X} \\ \mathrm{RNH_2} \\ \mathrm{P} \\ \mathrm{DCU} \\ [\mathrm{P{:}C}] \\ [\mathrm{DCU{:}C}] \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{array} \quad \begin{array}{c} \mathrm{RCO_2H} \rightleftharpoons \mathrm{X} \\ \mathrm{X} + \mathrm{RNH_2} \rightleftharpoons \mathrm{P} + \mathrm{DCU} \\ [\mathrm{P:C}] \rightleftharpoons \mathrm{P} \\ [\mathrm{DCU:C}] \rightleftharpoons \mathrm{DCU} \end{array} \tag{5.124}$$

The subnetwork has no emergent cycles and admits a mass-like conservation law

$$L^+ = n_{\mathrm{RCO_2H}} + n_{\mathrm{RNH_2}} + n_{\mathrm{X}} + n_{\mathrm{P}} + n_{\mathrm{DCU}} + n_{[\mathrm{P:C}]} + [\mathrm{DCU:C}], \tag{5.125}$$

meaning it is neither an allocatalytic nor autocatalytic matrix. In the original scheme, RNH_2 and RCO_2H were food sources as well. Removing them would break autonomy, since this would yield a reaction:

$$\emptyset \rightleftharpoons X. \tag{5.126}$$

If we were to consider X as abundant as well (due to a chemostatting via RCO_2H and DCC), a bimolecular $\emptyset$ reaction is obtained

$$\emptyset \rightleftharpoons P + DCU. \tag{5.127}$$

again, autonomy is not preserved.

Frustrated amplification and bimolecular $\emptyset$-reactions

If we wish to characterize such 'frustrated amplification' in terms of a submatrix, we can no longer use the SFA criterion. The dual role of food and feedback in Eq. (5.122) leads to a bimolecular $\emptyset$-reaction, if we remove $[DCC : C]$ from the description:

$$\emptyset \rightleftharpoons C + DCC. \tag{5.128}$$

Similarly, we can turn the consumption of C in bimolecular $\emptyset$-reactions:

$$C + P \rightleftharpoons \emptyset, \tag{5.129}$$
$$C + DCU \rightleftharpoons \emptyset. \tag{5.130}$$

Contrary to unimolecular $\emptyset$ reactions, which are a trivial means of breaking mass-conservation that are not prone to feedback, bimolecular $\emptyset$ reactions can be acted upon. Let us now define a property to formalize this:

Definition 5.8.1 — Pseudoautonomy. A stoichiometric submatrix $\boldsymbol{\nu}_*$ is pseudo-autonomous, if it contains bimolecular $\emptyset$-reactions, but no unimolecular $\emptyset$-reactions.

Using this definition, a network is either autonomous, pseudoautonomous, or neither.

If we now consider $[DCC : C], RCO_2H$ and RNH_2 as 'food' and $[P : C]$, $[DCU : C]$ as 'waste', we can write a stoichiometric submatrix

$$\tilde{\boldsymbol{\nu}} = \begin{array}{c} \\ \text{C} \\ \text{DCC} \\ \text{X} \\ \text{P} \\ \text{DCU} \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \begin{array}{l} \\ \emptyset \rightleftharpoons C + DCC \\ DCC \rightleftharpoons X \\ X \rightleftharpoons P + DCU \\ \emptyset \rightleftharpoons C + P \\ \emptyset \rightleftharpoons C + DCU \end{array} \tag{5.131}$$

which is pseudoautonomous and has no mass-like conservation laws. Moreover, $\tilde{\boldsymbol{\nu}}$ is invertible

$$\tilde{\boldsymbol{\nu}}^{-1} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} \text{C} & \text{DCC} & \text{X} & \text{P} & \text{DCU} \end{array} \\ \begin{pmatrix} -1 & 2 & 2 & 1 & 1 \\ -1 & 1 & 2 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 0 & 1 \\ -1 & 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \tag{5.132}$$

The replication vectors $\tilde{\boldsymbol{\nu}}^{-1} = (\boldsymbol{g}^{C}, \boldsymbol{g}^{DCC}, \boldsymbol{g}^{X}, \boldsymbol{g}^{P}, \boldsymbol{g}^{DCU})$ now elegantly capture how the replication of C is misaligned with the replication rest of the network. With respect to the other species,

C performs its reactions in the opposite sense. The replication vectors are linearly independent, so formally we can write a reaction vector $\boldsymbol{\Gamma} = \boldsymbol{g}^{\mathrm{C}} + \boldsymbol{g}^{\mathrm{DCC}} + \boldsymbol{g}^{\mathrm{X}} + \boldsymbol{g}^{\mathrm{P}} + \boldsymbol{g}^{\mathrm{DCU}}$ that replicates each species. In practice, we do not expect the species to replicate collectively, as detailed in the preceding sections.

We may wonder why this frustration of self-replication happens for this system, but not for the closely similar autocatalytic network. In the autocatalytic network, displacement reactions of the form

$$[\mathrm{DCC} : \mathrm{C}] + \mathrm{P} \rightleftharpoons [\mathrm{P} : \mathrm{C}] + \mathrm{DCC}, \tag{5.133}$$

yield the same result as the composition of a decomplexation and a complexation reaction in the chemical amplification network

$$[\mathrm{DCC} : \mathrm{C}] \rightleftharpoons \mathrm{C} + \mathrm{DCC}, \tag{5.134}$$

$$\mathrm{C} + \mathrm{P} \rightleftharpoons [\mathrm{P} : \mathrm{C}]. \tag{5.135}$$

The critical difference comes with the introduction of the dynamic species C. On the one hand, C plays the role of food that must be abundant enough to not limit the reaction rate. On the other hand, C is coupled to species DCC, which is intended to accumulate autocatalytically, and thus requires C to rapidly diminish in abundance. In the autocatalytic network, this problem is sidestepped by not having C as a substrate in the first place, but directly reacting with $[\mathrm{DCC} : \mathrm{C}]$.

Bibliography

Articles

[1] Raphaël Plasson et al. “Autocatalysis : At the Root of Self-Replication”. In: *Artif. Life* 17 (2011), pp. 219–236.

[3] R Bundschuh and U Gerland. “Dynamics of intramolecular recognition : Base-pairing in DNA / RNA near and far from equilibrium”. In: *Eur. Phys. J. E.* 19 (2006), pp. 319–329.

[4] David Weininger. “SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules”. In: *J. Chem. Inf. Comput. Sci.* 28 (1988), pp. 31–36.

[5] David Angeli, Patrick De Leenheer, and Eduardo D Sontag. “A Petri net approach to the study of persistence in chemical reaction networks”. In: *Math. Biosci.* 210 (2007), pp. 598–618.

[6] Manoj Gopalkrishnan. “Catalysis in Reaction Networks”. In: *Bull. Math. Biol.* 73.12 (2011), pp. 2962–2982.

[7] Abhishek Deshpande, Manoj Gopalkrishnan, and Computer Science. “Autocatalysis in Reaction Networks.” In: (2014). arXiv: 1309.3957v4.

[8] Michaelis Menten et al. “Thermodynamically consistent coarse graining of biocatalysts beyond Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis – Menten”. In: *New J. Phys.* 20 (2018), p. 042002.

[9] William Stillwell. “Facilitated Diffusion of Amino Acids across Bimolecular Lipid Membranes as a Model for Selective Accumulation of Amino Acids in a Primordial Protocell”. In: *BioSystems* 8 (1976), pp. 111–117.

[10] Stuart A. Kauffman. “Autocatalytic Sets of Proteins”. In: *J. Theor. Biol.* 119 (1986), pp. 1–24.

[11] Wim Hordijk, Stuart A Kauffman, and Mike Steel. “Required Levels of Catalysis for Emergence of Autocatalytic Sets in Models of Chemical Reaction Systems”. In: *Int. J. Mol. Sci* 12 (2011), pp. 3085–3101.

[12] F Schlögl. “Chemical reaction models for non-equilibrium phase transitions”. In: *Z. Phys.* 253.2 (1972), pp. 147–161.

[13] R. Aris, P. Gray, and R. Scott. “Modelling cubic autocatalysis by successive bimolecular steps”. In: *Chem. Eng. Sci.* 43 (1988), pp. 207–2011.

[14] G.B. Cook et al. “Bimolecular routes to cubic autocatalysis”. In: *J. Phys. Chem.* 93 (1989), pp. 2749–2755.

[15] P Gaspard. “Fluctuation theorem for nonequilibrium reactions”. In: *J. Chem. Phys.* 120.19 (2004), pp. 8898–8905.

[16] Melissa Vellela and Hong Qian. “Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system : the Schlögl model revisited”. In: *J. R. Soc. Interface* 6 (2009), pp. 925–940.

[17] A. Lemarchand and B. Nowakowski. “Do the internal fluctuations blur or enhance axial segmentation ?” In: *EPL* 94 (2011), p. 48004.

[18] P. Dziekan et al. “Reaction-diffusion approach to prevertebrae formation : Effect of a local source of morphogen”. In: *J. Chem. Phys.* 139 (2013), p. 114107.

[19] P Dziekan et al. “Effect of a Local Source or Sink of Inhibitor on Turing Patterns Effect of a Local Source or Sink of Inhibitor on Turing Patterns”. In: *Commun. Theor. Phys.* 62 (2014), pp. 622–630.

[21] Basile Nguyen, Udo Seifert, and Andre C Barato. “Phase transition in thermodynamically consistent biochemical oscillators”. In: *J. Chem. Phys.* 149 (2018), p. 045101. arXiv: `arXiv:1804.01080v2`.

[22] Yu Liu and David Sumpter. “Mathematical modeling reveals spontaneous emergence of self-replication in chemical reaction systems”. In: *J. Biol. Chem.* 293 (2018), pp. 18854–18863.

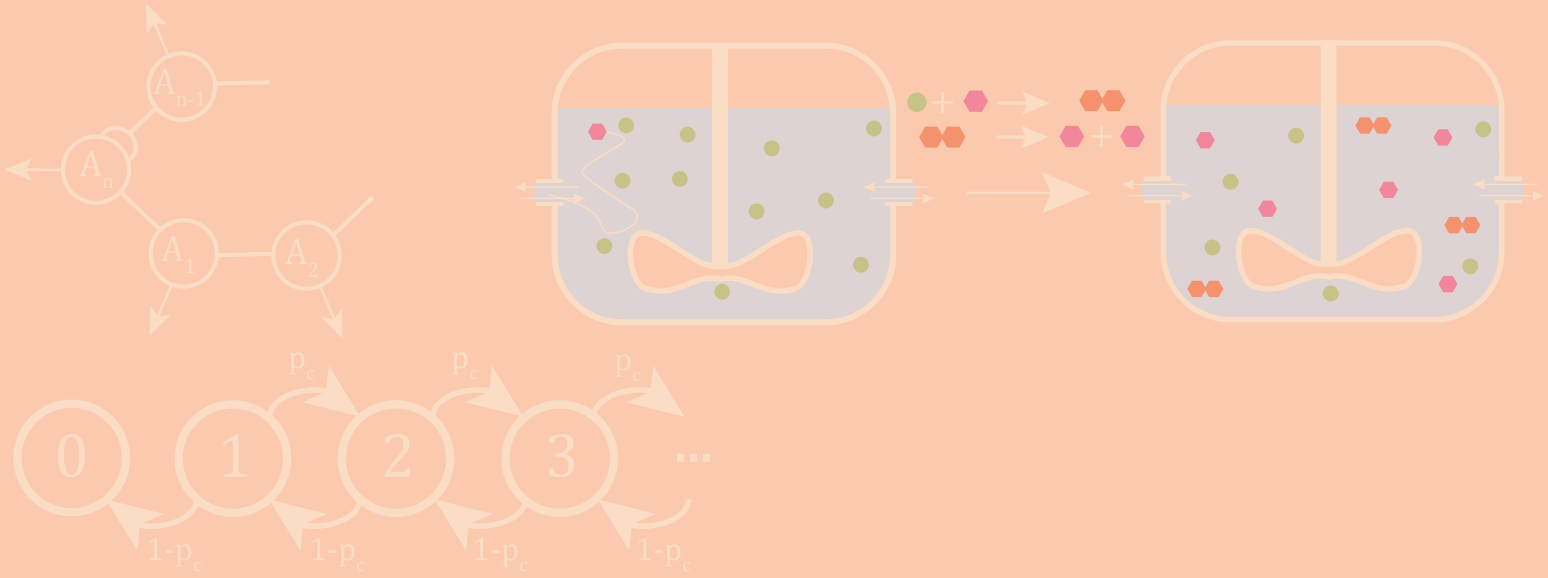
[23] Jian Chen et al. “Chemistry: Amplification by compartmentalization”. In: *Nature* 415.6870 (2002), pp. 385–386.

[24] Jian Chen et al. “Chemical amplification with encapsulated reagents”. In: *Proc. Natl. Acad. Sci. U. S. A.* 99.5 (2002), pp. 2593–2596.

Books

[2] A. D. McNaught and A. Wilkinson, eds. *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*. Oxford: Blackwell Scientific Publications, 1997.

[20] G Nicolis and I Prigogine. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*. New York: Wiley, 1977.



6. Autocatalytic Chemical Evolution

In extant biochemistry, evolution has become close to synonymous with genes. An illustrative quote comes from a joint paper between proponents and skeptics of niche construction as an extension of evolution[1]: *evolutionary processes are those that change gene frequencies*.

In origins of life and prebiotic chemistry, we are faced with the problem that we do not have 'genes' to begin with. To account for biological evolution today, many feel that there must be a selection process that mimics evolution, starting from prebiotic chemistry. In prebiotic chemistry, 'chemical evolution' refers to an evolution-like process operating on prebiotic molecules, with definitions varying strongly from author to author [2]. Many authors have speculated on what such an evolution may look like and how similar it must be to modern Darwinian evolution. Unfortunately, no satisfactory description of chemical evolution has been proposed, and some argue that a pre-Darwinian evolution process is implausible[3]: .. *the concept of "pre-Darwinian evolution" appears questionable, in particular because it is unlikely if not impossible that any evolution in complexity over time may work without multiplication and heritability allowing the emergence of genetically and ecologically diverse lineages on which natural selection may operate.*.

In this section, we first discuss some particular attempts to address this question, notably GARD[4, 5, 6], autocatalytic sets of polymers[7, 8, 9, 10] and an evolving autocatalytic metabolism in a single reactor [11]. A common feature of these models is autocatalysis, but they all use the phenonemon in a different way. These models insist on particular chemical systems and network structures, which provides an elegant approach for studying a particular mechanism.

In the formulation of scenarios, however, we should reconsider how closely we are tied to the restrictions of these models. These restrictions are first and foremost a means of making the model practical and tractable. Such considerations provide little justification to restrict a prebiotic scenario to one type of chemistry or autocatalytic mechanism.

As was shown in the last chapter, we have not even considered all types of autocatalysis yet. In addition, there is considerable confusion on what we should call autocatalysis[12, 13]. Here, we will address these issues, by showing how autocatalytic mechanisms in the literature fit in the general picture of stoichiometric autocatalysis. By uniting them in a common language and framework, we hope to pave the way for a more united treatment of the phenomena that some refer

to as ‘chemical evolution’.

The need for such a treatment is echoed by a clear impasse in OOL: the apparent implausibility of kickstarting a form of genetic evolution directly is an important motivation to look for forms of chemical evolution. Simultaneously, no other autocatalytic evolution mechanism convincingly provides enough evolutionary capacity by itself[14, 15, 16, 8, 3].

It is hoped that the present discussion may provide a fresh starting point in an overall effort towards a synthesis of chemical evolution. A first extension to this effort will be considered in Chapter 8, where a small number of multilevel selection mechanisms (as found in GARD and the stochastic corrector) is discussed and analyzed in detail.

Starting from a stoichiometric matrix with nonambiguous elementary reactions, we study the conditions for the nucleation and survival of new autocatalytic cycles in terms of microscopic rates. In doing so, we obtain a more rigorous description of the evolving autocatalytic metabolism, of which GARD and autocatalytic sets are particular examples. Subsequently, we introduce multiple environments and molecular transport processes. For all of this, we can use the framework of stoichiometric autocatalysis introduced in the last chapter.

In doing so, autocatalytic multicompartment cycles radically increase evolvability: we must now not consider a single evolving reactor doing pure chemistry, but an extensive ecosystem of chemically distinct microenvironments performing selective exchange. This may provide the combinatorial diversity to enforce further evolvability, while simultaneously keeping the local chemistries sufficiently clean.

The viability of multicompartment autocatalysis hinges on efficient and selective exchange between different compartments. It becomes advantageous for these compartments to be close, to be numerous, and even to be connected.

6.1 Chemical Evolution

As a concept, a notion of ‘chemical evolution’ has emerged a number of times over the last 100 years. C. Malaterre observes that the concept really started to gain traction[2] after its discussion by Melvin Calvin in “chemical evolution and the Origin of Life”[17]. Malaterre distinguishes two academic approaches that have been taken to chemical evolution: a ‘historical’ descriptive approach and theoretical attempts to truly define the process. Indeed, chemical evolution is the staple deux-ex-machina that is invoked to bring prebiotic scenarios to their biotic conclusion.

Chemical evolution has been defined many times over to fit particular scenarios or models. At present, there is no universally accepted definition. However, the different attempts at discussing such a process all revolve around molecules achieving increasingly elaborate states of organization in an evolution-like manner.

In the following, some models for chemical evolution will be discussed. While they are all quite different in context (focusing on particular types of molecules, networks and reactor settings) and somewhat limited in their scope, they show a key similarity: every single one of them performs stoichiometric autocatalysis.

One open question that is currently the object of active research[18] is to see how the cross-catalytic growth in GARD and autocatalysis in autocatalytic sets are related. As we will show, they are both examples of coarse-grained reaction networks, and imposing nonambiguity shows that they are both autocatalytic in the same stoichiometric sense.

6.1.1 Autocatalytic sets

Autocatalytic sets were pioneered by S. Kauffman in 1971[19]. At the time, responses were mixed: a remark by a disinterested chemist led him to abandon the line of research for over a decade[18]. Later, the work was picked up again, and was extended along several lines, notably by Kauffman[7],

W. Hordijk and M. Steel[9, 10] and has notably been explored in models for polymerization by Giri et al [20] and gene-free evolution by Vasas et al. [8].

Hordijk and Steel placed the autocatalytic sets framework on more rigorous mathematical footing and provide the following definition for what constitutes an autocatalytic set[9, 10, 12], which we will repeat here:

Definition 6.1.1 — Autocatalytic set (RAF set). An autocatalytic set (or RAF set) is a set of reactions $r \in \mathscr{R}$, that is:
1. F-generated (F): all reactants consumed by reactions $r \in \mathscr{R}$ are part of the food set f or can be made directly from it, using only reactions in $\mathscr{R}$.
2. Reflexively autocatalytic (RA): each reaction $r \in \mathscr{R}$ is catalyzed by at least one species, that is part of the food set f or can be made directly from it using only reactions in $\mathscr{R}$.

Note that not every reaction network that exhibits autocatalysis is an autocatalytic set. As pointed out by W. Hordijk, Toy Formose is not an autocatalytic set, because not every reaction is catalyzed[12], it is only an autocatalytic cycle. The converse, however, is true: every autocatalytic set is a reaction network that exhibits autocatalysis (or at least stoichiometric autocatalysis).

Our stoichiometric matrix framework differs from the RAF framework, by having only reactions, from which catalysis follows by their network motif. In RAF sets there is an explicit distinction between reactions and catalysis. As noted in Sec. 2.2, we can always decompose a reaction in more steps, and catalysis itself must occur in cycles composed of two or more steps. Catalysis will not act on every substep of this cycle (e.g. the diffusive approach of reactants or various steps in the reaction mechanism), so such decompositions are best avoided in RAF sets to satisfy the RA-condition.

Coarse-grained networks in autocatalytic sets

In terms of dynamics, a catalyst in a RAF set should accelerate the rate-limiting step in a series of bundled reaction steps bundled in a single reaction $r \in \mathscr{R}$. Kauffman made an observation of a similar kind [18], noting that some biochemical reactions (in E. Coli the number seems to be 3) are entirely uncatalyzed, they are rapid enough on their own. The appropriate choice of reactions and species to model a system under such coarse-graining can depend on the system composition, at certain concentrations other substeps may become limiting.

This coarse-graining must be done with care. If we were to fuse the 3 reactions in the Toy Formose model (Fig. 6.1), one ends up with a single net reaction

$$2\mathrm{C}_1 \xrightarrow{C_2} \mathrm{C}_2, \tag{6.1}$$

which is catalyzed by C_2.

In this single-reaction model, Toy Formose is a RAF. For real formose, the single-reaction approximation can be quite appropriate for the $\mathrm{C}_2 \longrightarrow \mathrm{C}_3 \longrightarrow \mathrm{C}_4 \longrightarrow 2\mathrm{C}_2$ pathway, when fragmentation of C_4 is the rate-limiting step. An increase in C_2 will rapidly lead to a corresponding increase in C_4, and hence increase the rate. To have this as a rate-limiting step, C_1, divalent metal ions and base need to be sufficiently abundant. Hordijk notes that real formose does not form more metal ions and that this prevents it from being a RAF.

Strictly speaking, this situation can be circumvented, if we allow the divalent metal ions to be part of the food set f. Such an intervention would not alter the definition of the RAF set and seems to be allowed within the framework. At any rate, Toy Formose does not suffer from this objection (since it is purely hypothetical) and it can have a representation in which it is a RAF and one where it is not. The separation of timescales dictates whether calling it a RAF is appropriate.

As shown in Fig. 6.1, whether an autocatalytic network (here: Toy Formose) is a RAF critically depends on the level of description. By decomposing Toy Formose in smaller reaction steps, we

move from an overall autocatalytic cycle to a mechanistic description of its substeps (A more realistic description of genuine formose would require even more more substeps, involving hydride shifts, deprotonation by catalytic base, etc.[21, 22]).

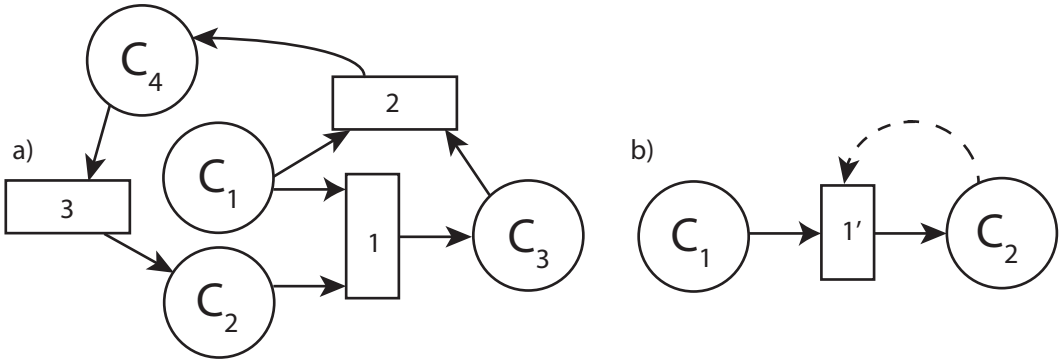


Figure 6.1: Two representations of the formose network using bipartite graphs. Dotted arrows signify catalysis, bold arrows map reactants to a reaction (stoichiometry not shown) a) a detailed, three-reaction Toy Formose. No reaction is catalyzed, not a RAF. b) a coarse-grained, single-reaction Toy Formose. The product catalyzes its own formation and forms a RAF.

Although we can describe the smallest autocatalytic cycle in Toy Formose as a RAF, the more general statement that formose is not an example of a RAF is also correct. To see this, we need to consider that in true formose there are more autocatalytic cycles, e.g. via C5, which can split according to $C_5 \longrightarrow C_2 + C_3$, and C6 via $C_6 \longrightarrow C_2 + C_4$ and $C_6 \longrightarrow C_3 + C_3$. To convert $C_2 \longrightarrow C_3 \longrightarrow C_4 \longrightarrow 2C_2$ to a RAF, we had to remove C3 and C4. To describe higher-order cycles as RAFs, we need these species in our description.

Abundance and polymer scenarios

An important question in RAF theory and for autocatalysis in general concerns the abundance of autocatalytic networks. Kauffman recalls that some of his key results for autocatalytic sets was inspired[18] by a Erdös-Rényi graphs[23]. These are random graphs, constructed by randomly connecting pairs of nodes with an independent probability p for each edge. As such graphs grow in size, the probability that they contain a cyclic subgraph grows rapidly.

By making random graphs for catalysts promoting the formation of other catalysts, Kauffman showed that large enough networks of catalysts and food molecules are expected to contain autocatalytic sets. The value of p is not expected to be very large: if we mix some arbitrary catalysts we should indeed be very surprised to see them catalyze each other's formation. Kauffman argues for values of p around $10^{-5} - 10^{-6}$. It is then necessary to have a high number ($O(1/p)$) of nodes (catalysts) to achieve appreciable connectivity.

Such a high number of nodes implies a highly diverse chemical mixture. In practice, scenarios with autocatalytic sets achieve such diversity by being formulated in terms of cross-catalytic networks of copolymers [7]. Some work has also been done on random graphs with links that show some extra structure, e.g. due to a complementarity rule in RNA, which has led to similar results[24]. A recent formulation of the model considers placing such autocatalytic sets in dividing coacervate compartments [8].

6.1.2 GARD

GARD stands for graded autocatalysis replication domain, and was originally introduced by Segré et al in 1998[4]. It is commonly presented as a model for amphiphile assemblies (e.g. micelles)

with a composition $\boldsymbol{n} = \{n_1, ..., n_s\}$, that follow an evolution equation

$$\frac{dn_i}{dt} = \left(k_i^+ \rho_i N - k_i^- n_i\right)\left(1 + \frac{1}{N}\sum_{j=1}^{s} \beta_{ij} n_j\right). \tag{6.2}$$

The surfaces are in contact with a reservoir, that contains species $\mathrm{Z_i}$ at concentration ρ_i and that can enter the surface, which has an area proportional to N. The incorporation happens with a base rate of k_i^+, but can be facilitated by other amphiphiles, for which the catalytic rate enhancement is characterized by β_{ij}.

A special ingredient in GARD is the division process, which splits a mature surface in two new ones, after achieving a maximal size N_{max}. A typical order of magnitude used in simulations is $N_{max} = 100$, which is a typical order of magnitude encountered for micelles. The new compartments are roughly equally sized partitions of the original mixture. The cross-catalytic network structure can introduce multiple (stable) attractors for the system composition in the growth-division process. Such a compositional attractor has been called a 'composome', or a compositional genome.

Composomes and Information

In principle, there are

$$\Omega = \binom{N+s-1}{N} \tag{6.3}$$

possible compositions for a size N. On several occasions, $\ln\Omega$ has been considered as a measure of the information in the composition. In the sense defined in Chapter 3, this can be given a thermodynamic interpretation: we can take a highly specific composition $\boldsymbol{n}$ and extract work from randomizing that composition with an appropriate engine. If all these compositions have the same free energy, a process can be constructed that yields $kT\ln\Omega$.

This also makes clear what the limits are of such an interpretation: the micelle composition $\boldsymbol{n}$ is strongly fluctuating and a stable composome state cannot be defined with single-molecule precision. Indeed, a composome is characterized by a compositional attractor and its vicinity[6]. Let us consider a protocol (analogous to Sec. 4.2) that extracts work by reversibly letting one stable composome interconvert between all possible N_{comp} stable composomes. If these composomes occupy the same volumes of phase space, it directly follows that we can extract a work of at most $kT\ln N_{comp}$.

In the GARD scenario, we do not aim to extract the maximum work from a composition, we aim to produce a set of distinct stable states. We see, however, that these ideas can be made analogous. Both in the thermodynamic sense and in the evolutionary sense, the natural quantity to study is the number of attractors N_{comp}. The number of such attractors is much smaller[16, 6] than Ω. In this sense, the analogy between the number of copolymer sequences and Ω is less appropriate, since copolymer sequences do not exhibit the inherent fluctuations that GARD assemblies do*.

Provided the compositional noise due to division and growth is sufficiently small, certain compositions can be stabilized. For $N_{max} = 100$, noise in inoculation and incorporation is expected to be considerable, especially with a large diversity of amphiphiles. In addition, if incorporation of amphiphiles is a memoryless process, it is expected to yield high noise, unless considerable differences in rate enhancement β_{ij} are in place to compensate for it (see also Sec. 8.4, for a discussion of noise in growth).

This has formed the object of some criticism: the inherently high noise (especially for $s \gg 1$) makes compositions unstable, hampering particular compositions to act on timescales deemed

*Letting $\boldsymbol{m}$ denote the number of monomer types and L the polymer length, we can in principle construct an engine that extracts a work equal to $kT\ln m^L$ per polymer, by interconverting between all other possible sequences.

necessary for evolutionary processes[25]. If in addition only a small number of attractors exists, these states will be rapidly explored and a stable steady-state. The latter was shown explicitly by the Perron-Frobenius theorem on the level of populations of attractor states. Lancet et al. [6] argue that other objects such as vesicles could be imagined, with much larger values of $N_{max} > 10^6$ (indeed, this is considered as an alternative path towards protocells[3]). Such a regime makes it very computationally intensive to study the model numerically, if the same scheme is adopted as used for $N_{max} = 100$.

Autocatalysis in GARD

In GARD, the parameters β_{ij} are phenomenological parameters informed by datasets on lipid systems, where *surface area, charge, ability to form complexes with neighboring molecules and intrinsic curvature*[6] were used to infer typical values for β_{ij}.

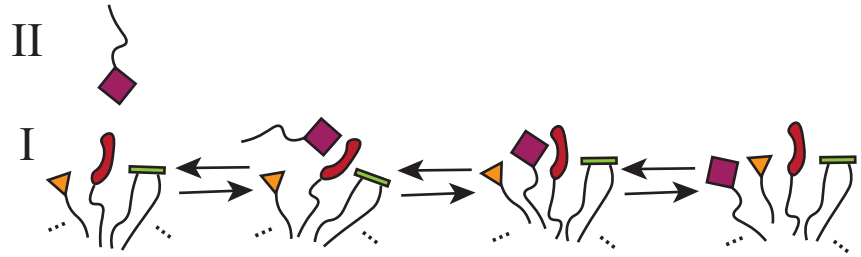


Figure 6.2: A catalytic incorporation mechanism: the red telephone amphiphile forms a complex with the purple square amphiphile, which mediates its incorporation in a micelle or vesicle.

A simple stoichiometric mechanism leading exactly to Eq. (6.2) can be found by considering what can give rise to a linear contribution for cross-catalysis. Let us consider amphiphiles A_1 and A_2, which can be in the micelle or reservoir, labelled I and II (see Fig. 6.2). To catalyze incorporation of the other, a complex is formed with a reservoir species, and subsequent dissociation takes place in the micelle

$$\mathrm{A^I}_1 + \mathrm{A^{II}}_2 \rightleftharpoons [\mathrm{A_1A_2}]^I \rightleftharpoons \mathrm{A^I}_1 + \mathrm{A^I}_2, \tag{6.4}$$

$$\mathrm{A^I}_2 + \mathrm{A^{II}}_1 \rightleftharpoons [\mathrm{A_2A_1}]^I \rightleftharpoons \mathrm{A^I}_1 + \mathrm{A^I}_2. \tag{6.5}$$

A simple model for this was proposed in Sec. 5.3.5, and a corresponding graph for GARD is shown here in Fig. 6.3. As shown in Sec. 5.3.5, removing the reservoir amphiphiles from the

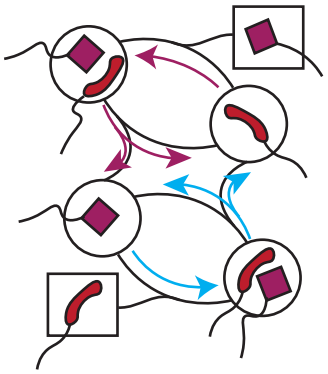


Figure 6.3: An autocatalytic network of co-assembling amphiphiles. Amphiphiles in square nodes are reservoir species, those in circular nodes represent amphiphiles in a micelle or membrane.

stoichiometric matrix $\boldsymbol{\nu}$ yields an invertible and autonomous matrix that admits SFA.

A question that has been raised is to which degree RAF-sets and GARD are connected[18, 6]. As outlined before, being a RAF or not depends critically on the level of description. Let us coarse-grain the exchange reactions, by removing the species $[A_1A_2]^I$ and $[A_2A_1]^I$:

$$A^I{}_1 + A^{II}{}_2 \rightleftharpoons A^I{}_1 + A^I{}_2, \tag{6.6}$$
$$A^I{}_2 + A^{II}{}_1 \rightleftharpoons A^I{}_1 + A^I{}_2, \tag{6.7}$$

which in alternative notation is written as

$$A^{II}{}_2 \overset{A^I{}_1}{\rightleftharpoons} A^I{}_2, \tag{6.8}$$
$$A^{II}{}_1 \overset{A^I{}_2}{\rightleftharpoons} A^I{}_1, \tag{6.9}$$

with the superscript indicating catalysis. Upon performing this coarse-graining operation, this minimal example turns into a RAF-set.

By proposing an alternative to information in copolymer sequences, GARD is often associated with 'metabolism first' and 'lipids first' scenarios. As a model, Eq. (6.2) can perfectly well apply to other molecular collectives, as was shown for e.g. amino acids [4]. Experimentally, composomes of cross-catalytic RNA have been realized by Vaidya et al[26]. This experimental system was fed with fresh solution containing RNA fragments, buffer and Mg^{2+}, the latter of which is indispensable for the proper functioning of the ribozymes formed. In a description without Mg^{2+} (whose complexation and release is typically rapid compared to the other reactions, so such a coarse-graining can be justified), this system can be shown to form a RAF set.

6.1.3 Evolving Metabolism: sequentially decorated autocatalysis

A more general, abstract approach to chemical evolution, was the consideration of some arbitrary chemistry, being sequentially decorated, through autocatalysis. Such an idea was proposed in the works of King [27], where, by a deduction through twelve progressive steps, conditions were outlined to have autocatalytic pathways.

Similar to our framework, King imposes reactions to be at most bimolecular. The networks King considers are formulated in terms of 'recycling' species, fed by reactants that are not part of the description (which would be the food set f in RAF-theory and chemostatted species in our framework). A production of 'waste' molecules was not considered, but they can be trivially removed by chemostatting as well.

Subsequently, King reasons in terms of the number of reactants consumed and products produced by individual reactions acting on recycling species. Starting from a branching reaction (converting one species to two), the trajectories of the two products back to that reaction can be traced. King argues that such a pathway will either lead to i) bimolecular rejoining and hence 0 net molecules ii) unimolecular rejoining, and hence 1 net new molecule.

King also considered the specificity of reactions and their effect on the final viability of the cycle. This is an important consideration that should receive our full attention: chemical reactions can provide autocatalysis, but they can also break it. King exemplifies this idea via a cyclic autocatalytic reaction, reminiscent of an n-step Toy Formose:

$$Z_1 \xrightarrow{1} \ldots \xrightarrow{n-1} Z_n \xrightarrow{n} 2Z_1. \tag{6.10}$$

A successfull cycle converts Z_1 to 2 Z_1. If any intermediate Z_k ($1 \leq k \leq n$) is broken down prematurely by a side-reaction, then a net conversion Z_1 to $\emptyset$ is performed. A net accumulation requires that on average, a Z_1 molecule yields more than one Z_1 molecule, which requires that cycles are successful at least half of the time.

King notes that if in every step i a fraction ζ_i is lost due to side reactions, the product of survival probabilities $(1-\zeta_i)$ must then verify

$$\prod_{i=1}^{n}(1-\zeta_i) > \frac{1}{2}, \tag{6.11}$$

for which King argued that, regardless of the distribution of selectivities, a smaller n is generally more like to yield viable autocatalysis.

Similar in spirit is a work by Bagley, Farmer and Fontana [11], which starts from an autocatalytic network where compositional fluctuations can trigger new autocatalytic reactions. The rare species that are not part of the network form 'the shadow'. Reactions in the shadow are represented by simple graphs and concern only 'rare' species, the 'parent' species (food) supplied by the autocatalytic network remain implicit.

By using simple graphs, the possible autocatalytic networks is considerably reduced, and replication becomes implicit. It has the advantage of leading to elegant reactions of the form Eq. (6.10). Together, these reactions model a birth-death process, for which a master equation was written, and solved for $t \to \infty$.

In turn, finding the extinction probability provided an efficient means of simulating 'chemical evolution'. When an autocatalytic molecule is generated, a random number generator indicates whether the cycle goes extinct or not, based on the extinction probability. It is then not necessary to run a stochastic simulation for the whole fixation process.

Another instructive model was proposed[28] and further developed[29, 30] by Jain and Krishna and revolved around a collective of species interacting through catalysis, which is represented by a graph. The evolution of such graphs involved the removal of the least fit node, followed by the introduction of a new node, which is randomly connected to the existing network.

6.1.4 A common thread: Autocatalysis

In the struggle for existence, effective reproduction is essential. For chemical networks, the same can be said, as the components that shape a successful chemistry must be maintained in the face of degradation and dilution. Moreover, for the network to spread to new places, its components must be multiplied beyond the numbers required for maintenance.

Based on the general definition for autocatalysis from stoichiometry, as introduced in the last section, this reproduction (if it can be captured in stoichiometric terms) inherently requires autocatalysis: the characteristic that a subset of chemicals converts food to make new chemicals, among which themselves, implies autocatalysis.

An intuitive way to see this, is to consider the following thought experiment, illustrated in Fig. 6.4. A CSTR reactor receives a constant influx of a certain chemical mixture, and the composition has reached a steady-state. Now, the mixture is perturbed by the occurrence of a rare molecule Z, which can be an impurity in the feed or a rare reaction product. Let us now wonder what can happen, given a residence time of τ. If i) Z is inert, the steady state does not change, and after τ, Z has left. If ii) Z reacts with other reactants in the mixture, it is converted to new compounds, which are each lost on a typical timescale τ. If iii) Z is (or is converted to) a catalyst that acts on the species present in the CSTR, it can introduce a considerable amount of new species during a time τ, after which it disappears. The species that were produced could in turn be any combination of i), ii), iii), but will disappear after τ. However, when any of these species is a copy of Z, hence, when Z is an autocatalyst (in the sense defined in Ch. 5), Z and its derivates can persist. Autocatalysis can alter the steady-state of the reactor by a small chemical perturbation, whereas other network patterns cannot.

This thought experiment is closely related to the evolving metabolism model of Bagley et al[11], where chemical evolution was envisaged as occurring in a large autocatalytic protometabolism,

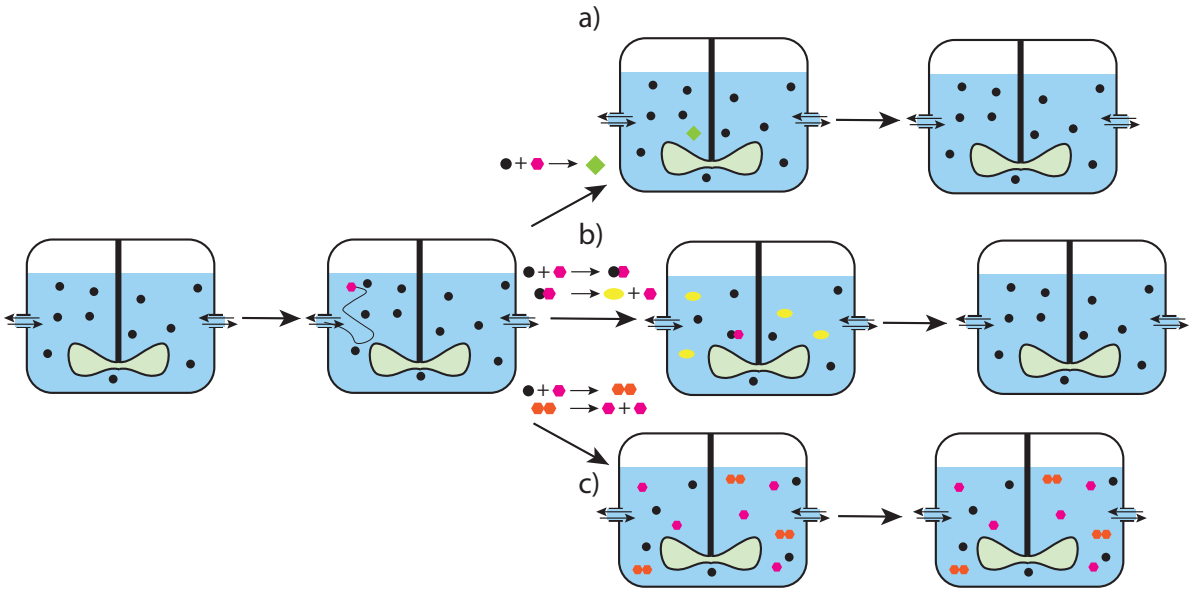


Figure 6.4: A CSTR responding to a chemical perturbation: the introduction of a rare molecule (pink hexagon). a) the rare molecule is consumed, forming a green square. b) the rare molecule performs allocatalysis, converting black circles to yellow ellipsoids. c) The rare molecule performs autocatalysis. a) and b) cannot persist in the face of reactor outflow and over time the CSTR reverts to the its original state. Autocatalysis however (c), can compensate for this degradation and be maintained.

which was sequentially decorated by new autocatalytic reactions.

6.2 Section: single-pot autocatalytic evolution

In this section, we will derive 'chemical evolution' in a single reactor, analogous to the model by Bagley et al[11], from the stoichiometric matrix. By considering how single-molecule perturbations can be amplified, we derive the autocatalytic network structures that can be triggered.

In elaborating this model, we get at the heart of the evolvability issue: Does chemistry itself have sufficient autocatalytic potential to 'evolve' in such a homogeneous single-reactor context? Part of the answer will come by looking at the requirements on the stoichiometric matrix for nucleating further autocatalysis, which we make explicit here.

6.2.1 Reactor setup

We start with a well-stirred CSTR reactor at steady-state, recieving a constant influx of chemical species denoted by a set $\boldsymbol{Y} = \{\mathrm{Y}_1, ..., \mathrm{Y}_{\mathrm{s_Y}}\}$. These species can react, and the species that are not flown in but are present in appreciable amount are called internal species $\boldsymbol{X} = \{\mathrm{X}_1, ..., \mathrm{X}_{\mathrm{s_X}}\}$. Furthermore, we consider a list of species $\boldsymbol{W} = \{\mathrm{W}_1, ..., \mathrm{W}_{\mathrm{s_W}}\}$ that, on average, are absent, but that may occasionally be present due to a rare reaction or impurities in the feed. For most W_k, we can then write:

$$N_{W_k} = 0, \tag{6.12}$$

and for any particular W_j that is present as a perturbation at a time t, we will consider

$$N_{W_j} = O(1). \tag{6.13}$$

Most W-species exist transiently. They are then converted back to an X or Y species, or will be inert. Some components, however, may be part of an autocatalytic network that can be nucleated by a single molecule. If such a network successfully increases the abundance of W and other members of the autocatalytic cycle, they cease to be rare. We can now extend the vector of X-species[†] with the vector $\boldsymbol{w}$:

$$\boldsymbol{X}' = \{\boldsymbol{X}, \boldsymbol{w}\}. \tag{6.14}$$

The components $\boldsymbol{w} \in \boldsymbol{W}$ are removed from $\boldsymbol{W}$. The new reaction network may in turn form a number of new rare species, described by the set $\boldsymbol{W}^+$, such that

$$\boldsymbol{W}' = \{\boldsymbol{W}/\boldsymbol{w}, \boldsymbol{W}^+\}. \tag{6.15}$$

If we suppose such events to be sufficiently far and few between in time, we can describe these events as going from steady-state to steady state (provided the system does indeed relax to a steady state).

Stoichiometric matrix decomposition

We can decompose the stoichiometric matrix $\boldsymbol{\nu}$ in three submatrices

$$\boldsymbol{\nu} = \begin{pmatrix} \boldsymbol{\nu}^Y \\ \boldsymbol{\nu}^X \\ \boldsymbol{\nu}^W \end{pmatrix}, \tag{6.16}$$

where the splitting allows to distinguish between 'food' ($\boldsymbol{Y}$) internal chemistry X, and 'rare chemistry' ($\boldsymbol{W}$). Macroscopic phenomena, entropy production and the steady state follow from reactions in $\boldsymbol{\nu}^Y$ and $\boldsymbol{\nu}^X$. In this interpretation, 'Chemical evolution' is the incorporation of rare species $\boldsymbol{W}$ in $\boldsymbol{X}$.

A chemical reaction can involve species from different sets, e.g.

$$\mathrm{X_i} + \mathrm{W_j} \rightleftharpoons \mathrm{W_k}. \tag{6.17}$$

However, such reactions should be in accord with the fact that a species W_k is typically absent. The following reaction

$$\mathrm{X_i} + \mathrm{Y_j} \rightleftharpoons \mathrm{W_k}, \tag{6.18}$$

yields a rare species, which is only consistent if the reaction is sufficiently disfavored (e.g. due to a high barrier or low concentrations), so that the species indeed remains rare. Let us now look at nucleation of autocatalytic cycles in $\boldsymbol{\nu}^W$.

Reactions in the space of rare species

Let us start by interpreting various types of reactions we can encounter in $\boldsymbol{\nu}^W$.

First of all, there can be $\emptyset$-reactions, that arise from the decomposition. Denoting by $\boldsymbol{Z} = \{\boldsymbol{X}, \boldsymbol{Y}\}$, these reactions are of the form

$$\sum_i \nu_{ij}^Z Z_i \rightleftharpoons \sum_i \nu_{ij}^W W_i. \tag{6.19}$$

Taking only the perspective of $\boldsymbol{W}$, we then have

$$\emptyset \rightleftharpoons \sum_i \nu_{ij}^W W_i. \tag{6.20}$$

[†]In principle, this modification may also make certain species in X disappear, but this is a lot more difficult and will not be pursued.

These reactions yield rare species in $\boldsymbol{W}$ from nonrare species in $\boldsymbol{Z}$. This is consistent with Eq. (6.13) if the reaction is sufficiently disfavored and thus rarely occurs in practice (e.g. due to a high barrier or low concentrations). Such $\emptyset$-reactions can be considered degradation pathways for a species $\boldsymbol{W}$, and performing them in reverse provides the occasional W_k.

Subsequently, there are effective unimolecular reactions. Unlike unimolecular reactions in $\boldsymbol{\nu}$, such reactions may involve species in $\boldsymbol{Z}$ as reaction partners, which gives three possibilities for unimolecular reactions

$$\mathrm{W_h} \rightleftharpoons \mathrm{W_j}, \tag{6.21}$$

$$\mathrm{W_h} + \mathrm{Z_i} \rightleftharpoons \mathrm{W_j}, \tag{6.22}$$

$$\mathrm{W_h} + \mathrm{Z_i} \rightleftharpoons \mathrm{W_j} + \mathrm{Z_k}, \tag{6.23}$$

which within $\boldsymbol{\nu}^W$ reduces to Eq. (6.21).

Finally, there are bimolecular reactions. Let us first consider effective branching reactions, which are

$$\mathrm{W_h} \rightleftharpoons \mathrm{W_j} + \mathrm{W_k}, \qquad \mathrm{W_h} + \mathrm{Z_i} \rightleftharpoons \mathrm{W_j} + \mathrm{W_k}, \tag{6.24}$$

which within $\boldsymbol{\nu}^W$ reduces to Eq. (6.25). Note that also this includes the case where $j = k$:

$$\mathrm{W_h} \rightleftharpoons 2\mathrm{W_j}. \tag{6.25}$$

Since species in $\boldsymbol{W}$ are absent on average, the reverse reaction is much less probable, thereby strongly favoring such reaction to go forward.

In a thermodynamic sense, this can be readily captured by

$$\Delta G = \Delta\mu^\circ + kT \ln \frac{(N_{\mathrm{W_j}} + 1)(N_{\mathrm{W_k}} + 1)}{N_T (N_{\mathrm{W_h}} - 1)}, \quad \Delta G = \Delta\mu^\circ + kT \ln \frac{(N_{\mathrm{W_j}} +)(N_{\mathrm{W_k}} + 1)}{N_{\mathrm{Z_i}} (N_{\mathrm{W_h}} - 1)}, \tag{6.26}$$

which for the former becomes $kT \ln N_T$, with N_T the total number of molecules in the system. In the latter case, we find $kT \ln N_{Z_i}$, which only makes the reaction reversible in the regime where N_{Z_i} is sufficiently small.

We can thus consider these reactions as irreversible

$$\mathrm{W_h} \longrightarrow \mathrm{W_j} + \mathrm{W_k}. \tag{6.27}$$

Finally, there can be reactions of the form

$$\mathrm{W_h} + \mathrm{W_i} \rightleftharpoons \mathrm{W_j} + \mathrm{W_k}. \tag{6.28}$$

Due to the rarity of species in $\boldsymbol{W}$, these reactions are kinetically improbable.

Autocatalytic Network Construction

Due to the low abundance of species in $\boldsymbol{W}$, the building blocks for an autocatalytic network are strongly reduced. Their composition is limited to unimolecular forward reactions as given by Eqs. (6.21)-(6.23). This means that the networks we will construct will correspond to solitary autocatalysis (see Ch.5).

In terms of the stoichiometric submatrix for $\boldsymbol{W}$, Eqs. (6.21)-(6.23) imply

$$\nu^W_{ij} = -\delta_{s,i} + \delta_{p,i}, \tag{6.29}$$

where s is the index of a substrate and p the index of a product, such that for reaction j, we have

$$W_s \rightleftharpoons W_p. \tag{6.30}$$

Furthermore, we have irreversible branching reactions, given by Eq. 6.25, which in the stoichiometric matrix take the form

$$\nu_{ij}^{W} = -\delta_{s,i} + \delta_{p,i} + \delta_{p',i}. \tag{6.31}$$

where p can be equal to p'.

Any sequence of unimolecular reactions will preserve the exact number of molecules. Such reactions have a mass-like conservation law $\boldsymbol{l}^{+}$. To establish autocatalysis in W, we therefore need at least one irreversible branching reaction. The simplest networks obtained by having a single braching step are shown in Fig. 6.5.

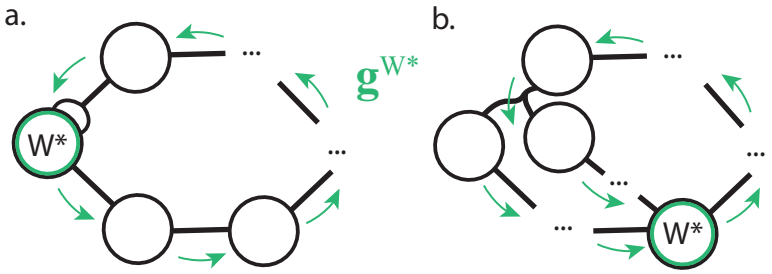


Figure 6.5: The two simplest autocatalytic network motifs containing only rare species, obtained either by a) producing 2 W* at the end of the cycle, or b) by producing 2 different products upon branching, which at some point are converted to the same species W*. The replication cycle $\boldsymbol{g}^{W*}$ is drawn for both networks.

6.3 Conditions for survival of autocatalytic networks

Whether an autocatalytic reaction can be maintained after the introduction of the first autocatalyst(s) (as in Fig. 6.4) is a statistical question involving kinetic competition between degradation and reproduction. In this section, this problem will be treated explicitly for some autocatalytic network motifs. From microscopic rates, the probability to finish (parts of) autocatalytic cycle along the graph can be found. From these probabilities a survival (fixation) probability can be found, through the appropriate mapping to a stochastic processes. For a single cyclic path (Sec. 6.3), this can be mapped to a birth-death process, which was also treated by Bagley et al.[11]. In Sec. 6.3 we show that more general cases can be mapped to a branching process. The two cases treated here (Fig. 6.5) are only the simplest cases. The branching process approach is generally applicable, however, as shown in Sec. 6.5.

A single cyclic path: path probability

Let us consider the simplest autocatalytic networks, which have one branching reaction, by starting with the class of networks given by Fig. 6.5a. When a first molecule is introduced in such a network, e.g. A_n in Fig. 6.6, it will need to go through intermediates A_1 to A_{n-1} before finally producing two copies of A_n. This also happens in Toy Formose:

$$C_2 \longrightarrow C_3 \longrightarrow C_4 \longrightarrow 2C_2. \tag{6.32}$$

Provided the reaction and degradation steps are memoryless[‡], such a process can be modeled as a Markov process. The fragmentation reaction $A_{n-1} \longrightarrow 2A_n$ is irreversible, as are all degradation processes. In principle, other steps can be reversible. For the nucleation of the (Toy) formose

[‡]Which we strive to achieve by making steps elementary and nonambiguous

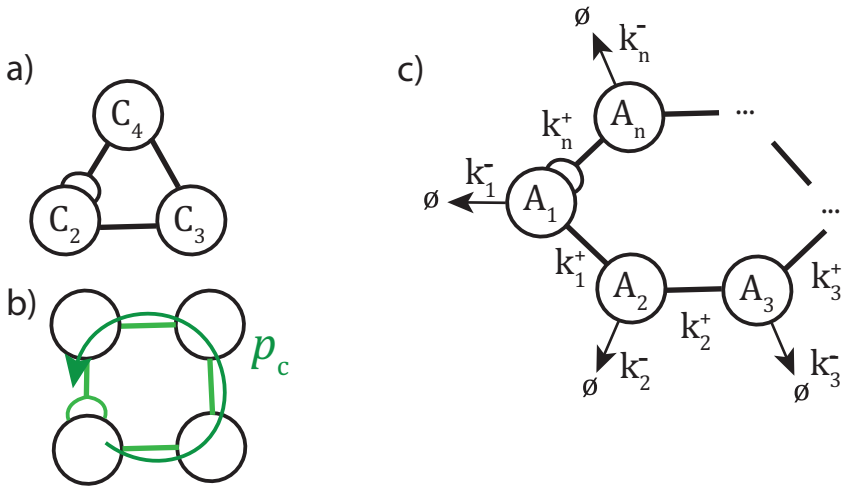


Figure 6.6: a) a three-membered forked branching network, illustrated by the Toy Formose reaction. b) Illustration of the replication path, which is successful with probability p_c, a central quantity in the birth-death process (see Fig. 6.7) c) General representation of a Forked Branching network with n members. Effective rate constants have been added for degradation and irreversible progression.

reaction, reactions can initially be considered irreversible to a good approximation (but with $\Delta\mu^\circ > 0$ for the fragmentation).

Let us now study the case where all steps are irreversible. At any intermediate A_s, one can transition to A_{s+1} or $\emptyset$. Let p_s^+ be the relative probability of successfully transitioning to A_{s+1}, and p_s^- for degradation, such that $1 = p_s^+ + p_s^-$. The relative probabilities of the two outcomes follow from effective rate constants

$$p_s^+ = \frac{k_s^+}{k_s^+ + k_s^-}, \tag{6.33}$$

$$p_s^- = \frac{k_s^-}{k_s^+ + k_s^-}. \tag{6.34}$$

The probability to successfully finish a cycle, p_c, is the probability of going through all steps without degradation

$$p_c = \prod_{s=1}^{n} p_s^+ = \prod_{s=1}^{n} \frac{k_s^+}{k_s^+ + k_s^-}. \tag{6.35}$$

For more elaborate network structures and reversible reactions, the expression for p_c becomes more complicated, but can still be expressed in terms of microscopic rates.

A single cyclic path: birth-death process

Let us now focus on the particular case where we start with a single A_1 autocatalyst. With probability p_c, the autocatalyst produces a copy of itself, providing two species starting again in state A_1, we will designate this transition from a state A_1 to $2A_1$ by a process

$$A_1 \xrightarrow{p_c} 2A_1 \tag{6.36}$$

Alternatively, A_1 or some intermediate A_k is degraded along the way (with probability $1 - p_c$). We will write the combination of all these processes as a net transition to a state $\emptyset$, through a process

$$A_1 \xrightarrow{1-p_c} \emptyset \tag{6.37}$$

The overall process can be then be represented as

$$\emptyset \xleftarrow{1-p_c} \mathrm{A}_1 \xrightarrow{p_c} 2\mathrm{A}_1. \tag{6.38}$$

We will refer to process (6.36) as a birth, and eqref (6.37) as a death.

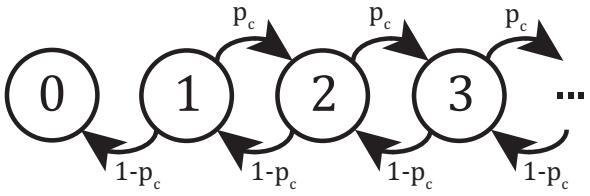


Figure 6.7: Survival of the network follows a 1D random walk or birth-death process, with birth probability p_c, corresponding to successful replication as shown in Fig. 6.6b. Nodes count the number of species N_m.

With every birth, the autocatalyst population increases by 1, with every death, it decreases by 1. We can now write a birth-death process[31] (see also Fig. 6.7)

$$N_m = N_{m-1} + X_{m-1} - 1, \tag{6.39}$$

Where N_m denotes the population size after m repetitions of the process, for an initial condition of only A_1 species $N_0 = N_{\mathrm{A}_1}$ (which will be further generalized afterwards). X_{m-1} is a random variable, and $X_{m-1} = 0$ with probability $1 - p_c$ and $X_{m-1} = 2$ with probability p_c.

Note that we are not modeling the actual kinetics, but exploiting a statistical property due to solitary autocatalysis: autocatalysts only react with the environment. This gives us independent trials for the process (6.38), which must give an extinction probability consistent with detailed stochastic kinetics, which we demonstrate in Sec. 6.5.

We start with $N_0 = 1$. Let us denote d_m as the total probability that the population has gone extinct, after m iterations. After one iteration, we either have $N_1 = 0$ or $N_1 = 2$, the latter of which independently go extinct with probability d_{m-1}^2. This then gives

$$d_m = (1 - p_c) + p_c d_{m-1}^2. \tag{6.40}$$

Since d_m increases monotonically and is bounded from above by 1, it must converge in the limit $m \to \infty$ to a limit we will call d

$$\lim_{m\to\infty} d_m = d \tag{6.41}$$

The extinction probability for $m \to \infty$ then becomes a quadratic equation in d, with solutions

$$d = 1 \ \vee \ d = \frac{1 - p_c}{p_c}. \tag{6.42}$$

We find the fixation probability as the probability to not go extinct $P_{fix} = 1 - d$, such that

$$P_{fix} = \begin{cases} 0, & p_c \leq \frac{1}{2}, \\ 2 - \frac{1}{p_c}, & p_c \geq \frac{1}{2}, \end{cases} \tag{6.43}$$

This result captures the survival in terms of microscopic rates, and was also found by Bagley et al.[11]. Extinction must becomes deterministic for $p_c < \frac{1}{2}$, where a population decays exponentially.

If we start from n A_1 species, the fixation probability P_{fix,nA_1} is the probability that not all of them go extinct: $P_{fix,n} = 1 - d^n$, hence

$$P_{fix,nA_1} = \begin{cases} 0, & p_c \leq \frac{1}{2}, \\ 1-(1-\frac{1}{p_c})^n, & p_c \geq \frac{1}{2}, \end{cases} \tag{6.44}$$

If we started from an initial state A_k instead of A_1, we can write a success rate $p_{c,k}$

$$p_{c,k} = \prod_{s=k}^{n} p_s^+ = \prod_{s=k}^{n} \frac{k_s^+}{k_s^+ + k_s^-}. \tag{6.45}$$

This headstart may increase the success rate of the first round: $p_{c,k} \geq p_{c,1} = p_c$, after which two A_1 are generated. Starting from a single A_k, we thus find a fixation probability

$$P_{fix,A_k} = p_{c,k} P_{fix,2A_1}. \tag{6.46}$$

Note that single-cycle network could have been more elaborate, e.g. with reversible reactions (see Sec. 6.5) and extra nodes that transiently deviate from the path. In such a case, we obtain a different expression for p_c. Its subsequent mapping to a birth-death process is unaltered, however. Let us now move to a situation that is described by a branching process.

Two cyclic paths: path probabilities

We will now study networks of the type described by Fig. 6.8a. Here the key fragmentation step produces two different autocatalysts, according to a reaction step

$$A_n \longrightarrow B_1 + C_1. \tag{6.47}$$

Through subsequent reaction steps, B_1 and C_1 can again become A_n

$$B_1 \quad \rightleftharpoons \quad ... \rightleftharpoons B_n \rightleftharpoons A_1 \rightleftharpoons ... \rightleftharpoons A_n, \tag{6.48}$$

$$C_1 \quad \rightleftharpoons \quad ... \rightleftharpoons C_n \rightleftharpoons A_1 \rightleftharpoons ... \rightleftharpoons A_n. \tag{6.49}$$

At some point, the path for B and C crosses at A_1. Fig. 6.8a shows simple example of this, encountered in formose, where C_5 produces C_2 and C_3. Here, the path for B is of length zero (that is: $B_1 = A_1$). We will now study the case where all forward reactions are irreversible, but our final result will turn out to be more general. When forward reactions are irreversible, we can write probabilities in the form of Eq. (6.35) for successful trajectories:
$p_A : A_1 \longrightarrow ... \longrightarrow B_1 + C_1$,
$p_B : B_1 \longrightarrow ... \longrightarrow A_1$,
$p_C : C_1 \longrightarrow ... \longrightarrow A_1$,
which are given by

$$p_A = \prod_{s=1}^{n} \frac{k_s^+}{k_s^+ + k_s^-}, \qquad p_B = \prod_{s=1}^{n} \frac{\kappa_s^+}{\kappa_s^+ + \kappa_s^-}, \qquad p_C = \prod_{s=1}^{n} \frac{k_s^{`+}}{k_s^{`+} + k_s^{`-}}. \tag{6.50}$$

Let us now take perspective of a B_1 species, which with a probability $p_A p_B$ completes a cycle, to return a B_1 and C_1 species

$$B_1 \xrightarrow{p_A p_B} B_1 + C_1. \tag{6.51}$$

Viewed in isolation, B_1 has then performed an allocatalytic cycle to produce C_1. Only C_1 can make new B_1 species, through

$$C_1 \xrightarrow{p_A p_C} B_1 + C_1. \tag{6.52}$$

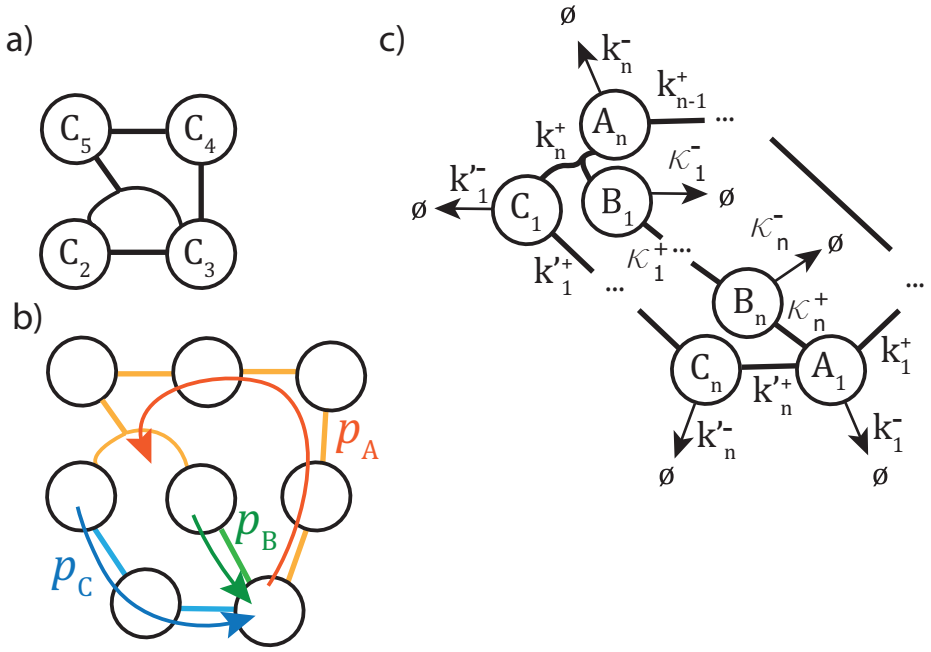


Figure 6.8: a) an example of a network with two branches. Since the common endpoint of the B and C branch is C_3, $p_B = 1$. b) schematic representation of the three trajectories that constitute the autocatalytic cycle, with their associated success probabilities. c) General representation of an autocatalytic cycle with two branches. Effective microscopic rate constants have been added for degradation and irreversible progression.

Let us denote P_n^B the probability that a B_1 molecule performs n successful cycles and is subsequently lost, leading to a net C_1 production of

$$B_1 \xrightarrow{P_n^B} nC_1, \tag{6.53}$$

where P_n^B follows from the number of successful trials of process (6.51) before degradation, which thus follows the geometric distribution

$$P_n^B = (p_A p_B)^n (1 - p_A p_B). \tag{6.54}$$

By the same argument, we can denote P_n^C the probability that a C_1 molecule performs n successful cycles and is subsequently being lost, leading to a net result of

$$C_1 \xrightarrow{P_n^C} nB_1. \tag{6.55}$$

Again, we can write

$$P_n^C = (p_A p_C)^n (1 - p_A p_C). \tag{6.56}$$

Now, we can combine these two, by considering the combination of a B_1 producing s C_1 molecules, and each C_1 producing, respectively, $n_1, n_2, .., n_s$ B_1 molecules:

$$B_1 \longrightarrow s\, C_1 \longrightarrow (n_1 + ... + n_s) B_1. \tag{6.57}$$

This can be written as an overall process that generates n new B_1 molecules, from a single B_1

$$B_1 \xrightarrow{\mathscr{P}_n^B} nB_1. \tag{6.58}$$

The distribution $\mathscr{P}_n^{\mathrm{B}}$ for Eq. (6.58) is then derived by considering all possible realizations of process (6.57):

$$\mathscr{P}_n^{\mathrm{B}} = \sum_{s=0}^{\infty} P_s^{\mathrm{B}} \sum_{n_1,\ldots,n_s}^{\infty} \prod_{k=0}^{s} P_{n_k}^{\mathrm{C}} \delta_{n_1+\ldots+n_s}^{n} \tag{6.59}$$

Upon substitution of Eqs. (6.54) and PNC1, we find

$$\mathscr{P}_n^{\mathrm{B}} = (1-p_{\mathrm{A}}p_{\mathrm{B}}) \sum_{s=1}^{\infty} (p_{\mathrm{A}}p_{\mathrm{B}})^s (1-p_{\mathrm{A}}p_{\mathrm{C}})^s (p_{\mathrm{A}}p_{\mathrm{C}})^n \binom{n+s-1}{n}, \quad n \geq 1 \tag{6.60}$$

$$\mathscr{P}_0^{\mathrm{B}} = (1-p_{\mathrm{A}}p_{\mathrm{B}}) \sum_{s=0}^{\infty} (p_{\mathrm{A}}p_{\mathrm{B}})^s (1-p_{\mathrm{A}}p_{\mathrm{C}})^s = \frac{1-p_{\mathrm{A}}p_{\mathrm{B}}}{1-p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{C}})} \tag{6.61}$$

Using the identity

$$\sum_{k=0}^{\infty} x^k \binom{k+n}{n} = \frac{1}{(1-x)^{n+1}}, \tag{6.62}$$

we can simplify $\mathscr{P}_n^{\mathrm{B}}$ to

$$\mathscr{P}_n^{\mathrm{B}} = \frac{p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{B}})(1-p_{\mathrm{A}}p_{\mathrm{C}})}{(1-p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{C}}))} \left(\frac{p_{\mathrm{A}}p_{\mathrm{C}}}{1-p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{C}})} \right)^n. \tag{6.63}$$

For further use, let us introduce the shorthand notation

$$\mathscr{P}_n^{\mathrm{B}} = \beta\alpha^n, \quad n \geq 1 \tag{6.64}$$

$$\mathscr{P}_0^{\mathrm{B}} = 1-\beta\frac{\alpha}{1-\alpha}, \tag{6.65}$$

with

$$\alpha = \frac{p_{\mathrm{A}}p_{\mathrm{C}}}{1-p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{C}})}, \qquad \beta = \frac{p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{B}})(1-p_{\mathrm{A}}p_{\mathrm{C}})}{(1-p_{\mathrm{A}}p_{\mathrm{B}}(1-p_{\mathrm{A}}p_{\mathrm{C}}))}. \tag{6.66}$$

Notice that we only need two variables: $\pi_{\mathrm{B}} = p_{\mathrm{A}}p_{\mathrm{B}}$ and $\pi_{\mathrm{C}} = p_{\mathrm{A}}p_{\mathrm{C}}$, which mark the respective probabilities for a B_1 and C_1 molecule to finish their cycle. In terms of these probabilities, we find a more generally applicable from of Eq. (6.63). For our present purposes it will be fruitful to not make this substitution, as we will be interested in the effect of p_A.

We now have the distribution that is needed to construct a branching process.

Two cyclic paths: branching process

Let us now compare the process (6.38) to the process (6.58). We see that the single cyclic path is a particular of Eq. (6.58), where there can only be 0 or 2 descendants. The more general case, where there can be n according $\mathscr{P}_n^{\mathrm{B}}$, is generally not amenable to analysis as a birth-death process, but rather makes use of the theory of branching processes, for which a single step is shown in Fig. 6.9.

Let us now study a population of autocatalysts according to such a process. Let N_m denote the number of autocatalysts after m repetitions of Eq. (6.58), where we start with only B_1 molecules: $N_0 = N_{\mathrm{B}_1,0}$. The stochastic process can be written as

$$N_{m+1} = N_m + X_m - 1, \tag{6.67}$$

Where X_m is a random variable drawn from $\mathscr{P}_n^{\mathrm{B}}$.

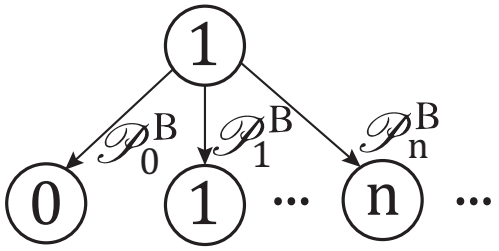


Figure 6.9: In a branching process, the number of descendants n is drawn from a distribution $\mathscr{P}_n^{\mathrm{B}}$, such that $N_{m+1} = N_m + n - 1$ with probability $\mathscr{P}_n^{\mathrm{B}}$. Starting at $N_0 = 1$, the population goes extinct in the next round with probability $\mathscr{P}_0^{\mathrm{B}}$.

To find the extinction probability, we can again use the total probability theorem[31] as we also used for Eq. 6.68. Let us again denote d_m the probability of extinction, starting from a single species B_1 ($N_0 = 1$), m steps from now. In terms of d_{m-1}, we can then write

$$d_m = \mathscr{P}_0^{\mathrm{B}} + \mathscr{P}_1^{\mathrm{B}} d_{m-1} + \mathscr{P}_2^{\mathrm{B}} d_{m-1}^2 + ... = \sum_{i=0}^{\infty} \mathscr{P}_k^{\mathrm{B}} d_{m-1}^k. \tag{6.68}$$

where every term counts the contribution to extinction if k descendants are spawned in the first round, which collectively go extinct in the next $m-1$ rounds with probability d_{m-1}^k. Recalling that $d_m \to d$ for $m \to \infty$ (Eq. (6.41)), we find

$$d = \sum_{i=0}^{\infty} \mathscr{P}_k^{\mathrm{B}} d^k = 1 - \beta \frac{\alpha}{1-\alpha} + \sum_{k=0}^{\infty} \beta (\alpha d)^k. \tag{6.69}$$

which reduces to

$$d = 1 - \beta \frac{\alpha}{1-\alpha} - \beta \frac{\alpha d}{1-\alpha d}. \tag{6.70}$$

Eq. (6.70) can be solved to give either $d = 1$, or

$$d = \frac{\beta}{\alpha - 1} + \frac{1}{\alpha} \tag{6.71}$$

The fixation probability again verifies $P_{fix} = 1 - d$. Let us now move to an example where P_{fix} can be expressed in terms of only α.

Example

As an illustration, let us consider the simple network given by Fig. 6.10. For irreversible reactions, the probabilities to successfully traverse the path's A and B are respectively:

$$p_{\mathrm{A}} = \prod_{s=1}^{n} \frac{k_s^+}{k_s^+ + k_s^-}, \qquad p_{\mathrm{B}} = \frac{\kappa_1^+}{\kappa_1^+ + \kappa_1^-} \frac{\kappa_2^+}{\kappa_2^+ + \kappa_2^-}. \tag{6.72}$$

Let us denote $\mathscr{P}_n^{\mathrm{A}}$ the probability of acquiring n new A_1 species from a single A_1 species

$$\mathrm{A}_1 \longrightarrow \mathrm{nA}_1 \tag{6.73}$$

This is due to successful production of B_1 species and their subsequent conversion to A_1. As shown in Sec. 6.3 we then find

$$\mathscr{P}_n^{\mathrm{A}} = (p_{\mathrm{A}} p_{\mathrm{B}})^n (1 - p_{\mathrm{A}}) \sum_{j=0}^{\infty} (p_{\mathrm{A}}(1 - p_{\mathrm{B}}))^j \binom{n+j}{j}. \tag{6.74}$$

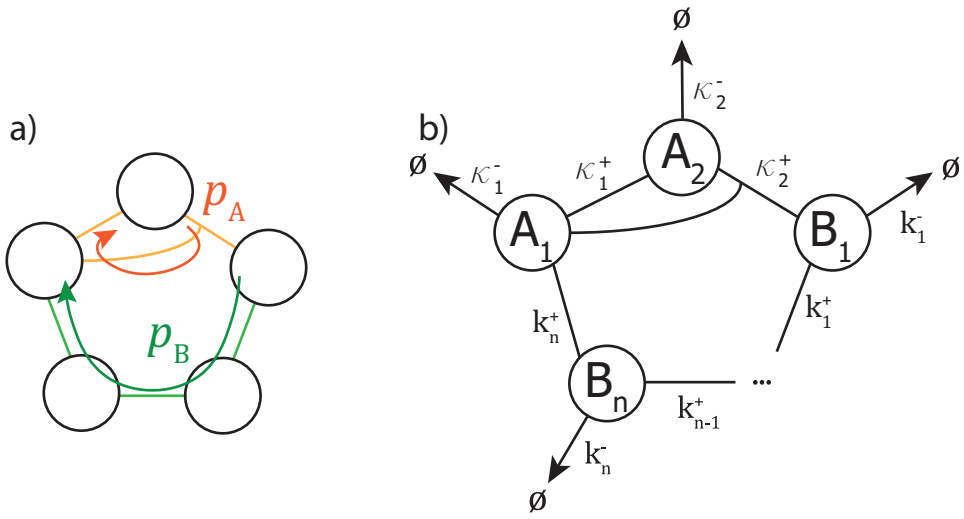


Figure 6.10: a) two paths with corresponding success probabilities. b) General case, with n states. Effective microscopic rate constants have been added for degradation and irreversible progression.

This expression simplifies to

$$\mathscr{P}_n^{\mathrm{A}} = \left(\frac{p_{\mathrm{A}}p_{\mathrm{B}}}{1-p_{\mathrm{A}}(1-p_{\mathrm{A}})}\right)^n \frac{1-p_{\mathrm{B}}}{1-p_{\mathrm{A}}(1-p_{\mathrm{B}})} = (1-\alpha)\alpha^n \tag{6.75}$$

As in Sec.6.3, we now define a branching process for N_m, starting only with molecules of A_1

$$N_{m+1} = N_m + X_{n,m} - 1 \tag{6.76}$$

where $X_{n,m}$ is a random variable drawn from $\mathscr{P}_n^{\mathrm{A}}$. Again, applying Eq. (6.69), we find for the extinction probability d

$$d = (1-\alpha)\sum_{k=0}^{\infty}(\alpha d)^k = \frac{1-\alpha}{1-\alpha d}. \tag{6.77}$$

Which admits the solution $d = 1$ and

$$d = \frac{1}{\alpha} - 1, \tag{6.78}$$

which leads to a fixation probability

$$P_{fix} = \begin{cases} 1-d = 2-\frac{1}{\alpha}, & \alpha \geq \frac{1}{2}, \\ 0, & \alpha < \frac{1}{2}, \end{cases} \tag{6.79}$$

Here, the parameter α plays the role that p_c plays in the network of Fig. (6.7), but it is composed of two path probabilities

$$\alpha = \frac{p_{\mathrm{A}}p_{\mathrm{B}}}{1-p_{\mathrm{A}}(1-p_{\mathrm{B}})}. \tag{6.80}$$

The fixation probability P_{fix} is traced in Fig. 6.11 for both cases.

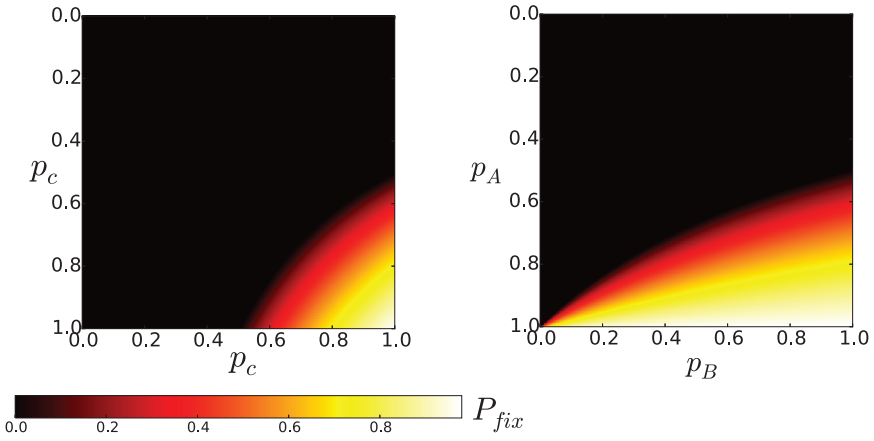


Figure 6.11: Comparison of P_{fix} for left) single-cycle networks such as Fig.(6.6) and right) networks with two cyclic branches as in Fig. (6.10). For comparison, p_c is traced on both axes for the single-cycle. For two cyclic branches, a low value of p_B can be compensated by a high value of p_A, but not vice versa.

The effect of p_A is quite different from that of p_B, as can be seen in Fig. 6.11. When $p_\mathrm{B} = 1$, survival of the double cycle network (Fig. 6.10) become equivalent to that of the single cycle network (Fig. 6.6).

In contrast, when when $p_\mathrm{A} \to 1$, the double-cycle network reaches deterministic fixation ($P_{fix} \to 1$), provided $p_\mathrm{B} > 0$. A high value of p_B ensures that B_1 will likely end up becoming A_1, whereas a high value of p_A guarantees that A_1 will typically produce a large amount of B_1 molecules before perishing, in an effective reaction (see Sec. 6.3)

$$\mathrm{A}_1 \to \mathrm{nB}_1. \tag{6.81}$$

Here, n is an exponentially distributed random variable, with an average

$$\langle n \rangle = p_\mathrm{A}/(1-p_\mathrm{A}). \tag{6.82}$$

Since $\langle n \rangle$ can be much larger than 2, generation of a new A_1 molecule from $\langle n \rangle$ B_1 can proceed at much lower rates of success p_B, since more attempts can be made. Indeed, the threshold at which P_{fix} becomes nonzero ($\alpha = 1/2$) corresponds to $p_\mathrm{B} = (1-p_\mathrm{A})/p_\mathrm{A}$, which is the point where on average one molecule is propagated: $p_\mathrm{B}\langle n \rangle = 1$.

This criterion confirms the picture also put forward by King[27] and Bagley et al.[11], a molecule should on average leave more than a single copy of itself. The ease at which this proceeds, however, is dependent on the network structure, as illustrated by our double cycle example network (Fig. 6.10). King's survival criterion Eq. (6.11) (also regularly put forward by others, e.g. [11, 16]) is not generally applicable, because the autocatalytic survival problem cannot generally be mapped to a birth-death process. As shown here, the problem can be treated more generally as a branching processes.

6.3.1 Selection of autocatalytic networks

The fixation probability of an autocatalytic cycle diminishes rapidly with an increased rate of side reactions, favoring species with few and slow side reactions.

In a CSTR, reactions need not only be specific, but also rapid with respect to outflow. A high specificity is thus no guarantee for survival, the reaction must outpace the outflow.

The network structure itself is instrumental in stabilizing survival. As demonstrated in Sec. 6.3, networks with multiple branches can improve their fixation rate by having an efficient branch that compensates for a less efficient one.

Dissipation and selection

Decorating a reaction network with further autocatalytic cycles is (typically) expected to lead to a network that increasingly dissipates free energy as the network introduces more and more autocatalytic cycles. At the same time, by the fact that nonspecific (having considerable side reactions) autocatalytic cycles are inherently very dissipative, there may be indirect selection against strong dissipation. A cycle that is specific lacks this inherent dissipation from side reactions, but it may still be strongly dissipative by its own chemistry. We can thus argue that there is a rudimentary selection for efficiency and against squandering resources.

In the works of Smith and Morowitz[32], chemical complexity is mentioned as a means of dissipating free energy: *More fundamentally, the formation of chemical complexity can be a means to the dissipation of free energy, not only indirectly through complex life but plausibly even in early geochemistry.* In their book, Smith and Morowitz explain well what they mean exactly by that and what the limits are of this statement. This is with good reason, because linking entropy production directly to life has proven fertile ground for misinterpretation.

By being far from equilibrium chemical systems can display impressive features, that require dissipation to fuel them. Increasingly, dissipation has been presented as a desirable feature in its own right, with more dissipation being considered better. We should be careful with such interpretations: free energy is a precious resource that can be expended (dissipated) only once. A most excellent source of dissipation is the direct degradation of energy-rich molecules. On the level of the cell and chemical networks, squandering resources is clearly not a viable strategy. Putting them to good use to be competitive, however, definitely is.

The same is true for an autocatalytic network that is being decorated: maximizing dissipation is not an 'objective' in itself. However, gradual evolution of a chemical network (or life itself) will lead the system to steadily tap into new resources that fuel new innovations. This may be expected to increase dissipation in the long run.

Triggering bistabilities in autocatalysis

So far, we have focused on autocatalytic networks that can be triggered by single-molecule perturbations. As shown in Sec. 5.5, this does not encompass all autocatalytic systems: some reactions require reactants to pass a concentration threshold, as was illustrated for the termolecular reaction in the Schlögl reaction[33], which performs a net reaction

$$2\mathrm{X} + \mathrm{E} \underset{-g}{\overset{g}{\rightleftharpoons}} 3\mathrm{X}. \tag{6.83}$$

Our criteria for triggering via rare autocatalysts (Eq. (6.13)) exclude triggering such reactions as perturbations of a steady-state. We may then wonder: can such reactions still have an effect in the CSTR model?

While we may not trigger such reactions by single-molecule perturbations, we may trigger them when large changes occur, such as when a new autocatalytic cycle is successfully triggered. Such an event will macroscopically alter the system composition, and the system may thereby accumulate enough autocatalysts that react with other autocatalysts as in Eq. 6.83. Such reactions are thus not excluded, but their incorporation follows a different mechanism.

6.4 Common obstacles in chemical evolution

In the last section, we derived a model similar to the 'evolving metabolic network' by Bagley et al[11]. Starting from the stoichiometric matrix and chemical considerations, we found that we can recover their model from microscopic considerations and extend it to other autocatalytic networks. Now that the model is formulated in a larger framework, it is more readily amenable to further extension.

Nevertheless, we should readily understand that the theory of the 'evolving metabolic network' is presently not considered as a dominant motif in abiogenesis. There are some good chemical reasons for this, most of which also apply to the GARD model, RAF-sets of polymers and other approaches. Let us here consider four important obstacles that these theories must overcome.

I. Variation and the rarity of autocatalysis

At present, autocatalysis does not seem to be an abundant chemical property. Consequently, it is not evident that there is enough autocatalysis to produce chemical networks that progressively nucleate new autocatalytic cycles. If we want evolutionary trajectories that are contingent, autocatalysis must become an even more abundant property.

In evolutionary terms, a lack of autocatalysis can be compared to a lack of variation. By extension, we must wonder if such a mode of evolution can be open-ended on the timescale of interest [25].

II. More species: Side reactions, trapping states, inactive complexes

A popular solution to the rarity of autocatalysis[7] is the addition of ever more random compounds to the mix. This intuition[18] is based on Erdös-Renyi graphs [23], where random graphs will, at a sufficient size and connectivity, form cyclic graphs. A diverse mixture will have many different reactions, some of which correspond to autocatalysis. By subsequent network decorations, a network will generate more diversity, thus laying the ground for future autocatalytic cycles [11].

Of course, the formation of cycles is just one consequence of an explosive growth in the number of reactions. In particular, more components also facilitate side reactions, as argued by Orgel[34] and treated in detail by Szathmary[35]. Side reactions are but one issue: molecules may bundle in kinetically trapped inactive states, a common occurrence when complementarity comes into play. In addition, to accommodate more components in a reaction volume, individual concentrations must decrease, which drastically reduces reaction rates. In the present model, such contributions will quickly reduce any fixation probability P_{fix} to 0.

Accommodating more types of species reduces the concentration at which each species can be present. Volume conservation guarantees this, but in a CSTR more strict conservation on the total mass influx applies as well[36]. Such progressive dilution of reactants decreases reaction rates. In a CSTR with a fixed residence time τ_r, this sequential decoration of networks would make overly diverse networks progressively less fit in the face of fixed degradation.

III. Chemical networks graphs and modeling

Traditionally, modelling approaches in the field have relied on artificial chemistry that equate to random graphs. Indeed, in Ref. [11] but also for example GARD and RAF-sets, the evolving metabolism is simulated using such assumptions for chemical networks.

In practice, chemistry is highly structured. Hard nucleophiles react with hard electrophiles, soft nucleophiles react with soft electrophiles (HSAB theory). Acids react with bases, oxidizing agents react with reducing agents, and so forth. The result of such reactions is normally weaker acids, weaker bases and overall less reactive species [37, 38].

We can study reactions, because particular functional groups and elements perform those reactions (relatively) systematically. In this sense, random graphs are often a poor proxy for chemistry: they lack this deeply correlated structure. Their interpretation and justification requires great care. For the GARD system, this challenge has been acknowledged: efforts are made to understand and justify the rates in the model [6].

Accounting for reactivity, conservation laws and other structural features of chemical networks remains an outstanding challenge.

IV. Model restrictions vs scenario restrictions

Many prebiotic scenarios insist on one type of autocatalytic evolution, mediated by a very narrow subset of possible chemistries. E.g. autocatalysis for lipids-first, RNA-world, peptides or RAF sets of polymers. It is by no means evident that we can assign such a key role to one type of molecule, neither from biology nor from chemistry.

A whole list of chemical evolution mechanisms in the literature have been argued not to be competent enough to provide the evolutionary potential for abiogenesis [3]. A hidden assumption in this analysis, however, is that all these mechanisms are mutually exclusive and have to do all of abiogenesis by themselves. That is indeed a large task for mechanisms that have all been criticized for their limited evolvability.

As shown in Sec.5.4, other stoichiometric processes such as evaporation, diffusion, partitioning can be the building blocks of autocatalytic cycles not accessible to pure chemistry. The systems of interest in origins of life are inherently open and communicating with an environment. This is not simply an argument in favor of multicompartment autocatalysis, it is an argument in favor of a vibrant mix of autocatalytic mechanisms. Needlessly reducing our scope to one mechanism and chemistry is paradoxical in view of obstacle I, the rarity of autocatalysis. The evolutionary challenge we wish to address would rather require us to widen our scope of compounds and mechanisms.

Towards a synthesis of chemical evolution

To come to a more complete description of chemical evolution, these and many other obstacles must be taken into consideration. In an attempt to provide a more complete description of chemical evolution, let us therefore consider removing some assumptions. In particular, we consider the following:
i) We do not impose a particular chemistry of interest.
ii) Our system need not be a CSTR
iii) Our system is part of a larger environment, with stoichiometric processes that describe its exchange.

For such a general system, we can, again, use a stoichiometric matrix formalism, in which we now also specify the local environment of a reactant. In such a formalism, a compound can simultaneously be a rare reactant in one compartment and an abundant reactant in another.

6.5 Spatial autocatalysis and fixation

The essential extension we have provided, is the introduction of more environments/compartments. Let us now consider how this changes our perspective on autocatalysis and chemical evolution, by explicitly considering the role of space in fixation.

Let us again consider the autocatalytic network described by Figs 5.7 and 5.6, which is described by the reactions

$$\mathrm{B}^{\mathrm{II}} + \mathrm{BC}^{\mathrm{II}} \rightleftharpoons [\mathrm{BCB}]^{\mathrm{II}}, \tag{6.84}$$
$$[\mathrm{BCB}]^{\mathrm{II}} \rightleftharpoons [\mathrm{BCB}]^{\mathrm{I}}, \tag{6.85}$$
$$[\mathrm{BCB}]^{\mathrm{I}} \rightleftharpoons \mathrm{B}^{\mathrm{I}} + \mathrm{BC}^{\mathrm{I}}, \tag{6.86}$$
$$\mathrm{D}^{\mathrm{I}} + \mathrm{BC}^{\mathrm{I}} \rightleftharpoons \mathrm{B}^{\mathrm{I}} + \mathrm{DC}^{\mathrm{I}}, \tag{6.87}$$
$$\mathrm{B}^{\mathrm{I}} \rightleftharpoons \mathrm{B}^{\mathrm{II}}. \tag{6.88}$$

Let us now consider a situation where compartment I is a small compartment that can exchange B and BCB with a spatially extended phase II. In Fig. 6.12, the corresponding autocatalytic network is drawn in terms of rare autocatalysts.

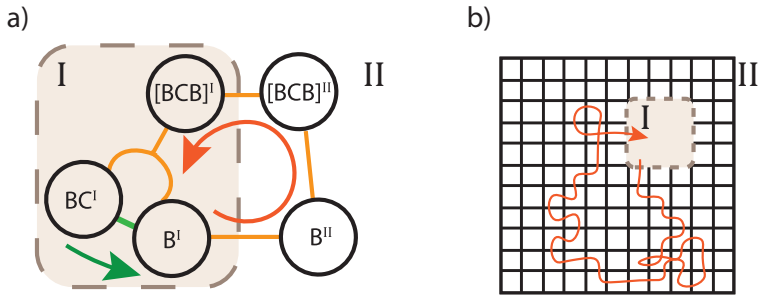


Figure 6.12: a) An illustration of spatial autocatalysis for reaction network given by Eqs (6.84)-(6.88). Orange and green arrows correspond to the A and B path, as discussed in Sec. 6.3. b) Phase I is a small compartment, e.g. a vesicle. Phase II is a spatially extended environment, e.g. a bulk volume phase. Completing the A path (orange) requires a return to phase I. Lattice sites are not drawn to scale.

Let us consider that compounds displace only by diffusion. As can be seen in 6.12b, an important step in the autocatalysis is then the return to the original compartment. As the dimension of phase II increases, returning will become a more time-consuming process, which makes degradation processes in phase II particularly detrimental. Even in the absence of degradation reactions, there can be 'effective degradation' by diffusion: a fraction of species may never diffuse back (e.g. for a discrete random walk on a cubic lattice ($d = 3$), the probability to return to the origin is 34 %. In our example, conditions are a bit milder: compartment I is much larger than a lattice site.).

We see that there are some inherent problems that may come up when an autocatalytic cycle involves exchange between different phases. Our example system will have difficulty sustaining autocatalysis, especially in the presence of degradation reactions. We will now discuss three solutions by which the fixation of such a multicompartment autocatalytic cycle can be improved: i) Spatial confinement, ii) Cooperation, iii) Network structure. Interestingly, all of these example can be argued to have a counterpart in biology, which will be discussed in Sec (6.5.4).

6.5.1 Spatial confinement

By confining compartment I in a small enough surrounding space, excursions of diverging length and time are excluded. Let us here consider the case of a compartment confined in a small cavity, with weak coupling to an outside environment, as shown by Fig. 6.13.

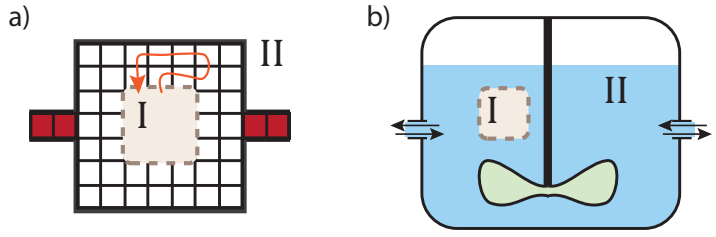


Figure 6.13: Two examples of a confined compartment I. a) Diffusion is limited to a small lattice due to physical impediments, leading to more rapid return. 'Effective degradation' by diffusion may still exist by escape of the confined phase II, which can happen here by diffusion past the red lattice sites. b) A well-mixed CSTR fluid II in contact with compartment I forms the well-mixed counterpart, of a).

In the spatial autocatalytic cycle (Fig. 6.12a), the environment introduces degradation pathways

for B and [BCB], through

$$\mathrm{B}^{\mathrm{II}} \rightleftharpoons \emptyset, \tag{6.89}$$

$$[\mathrm{BCB}]^{\mathrm{II}} \rightleftharpoons \emptyset. \tag{6.90}$$

We will study this system with mean-field model, appropriate when exchange between I and II is slow.

When species displace by diffusion, well-mixed situations can occur, when spatial relaxation of gradients occurs on a faster timescale $\tau_{mix} \propto L^2/D$ than exchange (τ_{ex}) between I and II. For our present discussion, we will consider the well-mixed CSTR limit (Fig. 6.13b) of a confined autocatalytic reaction. We can then use a coarse-grained description for degradation, by only specifying the phase (I, II) in which a species resides

$$\mathrm{B}^{\mathrm{II}} \xrightarrow{\kappa^{\emptyset}} \emptyset, \quad [\mathrm{BCB}]^{\mathrm{II}} \xrightarrow{\kappa^{\emptyset}} \emptyset. \tag{6.91}$$

Degradation rates in II follow the mean residence time τ_r: $k^{\emptyset} = k^{\emptyset} = 1/\tau_r$ (Sec. 3.2).

Similarly, a well-mixed description can be made for exchange between I and II

$$\mathrm{B}^{\mathrm{I}} \underset{\kappa_{\mathrm{B}}^{-}}{\overset{\kappa_{\mathrm{B}}^{+}}{\rightleftharpoons}} \mathrm{B}^{\mathrm{II}}, \quad [\mathrm{BCB}]^{\mathrm{I}} \underset{\kappa_{[\mathrm{BCB}]}^{-}}{\overset{\kappa_{[\mathrm{BCB}]}^{+}}{\rightleftharpoons}} [\mathrm{BCB}]^{\mathrm{II}}. \tag{6.92}$$

Supposing I and II are similar environments (same solvent, etc.), detailed balance requires that $\kappa_{\mathrm{B}}^{+}/\kappa_{\mathrm{B}}^{-} = V^{\mathrm{II}}/V^{\mathrm{I}}$ and $\kappa_{[\mathrm{BCB}]}^{+}/\kappa_{[\mathrm{BCB}]}^{-} = V^{\mathrm{II}}/V^{\mathrm{I}}$.

Finally, we have the chemical reactions

$$\mathrm{B}^{\mathrm{I}} + \mathrm{BC}^{\mathrm{I}} \rightleftharpoons [\mathrm{BCB}]^{\mathrm{I}}, \quad \mathrm{B}^{\mathrm{II}} + \mathrm{BC}^{\mathrm{II}} \rightleftharpoons [\mathrm{BCB}]^{\mathrm{II}}, \quad \mathrm{D}^{\mathrm{I}} + \mathrm{BC}^{\mathrm{I}} \rightleftharpoons \mathrm{B}^{\mathrm{I}} + \mathrm{DC}^{\mathrm{I}}. \tag{6.93}$$

Removing abundant species ($\mathrm{D}^{\mathrm{I}}, \mathrm{BC}^{\mathrm{II}}$), and rare waste products (DC^{I}) from the description, we find an effective description

$$[\mathrm{BCB}]^{\mathrm{I}} \xrightarrow{k_{[\mathrm{BCB}]}^{\mathrm{I}}} \mathrm{B}^{\mathrm{I}} + \mathrm{BC}^{\mathrm{I}}, \quad \mathrm{B}^{\mathrm{II}} \underset{k_{[\mathrm{BCB}]}^{\mathrm{II}}}{\overset{k_{\mathrm{B}}^{\mathrm{II}}}{\rightleftharpoons}} [\mathrm{BCB}]^{\mathrm{II}}, \quad \mathrm{BC}^{\mathrm{I}} \xrightarrow{k_{\mathrm{BC}}^{\mathrm{I}}} \mathrm{B}^{\mathrm{I}}. \tag{6.94}$$

With the inclusion of the volumes we can then write the relevant transition rates w on the level of single molecules, which for compartment I yields

$$w_{\mathrm{B}^{\mathrm{I}} \to \mathrm{B}^{\mathrm{II}}} = \kappa_{\mathrm{B}}^{+}, \quad w_{[\mathrm{BCB}]^{\mathrm{I}} \to [\mathrm{BCB}]^{\mathrm{II}}} = \kappa_{[\mathrm{BCB}]}^{+}, \quad w_{[\mathrm{BCB}]^{\mathrm{I}} \to \mathrm{B}^{\mathrm{I}}} = k_{[\mathrm{BCB}]}^{\mathrm{I}}, \quad w_{\mathrm{BC}^{\mathrm{I}} \to \mathrm{B}^{\mathrm{I}}} = k_{\mathrm{BC}}^{\mathrm{I}}. \tag{6.95}$$

Analogously, we find for phase II

$$w_{\mathrm{B}^{\mathrm{II}} \to \mathrm{B}^{\mathrm{I}}} = \kappa_{\mathrm{B}}^{-}, \quad w_{[\mathrm{BCB}]^{\mathrm{II}} \to [\mathrm{BCB}]^{\mathrm{I}}} = \kappa_{[\mathrm{BCB}]}^{-}, \quad w_{\mathrm{B}^{\mathrm{II}} \to [\mathrm{BCB}]^{\mathrm{II}}} = k_{\mathrm{B}}^{\mathrm{II}}, \quad w_{[\mathrm{BCB}]^{\mathrm{II}} \to \mathrm{B}^{\mathrm{II}}} = k_{[\mathrm{BCB}]}^{\mathrm{II}}, \tag{6.96}$$

and for the degradation process we write

$$w_{\mathrm{B}^{\mathrm{II}} \to \emptyset} = w_{[\mathrm{BCB}]^{\mathrm{II}} \to \emptyset} = k^{\emptyset}. \tag{6.97}$$

We can now return to our stochastic models for the survival of autocatalytic cycles derived in Sec.6.3, applied to the network given in Fig. 6.12a.

Supposing $N_{\mathrm{D}}^{\mathrm{I}}, N_{[\mathrm{BCB}]}^{\mathrm{II}} \gg 1$, and initially $N_{[\mathrm{BCB}]}^{\mathrm{I}} = N_{\mathrm{B}}^{\mathrm{I}} = N_{\mathrm{BC}}^{\mathrm{I}} = N_{\mathrm{DC}}^{\mathrm{I}} = N_{[\mathrm{BCB}]}^{\mathrm{II}} = N_{\mathrm{B}}^{\mathrm{II}} = 0$, reactions III and IV in Fig. 5.6 are (initially) irreversible.

BC^{I} converts to B with no intervention of side reactions and no reverse reactions. The fixation probability P_{fix} can then be found from the probability that a B^{I} successfully traverses, incorporates

BC to become [BCB], and returns to irreversibly form $2B^{I}$. Denoting this probability with p_B, the fixation probability follows directly from Eq. (6.44).

$$P_{fix} = \begin{cases} 0, & p_B \leq \frac{1}{2}, \\ 2 - \frac{1}{p_B}, & p_B \geq \frac{1}{2}, \end{cases} \tag{6.98}$$

Note however, that B and [BCB] can go back and forth between I and II, and interconvert reversibly in phase II, which needs to be taken into account for determining p_B. Noting that there is no degradation pathway for B^{I} and that it can only transit to B^{II}, we can ignore it in our calculation and consider the total success of all trajectories starting from B^{II}:

$$p_B = p^+_{+1} \sum_{k=0}^{\infty} (p^+_{-2} p^+_{+1} + p^+_{+2} p^+_{-3})^k p^+_{+2} p^+_{+3} = \frac{p^+_{+1} p^+_{+2} p^+_{+3}}{1 - (p^+_{-2} p^+_{+1} + p^+_{2+} p^+_{3-})} \tag{6.99}$$

where p^+_{+k} is the probability to successfully advance forward from node k to $k+1$, and p^+_{-k} the probability to successfully return from k to $k-1$. Nodes have the following correspondence with compounds 1: B^{II}, 2: $[BCB]^{II}$, 3: $[BCB]^{I}$. We can express p^+_{+k} in terms of transition rates w

$$p^+_{1+} = \frac{w_{B^{II} \to [BCB]^{II}}}{w_{B^{II} \to [BCB]^{II}} + w_{B^{II} \to \emptyset}} \tag{6.100}$$

$$p^+_{1-} = \frac{w_{[BCB]^{II} \to B^{II}}}{w_{[BCB]^{II} \to B^{II}} + w_{[BCB]^{II} \to \emptyset} + w_{[BCB]^{II} \to [BCB]^{I}}} \tag{6.101}$$

$$p^+_{2+} = \frac{w_{[BCB]^{II} \to [BCB]^{I}}}{w_{[BCB]^{II} \to B^{II}} + w_{[BCB]^{II} \to \emptyset} + w_{[BCB]^{II} \to [BCB]^{I}}} \tag{6.102}$$

$$p^+_{2-} = \frac{w_{[BCB]^{I} \to [BCB]^{II}}}{w_{[BCB]^{I} \to [BCB]^{II}} + w_{[BCB]^{I} \to B^{I}}} \tag{6.103}$$

$$p^+_{3+} = \frac{w_{[BCB]^{I} \to B^{I}}}{w_{[BCB]^{I} \to [BCB]^{II}} + w_{[BCB]^{I} \to B^{I}}} \tag{6.104}$$

In Fig 6.14, some trajectories are shown for this process, which was modeled using Gillespie's Algorithm[39].

We follow the populations for 10000 transitions. By following the fraction of nonextinct populations as a function of time, it was checked that this yields satisfactory convergence towards the fixation probability P_{fix}, which agrees well with the analytical result (Eq. 6.99) as shown in Fig. 6.15.

6.5.2 Cooperation

In Fig. 6.12, fixation of an autocatalytic cycle requires a freshly formed [BCB] compound to return to its original compartment. If there are multiple compartments with a similar composition, as shown in Fig. 6.16, this requirement can be relaxed: when a released B molecule forms [BCB], it can enter any other compartment, to yield an overall production of B. In Ch.5, we also considered an analogous example where the coupling of two chemically distinct compartments can lead to autocatalysis, similar to syntrophy in ecology. These considerations are reminiscent of a model by Lemarchand and Jullien [40] using autocatalytically assembling vesicles which exhibit chemically mediated ecological behavior.

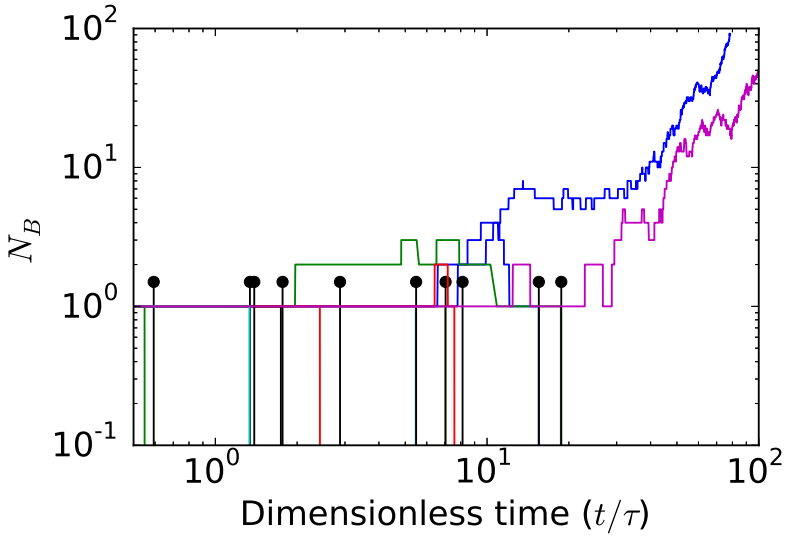


Figure 6.14: 12 stochastic trajectories for $N_{\mathrm{B}} = N_{\mathrm{B^I}} + N_{\mathrm{B^{II}}} + N_{\mathrm{[BCB]^I}} + N_{\mathrm{[BCB]^{II}}}$, starting with $N_{\mathrm{B^{II}}} = 1$, $V^I = V^{II} = 1.0$ and setting other concentrations of B-species to 0. Rate constants are $\kappa_{\mathrm{B}}^{+} = \kappa_{\mathrm{B}}^{-} = 1.0$, $\kappa_{\mathrm{[BCB]}}^{+} = \kappa_{\mathrm{[BCB]}}^{-} = 0.5$, $k_{\mathrm{[BCB]}}^{\mathrm{I}} = k_{\mathrm{[BCB]}}^{\mathrm{II}} = 10.0$, $k_{\mathrm{B}}^{I} = 10.0$, $k_{\mathrm{B}}^{\mathrm{II}} = 10.0$, $k_{\mathrm{BC}}^{\mathrm{I}} = 100.0$, $k^{\emptyset} = 0.15$. A trajectory consists of 10000 subsequent transitions, unless N_{B} reaches 0 before. Note that only 2 trajectories reach an exponential growth regime, black dots mark the points where the other 10 trajectories go extinct. Time is rendered dimensionless by $\tau = 1/\kappa_{\mathrm{B}}^{+}$.

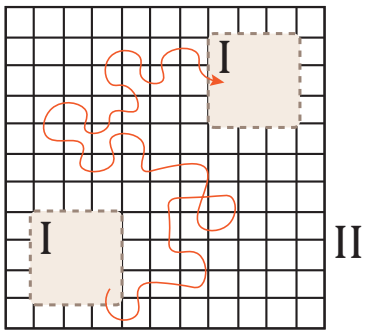


Figure 6.16: Spatial autocatalysis can be mediated by two compartments performing the same chemistry. A compartment recovers autocatalysts released by its neighbor. In its absence, some of these autocatalysts would likely diffuse away or degrade. Now, they are put to use to further enhance the autocatalytic cycle in Fig. 6.12a. This provides the basis for cooperation.

Cooperation in the well-mixed limit

In a well-mixed volume II with N compartments of volume V^{I}, the number of exchange processes between II and I are N times more numerous than for the single-compartment case. We can then rewrite the transition rate for transferring to a compartment of type I as

$$w_{\mathrm{[BCB]^{II}}\to\mathrm{[BCB]^{I}}} = N\kappa_{\mathrm{[BCB]}}^{-}, \qquad w_{\mathrm{B^{II}}\to\mathrm{B^{I}}} = N\kappa_{\mathrm{B}}^{-}. \tag{6.105}$$

amounting to a net increase in uptake.

The competing degradation processes, occurring with rates $k_{\mathrm{B}}^{\emptyset}, k_{\mathrm{[BCB]}}^{\emptyset}$, are left untouched by the increase in compartments. Consequently, the losses incurred due to degradation are reduced with

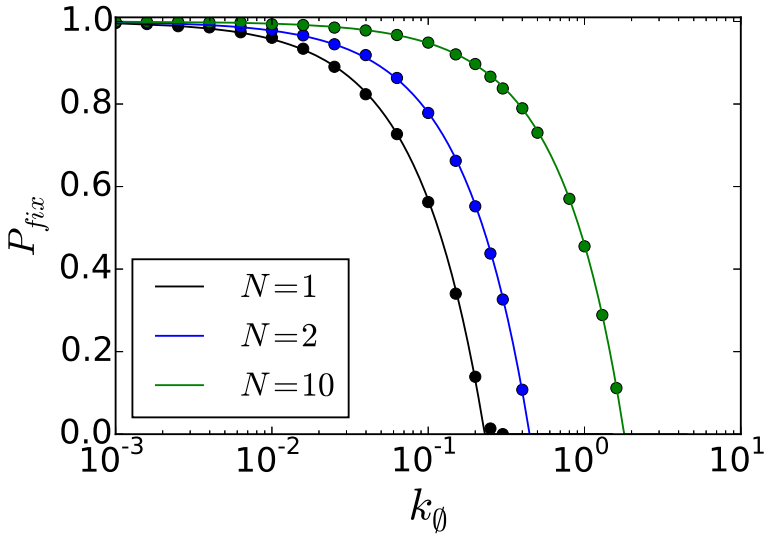


Figure 6.15: Fixation probability P_{fix} as function of degradation rate $k_\emptyset$, for N compartments. Other parameters remain fixed at values specified under Fig. 6.14. Lines correspond to the exact solution (Eq. 6.99), each circular dot corresponds to simulations as specified under Fig. 6.14, in which fixation was approximated by survival after 10000 transitions.

respect to the single-compartment case, since the relative probability to go from $[\mathrm{BCB}]^{\mathrm{II}}$ to $[\mathrm{BCB}]^{\mathrm{I}}$ is now increased, leading to a higher success rate p_{B} (Eq. (6.99)).

In Fig. 6.15, the exact expression for p_{B} is traced for $N = 2$ and $N = 10$, and compared with numerical simulations.

Cooperation without mixing

The ideal mixing approximation may not always be appropriate. In this section, we will instead advance a more general statistical argument.

Let us consider diffusion from a compartment α to a compartment β, with an absorbing boundary condition. This process is successfulwith a probability $p_{\alpha\to\beta} = p_e$. The compound returns to α without going through β with probability $p_{\alpha\to\alpha} = p_r$. With probability $p_\emptyset = 1 - p_r - p_e$ it diffuses away to infinity or is degraded. If $V^\alpha = V^\beta$ and the environment is isotropic, the reverse trajectories from β to α are equally likely, $p_{\alpha\to\beta} = p_{\beta\to\alpha} = p_e$.

A particle going from volume α to volume β, will only subsequently return to α with probability p_e. Considering returns from β to β, the total probability of returning to α becomes

$$\bar{p}_c = p_c + p_e^2 \sum_{k=0}^{\infty} p_c^k = p_c + \frac{p_e^2}{1-p_c}. \tag{6.106}$$

If α is the only reactive compartment, $\bar{p}_c$ must become larger than $1/2$ to have any chance of viable autocatalysis. If β is reactive as well, the probability π to end up either in α or in β is

$$\pi = p_c + p_e = \bar{p}_c + p_e \left(1 - \frac{p_e}{1-p_c}\right) \tag{6.107}$$

Let us now consider we have N equidistant compartments. Initially, $N-1$ compartments can be reached, each with probability p_e^N. A final return step can only be to the original compartment, happening with probability p_e^N, while in the meantime the particle can move around between the

$N-1$ other compartments. We can then write

$$\bar{p}_c^N = p_c^N + (N-1)(p_e^N)^2 \sum_{k=0}^{\infty} (p_c^N + (N-2)p_e^N) = p_c^N + \frac{(N-1)(p_e^N)^2}{1-p_c^N+(N-2)p_e^N}, \tag{6.108}$$

$$\pi^N = p_c^N + (N-1)p_e^N = \bar{p}_c^N + (N-1)p_e^N \left(1 - \frac{p_e^N}{1-p_c^N-(N-2)p_e^N}\right) \tag{6.109}$$

where p_c^N is now a return probability without traversing any of the other $N-1$ compartments, and p_e^N the probability of going to another compartment without passing through any compartments. Using the total probability

$$p_c^N + (N-1)p_e^N + p_\emptyset^N = 1, \tag{6.110}$$

we then obtain

$$\bar{p}_c^N = p_c^N + (N-1)p_e^N \frac{p_e^N}{p_e^N + p_\emptyset^N}, \tag{6.111}$$

$$\pi^N = p_c^N + (N-1)p_e^N = \bar{p}_c^N + (N-1)p_e^N \left(\frac{p_\emptyset^N}{p_e^N + p_\emptyset^N}\right). \tag{6.112}$$

Since $\pi^N > \bar{p}_c^N$, it follows that having reactive neighboring droplets is strictly better.

We can write an approximate correspondence between p_e and p_e^N, by considering all indirect trajectories through the $N-2$ nodes that do not touch α and β

$$p_e \approx p_e^N + (p_e^N)^2 \sum_{k=0}^{\infty} (N-2)(p_e^N)^k = p_e^N + p_e^N \frac{p_e^N}{1-(N-2)p_e^N}. \tag{6.113}$$

Eq. (6.113) can be rearranged to give a quadratic equation

$$(N-1)(p_e^N)^2 - [(N-2)p_e + 1]p_e^N + p_e = 0, \tag{6.114}$$

of its two solutions, the physical one corresponding to our model is

$$p_e^N = \frac{(N-2)p_e + 1 + \sqrt{(N-2)^2 p_e^2 - 2(N-1)p_e + 1}}{2(N-1)}. \tag{6.115}$$

In the absence of transport boundaries and for uniform degradation, such a correspondence becomes exact, and p_e^N is well approximated by p_e, as long as p_e^N and $(N-2)p_e^N$ are not too large. Under those same approximations, $\bar{p}_c^N = \bar{p}_c$, so that we can directly compare the

$$\pi^N - \pi = (N-1)p_e^N \left(\frac{p_\emptyset^N}{p_e^N + p_\emptyset^N}\right) - p_e \left(\frac{p_\emptyset}{p_e + p_\emptyset}\right) \tag{6.116}$$

Let $p_e^N = \varepsilon$, where $0 < \varepsilon \ll 1$. From Eq. (6.115) it follows $p_e^N \to p_e$. By Eq. (6.108), we have $p_c^N \to p_c$. From total probability (Eq. (6.110)) we then have $p_\emptyset^N \to p_\emptyset$, s.t.

$$\pi^N - \pi = (N-2)\varepsilon \left(\frac{p_\emptyset}{\varepsilon + p_\emptyset}\right). \tag{6.117}$$

To leading order, capture increases linearly with the number of surrounding compartments. This increase in capture can make autocatalysis viable that would not survive for a single compartment.

6.5.3 Network robustness against loss

So far, our solutions focused on environmental effects that could improve the survival of the BCB model reaction. Another factor we can consider for improvement is the reaction network itself: by being appropriately structured, the network dependence on the successful return of autocatalysts can be drastically lowered.

An instructive consideration is to avoid having a molecule that has to fetch itself, and instead have the molecule fetch an allocatalyst that can act within the compartment. A simple model we can construct for this situation uses a species F, to fetch a precatalyst P. Taken back to the compartment, P irreversibly reacts with reservoir species D, to form a catalyst C that converts a reservoir species X to more F.

$$\mathrm{F^I} \rightleftharpoons \mathrm{F^{II}}, \quad \mathrm{F^{II}} + \mathrm{P} \rightleftharpoons \mathrm{FP^{II}}, \tag{6.118}$$

$$\mathrm{FP^{II}} \rightleftharpoons \mathrm{FP^I}, \quad \mathrm{FP^I} \rightleftharpoons \mathrm{F^I} + \mathrm{P^I}, \tag{6.119}$$

$$\mathrm{F^{II}} \longrightarrow \emptyset, \quad \mathrm{FP^{II}} \longrightarrow \emptyset, \tag{6.120}$$

$$\mathrm{P^I} + \mathrm{D^I} \longrightarrow \mathrm{C^I}, \quad \mathrm{C^I} \longrightarrow \emptyset, \tag{6.121}$$

$$\mathrm{C^I} + \mathrm{X^I} \rightleftharpoons \mathrm{CX^I}, \quad \mathrm{CX^I} \rightleftharpoons \mathrm{C^I} + \mathrm{F^I}. \tag{6.122}$$

This leads to a network reminiscent of the simple network in Fig. 5.4, in which a composition of two allocatalytic reactions yields an autocatalytic reaction. This type of autocatalytic network motif is characteristic for GARD[6].

Let p_C denote the probability of a successful catalytic cycle for C, which corresponds to binding X and then producing one F molecule. The probability of an unsuccessful cycle $(1-p_\mathrm{C})$, corresponds to degradation of C.

Let P_n^C be the probability to perform a successful catalytic cycle n times, followed by degradation. Such a process yields a geometric distribution

$$P_n^\mathrm{C} = (1-p_\mathrm{C})p_\mathrm{C}^n \tag{6.123}$$

Typically, a C molecule produces $p_\mathrm{C}/(1-p_\mathrm{C})$ F molecules before being lost to degradation.

Let us denote p_F the probability that an F molecule successfully performs a fetch cycle, irreversibly releasing P (due to the low abundances of P^I and F^I, the rate of their reverse reactions can be considered negligible), which is then converted to C. The probability of fetching n catalysts in this way, is then

$$P_n^\mathrm{F} = (1-p_\mathrm{F})p_\mathrm{F}^n, \tag{6.124}$$

which is of the exact same functional form as Eq. (6.123): both describe catalytic processes (fetching or direct production) competing with degradation.

Taken together, we can define $\mathscr{P}_n^\mathrm{C}$, the probability that a catalyst C yields n new catalysts C, through the fetch molecules (transport catalyst) F it produced.

$$\mathscr{P}_n^\mathrm{C} = \sum_{s=0}^{\infty} P_s^\mathrm{C} \sum_{n_1,\ldots,n_s}^{\infty} \prod_{k=0}^{s} P_{n_k}^\mathrm{F} \delta_{n_1+\ldots+n_s}^n \tag{6.125}$$

The resulting expression was derived for a more general case in Sec. 6.3. Taking $p_\mathrm{A}=1$ in that derivation, we directly obtain:

$$\mathscr{P}_n^\mathrm{C} = \beta\alpha^n, \quad n \geq 1 \tag{6.126}$$

$$\mathscr{P}_0^\mathrm{C} = 1-\beta\frac{\alpha}{1-\alpha}. \tag{6.127}$$

with

$$\alpha = \frac{p_F}{1-p_C(1-p_F)}, \qquad \beta = \frac{p_C(1-p_C)(1-p_F)}{1-p_C(1-p_F)}. \tag{6.128}$$

We can then use the probability to go extinct within m rounds, d_m, as defined in Eq. (6.68), which in the limit $m \to \infty$ converges to d, and for which it was shown that

$$d = \frac{\beta}{\alpha - 1} + \frac{1}{\alpha} \tag{6.129}$$

which here simplifies to

$$d = \frac{1-p_C}{p_F} \tag{6.130}$$

At the threshold for guaranteed extinction, $d = 1$, we then have

$$p_F = 1 - p_C \tag{6.131}$$

An F molecule will on average have provided $p_F/(1-p_F)$ catalyst molecules, before degrading. Similarly, a catalyst will have provided $p_C/(1-p_C)$ fetch molecules F.

Overcoming extinction then requires that the fetch molecules F generated by one species of C, bring back (on average) more than one catalayst C:

$$\frac{p_F p_C}{(1-p_F)(1-p_C)} > 1. \tag{6.132}$$

Indeed, plugging in the threshold (Eq. (6.131)) for extinction $p_F = 1 - p_C$, a catalyst will replace itself, on average, with exactly one new catalyst. The fixation probability P_{fix} is then

$$P_{fix} = \begin{cases} 1 - d = \frac{p_F + p_C - 1}{p_F}, & p_F + p_C \geq 1, \\ 0, & p_F + p_C < 1, \end{cases} \tag{6.133}$$

The key point here is that to reach fixation, the internal catalyst C can compensate for excessive degradation of F, and vice versa. While the spatial diffusion of F may be perilous due to fluxes, diffusion and external degradation, the chemical milieu can be quite accommodating to C, thus ensuring it will produce a large amount of F before being lost.

6.5.4 Biological examples

In biology, all of the aforementioned phenomenology can be found. (Micro)organisms actively pursue nutrients, which may be poorly accessible at first.

A first example is Fe^{3+}. Iron is an essential nutrient, forming part of several key enzymes, such as ferredoxins used in oxidative phosphorylation. At pH 7, most Fe^{3+} forms a precipitate, and equilibrium concentrations are 10^{-18}M which makes it a key limiting nutrient[41]. Since its role is catalytic (Iron is essential to reach metabolic closure), the rationale of Sec 6.5.3. applies: as long as we are careful with the catalytic iron we have, we can sacrifice considerable resources to procure more. This is evident, when looking at the rich diversity of siderophores (literally: iron-carriers) that have evolved to capture extra iron[41].

Siderophores are relatively simple molecules, found particularly with microbia (e.g. fungi, bacteria), but some plants also excrete them (phytosiderophores). Over 300 siderophores have been identified [41], and some bacteria are capable of exploiting siderophores of other species (xenosiderophores). In the simple cooperation model in Sec. 6.5.2, it became clear that excretion of transport catalysts improves autocatalysis and nutrient acquisition on the collective level. Biological

agents may regulate themselves to deviate from such behavior. Siderophores have been the object of sociobiological and behavioral studies, to assess if producers of siderophores are being exploited by non-producers [42]. While some proof is found for such behavior, it should be remarked that this exploitation is not very detrimental to producers: siderophores that have moved far enough to be 'stolen', are typically too far away to appreciably diffuse back[41]. In addition, siderophores are relatively simple, small molecules, expected to have a low unit cost in metabolic terms.

Another important strategy, confinement, is especially important when valuable compounds are excreted. Decomposers (e.g. in the soil) have the arduous task to degrade a whole pastiche of inert biopolymer bonds (e.g. cellulose, lignin [43]), to penetrate towards more nutritious and easily degradable material. The degradation of these bonds requires highly specialized large proteins. Given the cost of producing such proteins, it is important that such a protein sticks around as long as possible. One confinement strategy to achieve this is by making ecto-enzymes: enzymes anchored to the surface.

More generally, degradation of nutritious polymers may occur extracellularly, outside of the surface, followed by the import the nutrients. To make sure degradation enzymes stick around, they can be trapped in polymer matrices characteristic of biofilms, which contain sites with a high affinity for these enzymes. In such a matrix, nutrients can flow through and are extracellularly digested, a sort of 'inverse fishnet'.

Exceptionally elegant cases of confinement are found for syntrophic (cross-feeding) organisms. The efficient coupling of syntrophic metabolism requires the organisms involved to exchange each other's nutrients at a minimal loss, which makes simple secretion a suboptimal solution. A considerable leap in efficiency is made, when a physical substrate confines these flows. Geobacter is known [44] to produce conductive wires to transport electrons to its partner. Other bacteria have been observed to build tubulin structures filled with ubiquinones, allowing for rapid hopping of electrons.

Varahan et al showed[45] that isogenic cells in yeast colonies can form specialized cell groups when exposed to low levels of glucose, with one group becoming gluconeogenic cells providing trehalose, which is consumed by cells utilizing a high pentose phosphate pathway. These metabolically complementary states are spatially organized and are a key to understanding its further propagation. It is suggested that such a resource strategy may maximize spatial expansion.

6.5.5 An afterthought

In this chapter and Ch.5, we have established a link between network structure and the forms of catalysis it can inhabit. These findings were then used to give a microscopic derivation of 'autocatalytic chemical evolution', along the lines of King [27] and Bagley et al[11]. We hope that this has opened new vistas for understanding autocatalysis and chemical evolution.

Work on prebiotic chemical evolution has thus far focused on single-pot chemistry. A notable exception to this trend is a work by Lemarchand and Jullien [40], in which autocatalytically assembling micelles and vesicles are considered, which can display all sorts of ecological interactions mediated by simple chemical reactions. We anticipate that such multicompartment considerations provides a rich

Our framework for stoichiometric autocatalysis provides a possible solution to a number of problems in chemical evolution: it does not require the autocatalysis to be formulated for a small subset of species, e.g. lipids, catalytic copolymers or for one environment. We may equally well consider a puddle, a rock pore, a slab of metal, a mineral surface, a contained gas, or all of them combined.

This makes autocatalysis a much less rare property. Autocatalysis forms the cornerstone of all proposed mechanisms in prebiotic chemical evolution. In any scenario, reducing the abundance of autocatalysis is something that should be done with great care. An interesting future direction

will be the further exploration of physical-chemical feedback mechanisms, to see what kind of feedbacks go beyond a stoichiometric description.

The problem of rare autocatalysis has previously only been considered in the single-pot context. In this situation, it was proposed that the problem can be overcome by having a large diversity of compounds. Insofar as stoichiometric autocatalysis is concerned, a large molecular diversity is indeed ideal to give an abundance of autocatalytic network motifs. As considered in this chapter, an equally essential problem is the question of viability: can an autocatalytic motif actually see the light of day?

Piling up all these diverse compounds in the same phase comes with clear drawbacks. These compounds can do many more things than forming autocatalytic sets: we should expect that catalysis of degradation, side-reactions and formation of trapping complexes grows much more rapidly. Such effects will quickly make many autocatalytic cycles unviable[16, 8]. This seems to lead to a paradox: either we have i) abundant autocatalytic motifs, but they are not viable, or we have ii) hospitable conditions for autocatalysis, but lack the diversity to make that property emerge.

It is this paradox that multicompartment may attempt to address. Spatially extended multi-compartment systems can accommodate many innovations and autocatalytic cycles. On a local level, however, the chemical diversity can be maintained at levels that are hospitable enough for autocatalysis. As noted in Ch. 5, there is inherently a larger space of autocatalytic reactions for the same set of compounds if we consider they can also selectively move to other, chemically distinct environments.

In multicompartment autocatalysis, innovations are not localized in one reactor. If locally no new autocatalytic cycles can be triggered, there is still a panoply of alternative locations where this may happen. Thereby, it sidesteps the obstacles that render a single diverse soup unviable. On the other hand, it still possesses the combinatorial diversity that motivated the single-pot approach.

Mechanistically, it was shown that autocatalysis leads to richer behavior than foreseen, with cooperation, syntrophy and parasitism emerging from small chemical networks. Such behavior suggests that the multicompartment approach may yield higher-order features which can be studied through the lens of ecology. Structurally, such a network evolution may be very different from the classical view of 'competing, self-replicating molecules'. Multicompartment networks can be tightly interlinked in their chemistry, interactions and survival. It is not clear whether a conception of individual evolving replicators that compete is an appropriate way of viewing such a system.

Insight in evolution often requires an appreciation of tradeoffs, an idea of the vulnerabilities that shape the fitness landscape, and what innovations overcome them. In Sec.8.4 and Ch.9, some simple ones have been considered. Network structure is an important determinant for the survival of an autocatalytic network. Confinement and collaboration are robust strategies that confer viability to networks, which may in turn provide fertile ground for new innovations.

Bibliography

Articles

[1] Thomas C Scott-Phillips et al. "The Niche Construction Perspective: A Critical Appraisal". In: *Evolution (N. Y).* 68.5 (2014), pp. 1231–43.

[2] Christophe Malaterre. "Chemical evolution and life". In: *Bio Web Conf.* 4 (2015), pp. 1–8.

[3] Marc Tessera. "Is pre-Darwinian evolution plausible ?" In: *Biol. Direct* 13 (2018), pp. 1–18.

[4] Daniel Segré et al. "Graded autocatalysis replication domain (GARD): kinetic analysis of self-replication in mutually catalytic sets". In: *Orig. Life Evol. Biosph.* 28 (1998), pp. 501–514.

[5] Daniel Segré, Dafna Ben-eli, and Doron Lancet. "Compositional genomes : Prebiotic information transfer in mutually catalytic noncovalent assemblies". In: *Proc. Natl. Acad. Sci. U.S.A.* 97.8 (2000), pp. 4112–4117.

[6] D Lancet, R Zidovetzki, and O Markovitch. "Systems protobiology: origin of life in lipid catalytic networks". In: *J.R.Soc.Interface* 15 (2018), p. 20180159.

[7] Stuart A. Kauffman. "Autocatalytic Sets of Proteins". In: *J. Theor. Biol.* 119 (1986), pp. 1–24.

[8] Vera Vasas et al. "Evolution before genes". In: *Biol. Direct* 7 (2012), pp. 1–14.

[9] Wim Hordijk, Jotun Hein, and Mike Steel. "Autocatalytic Sets and the Origin of Life". In: *Entropy* 12 (2010), pp. 1733–1742.

[10] Wim Hordijk, Stuart A Kauffman, and Mike Steel. "Required Levels of Catalysis for Emergence of Autocatalytic Sets in Models of Chemical Reaction Systems". In: *Int. J. Mol. Sci* 12 (2011), pp. 3085–3101.

[12] Wim Hordijk. "Autocatalytic Sets and RNA Secondary Structure". In: *J. Mol. Evol.* 84.4 (2017), pp. 153–158.

[13] Donna G. Blackmond. "An examination of the role of autocatalytic cycles in the chemistry of proposed primordial reactions". In: *Angew. Chemie - Int. Ed.* 48.2 (2009), pp. 386–390.

[14] L. E. Orgel. "Evolution of the genetic apparatus: A review". In: *J. Mol. Biol.* 38 (1968), pp. 381–393.

[15] Gerald F Joyce. "The antiquity of RNA-based evolution". In: *Nature* 418 (2002), pp. 214–221.

[16] E Szathmáry. "The origin of replicators and reproducors". In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361 (2006), pp. 1761–1776.

[17] Melvin Calvin. "Chemical Evolution and the Origin of Life". In: *Am. Sci.* 44.3 (1956), pp. 248–263.

[19] Stuart A. Kauffman. "Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems". In: *J. Cybern.* 1.1 (1971), pp. 71–96.

[20] Varun Giri and Sanjay Jain. "The Origin of Large Molecules in Primordial Autocatalytic Reaction Networks". In: *PLoS One* 7.1 (2012), e29546.

[21] Chandrakumar Appayee and Ronald Breslow. "Deuterium Studies Reveal a New Mechanism for the Formose Reaction Involving Hydride Shifts". In: *J. Am. Chem. Soc.* 136 (2014), pp. 3720–3723.

[22] Liang Cheng, Charles Doubleday, and Ronald Breslow. "Evidence for tunneling in base-catalyzed isomerization of glyceraldehyde to dihydroxyacetone by hydride shift under formose conditions". In: *Proc. Natl. Acad. Sci. U. S. A.* 112.14 (2015), pp. 4218–4220.

[23] Paul Erdös and Alfred Rényi. "On the Evolution of Random Graphs". In: *Publ. Math. Inst. Hungary. Acad. Sci.* 5 (1960), pp. 17–61.

[25] Vera Vasas, Eörs Szathmáry, and Mauro Santos. "Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life". In: *Proc. Natl. Acad. Sci. U.S.A.* 107.4 (2010), pp. 1470–1475.

[26] Nilesh Vaidya et al. "Spontaneous network formation among cooperative RNA replicators". In: *Nature* 491.7422 (Oct. 2012), pp. 72–77.

[27] G.A.M. King. "Recycling, Reproduction, and Life's Origins". In: *BioSystems* 15 (1982), pp. 89–97.

[28] Sanjay Jain and Sandeep Krishna. “Autocatalytic Sets and the Growth of Complexity in an Evolutionary Model”. In: *Phys. Rev. Lett.* 81.2 (1998), pp. 5684–5687.

[29] Sanjay Jain and Sandeep Krishna. “Emergence and growth of complex networks in adaptive systems”. In: *Comput. Phys. Commun* 121-122 (1999), pp. 116–121.

[30] Sanjay Jain and Sandeep Krishna. “A model for the emergence of cooperation, interdependence, and structure in evolving networks”. In: *Proc. Natl. Acad. Sci. U.S.A.* 98.2 (2001), pp. 543–547.

[33] F Schlögl. “Chemical reaction models for non-equilibrium phase transitions”. In: *Z. Phys.* 253.2 (1972), pp. 147–161.

[34] Leslie E Orgel. “The implausibility of Metabolic Cycles on the prebiotic earth”. In: *Plos Biol.* 6.1 (2008), pp. 5–13.

[35] E. Szathmary. “The evolution of replicators”. In: *Philos. Trans. R. Soc. B Biol. Sci.* 355.1403 (2000), pp. 1669–1676.

[36] Alex Blokhuis, David Lacoste, and Pierre Gaspard. “Reaction kinetics in open reactors and serial transfers between closed reactors”. In: *J. Chem. Phys.* 148 (2018), p. 144902.

[39] D T Gillespie. “Exact Stochastic Simulation of Coupled Chemical-reactions”. In: *J. Phys. Chem* 81 (1977), p. 2340.

[40] A Lemarchand and L Jullien. “Competition and Symbiosis in a Chemical World”. In: *J. Phys. Chem. B* 108 (2004), pp. 11782–11791.

[41] Gabriel E Leventhal, Martin Ackermann, and Konstanze T Schiessl. “Why microbes secrete molecules to modify their environment : the case of iron-chelating siderophores”. In: *J. R. Soc. Interface* 16 (2019), p. 20180674.

[42] Elena Butaité et al. “Siderophore cheating and cheating resistance shape competition for iron in soil and freshwater Pseudomonas”. In: *Nat. Commun.* 8 (2017), p. 414.

[44] Kazuya Watanabe. “Microbial interspecies interactions : recent findings in syntrophic consortia”. In: *Front. Microbiol.* 6.May (2015), p. 477.

[45] Sriram Varahan et al. “Metabolic constraints drive self-organization of specialized cell groups”. In: *Elife* 8 (2019), e46735.

Books

[18] Stuart A. Kauffman. *A world beyond physics: the emergence & evolution of life*. Oxford University Press, 2019.

[24] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press, 1993.

[31] S Karlin and H M Taylor. *A first course in stochastic processes*. New York: Academic Press, 1975.

[32] Eric Smith and Harold Morowitz. *The origin and nature of life on earth: The Emergence of the Fourth Geosphere*. Cambridge, UK: Cambridge University Press, 2016.

[37] Francis A. Carey and Richard J. Sundberg. *Advanced Organic Chemistry Part A: Structure and Mechanisms*. New York: Springer, 2006.

[38] T.W. Graham Solomons. *Organic Chemistry*. Sixth edit. New York: John Wiley & Sons, 1995.

[43] F. Blaine Jr Metting, ed. *Soil microbial ecology: aplications in agricultural and environmental management*. New York: Wiley, 1992.

7. Prebiotic Polymerization Scenarios

7.1 Prebiotic Polymer Scenarios

A recurring element in origins-of-life scenarios is that, sooner or later, a repertoire of functional polymers is formed. The RNA-world hypothesis, for example, proposes that RNA molecules were one of the first functional polymers, by exploiting their capacities for catalysis and information storage [1, 2]. In particular, the ligation of monomers to form polymers (polymerization) is an activity that some of such functional polymers are thought to have. Similar hypotheses exist for e.g. amino acids [3, 4, 5] and XNA's [6, 7, 8, 9]. The scope can be widened a little further when considering Krishnamurthy's view that today's biochemistry may have been selected from many prebiotic precursors that have not been preserved[10]. An interesting alternative form of information storage was proposed by D.Segré et al. [11], where autocatalysis and the population composition of lipid membranes is proposed as a means of information storage. This idea has seen implementation in an RNA-polymer system [12].

In those polymer scenarios, one generally wishes functional polymers to provide: i) a catalytic repertoire, ii) heredity and iii) evolvability. Condition i) seems to be readily fulfilled by DNA [13], RNA [14] and various XNAs [15] as well as oligopeptides. Condition ii) however, is far from trivial: making persistent copies of long polymer sequences without proofreading mechanisms and polymerases seems to be out of the question. At the same time, even rudimentary polymerases are prohibitively long, making them unlikely to appear in a random polymer pool and, without error correction machinery or multilevel selection[16, 17] hard to maintain (Eigen's Paradox) [18]. For evolvability to come into play, this heredity seems to be an essential requirement.

Since heredity is such a central issue, scenarios often introduce new strategies to make this property emerge. Many of these strategies boil down to either of the following two situations:
i) the generation of wider length distributions, with the hope of finding functional machinery to obtain heredity, such as a polymerase [19, 20, 21, 22],
ii) a chemistry promoting the replication of particular subsets of sequence space (e.g. in reflexively autocatalytic networks [4, 23, 24]).

It is important to appreciate that these are nontrivial things to achieve and subsequently evolve,

especially from a simple starting point. The idea that the origins of life required extremely rare events to yield sufficiently functional self-replicating objects can be traced back to the ideas of Oparin and Haldane [3], who imagined that a rare peptide had to emerge. With the discovery of DNA, concepts like sequence complementarity and templates made its entry. With the discovery of RNA ribozymes, a species was identified that has, in principle, the ingredients to facilitate catalysis and self-replication [1]. This self-replication and catalysis is not trivial however, it is postulated to require complex functional species. Like the Oparin-Haldane scenario, RNA and other polymer scenarios envision that a long time must pass in which prebiotic chemistry accumulates monomers and chemical activation capacities with which it starts forming (quasi-)random polymer species. Eventually (Joyce's timeline suggests it takes up to 200 million years [25]), such a process is thought to lead to the supposedly long and complex functional species that provide the transit to e.g. a pre-RNA world and evolution of self-replicating polymer sequences.

Such prebiotic polymer scenarios have obvious thermodynamic challenges to overcome: for many biomonomers of interest, polymerization is thermodynamically disfavored. The repeated assembly of a small subset of sequences goes strongly against the entropic drive to erase this information by making sequences as random as possible. With the framework of nonequilibrium thermodynamics and stochastic thermodynamics, these thermodynamic aspects can now be captured. Free energy and material resources are inherently limited, providing clear constraints on what a scenario can do, and how much of it can be done.

In the upcoming sections, various aspects of polymers, length distributions, sequence space and nonequilibrium thermodynamics will be discussed. In Sec. 7.2 we will build up a framework for recombination of polymers in closed systems, providing a thermodynamic description of sequence and length distributions. The contents of this section are published in Ref. [26] and we will largely follow the lines of the original publication.

In Sec. 7.3, recombination reactions are opened to exchange with reservoirs, leading to nonequilibrium regimes in which efficient net polymerization can be achieved. In Sec. 7.4, a joint published work (Ref. [27]) with the groups of D. Baum and N. Lehman is discussed, in which experiments with RNA adsorption on mineral surfaces are described by a toy model with multisite adsorption. Contrary to what has been argued in the literature, we show that minerals preferentially adsorb larger polymers, which provides another pathway to the acquisition of longer polymers.

In Sec. 7.5, we provide a thermodynamically explicit description of some common prebiotic polymer scenarios using ligation, fragmentation and activation. The key step of activation is needed to have a thermodynamically consistent description, and the relative strengths of ligation, activation and fragmentation can lead to ligation-limited and activation-limited regimes. A large class of models in the literature are shown to work in a strongly ligation-limited regime. The framework quantifies the rate of generation of random sequences for every length for arbitrary polymer distributions and links it to the entropy production for this process. Coupled with the length distribution, we can then establish a mean time and free energy cost for finding a particular species, providing an absolute lower bound on random dissipative searching through sequence space, irrespective of e.g. catalysis, and this leads to a tradeoff between dissipation rate, time and the complexity of a species.

7.2 Recombination of polymers in closed systems

In this section, we will study some properties of 'recombination' reactions. The technical discussions will largely follow Ref. [26], which can be consulted for further detail.

7.2.1 Recombination reactions

The reactions referred to as recombinations are depicted in Fig 7.1. In the first mechanism, two polymer chains are interchanged, and we refer to it as 'chain exchange'. It is equivalent to 'metathesis' in synthetic organic chemistry, with examples such as disulfide metathesis and olefin metathesis. The latter has become such a valuable tool for chemists that it was awarded the 2005 Nobel Prize in Chemistry.[28]

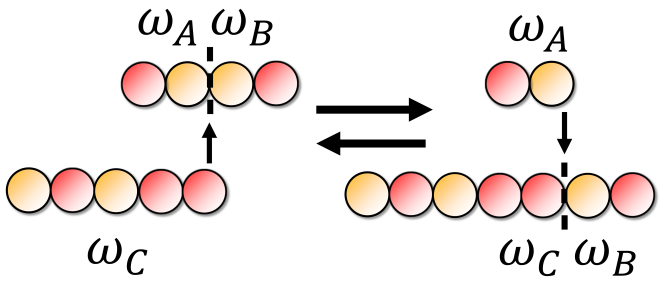


Figure 7.1: Representation of Attack-Exchange reaction: $\omega_A\omega_B + \omega_C \rightleftarrows \omega_C\omega_B + \omega_A$ for the case that two monomer types are present: $m = 2$.

The second scheme, 'attack-exchange' involves one reactant picking up the group to be transferred from the other. Such a reaction is preceded by a 'trans-' prefix in synthetic chemistry (e.g. transesterification, transamination, transamidation, etc.). These reactions are key players in dynamic covalent chemistry and are applied in novel materials, such as vitrimers[29, 30], which combine the properties of thermosets and thermoharders. Modern biology employs transesterification in splicing pathways in which terminal hydroxyl groups attack phosphodiesters[31](see Fig. 7.2).

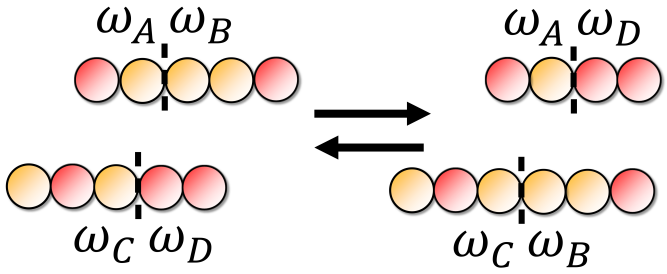


Figure 7.2: Representation of Chain-Exchange reaction: $\omega_A\omega_B + \omega_C\omega_D \rightleftarrows \omega_A\omega_D + \omega_C\omega_B$.

7.2.2 In prebiotic scenarios

In RNA, a plethora of pathways have been laid bare for recombination of small strands [32], in the absence of any enzyme. Such reactions do not require an energy source or abundant monomer supplies and they are sequence-selective. Such reactions may be quite slow[33].

A synthetic ribozyme catalyzing transesterification, Azoarcus ribozyme, was developed [34], of which a version was developed that could be cut in pieces that self-assemble in an active, noncovalent complex. This noncovalent complex can in turn recombine bonds in other noncovalent complexes to yield a covalent complex (plus small strands as leaving groups). By changing sequence specificities, cooperative RNA networks were assembled [12].

It has also been argued that such reactions may have provided 'sexual evolution' of genetic polymers very early on[35]. In modern genetics, crossing over of genetic material in chromosomes is a more elaborate version of this, whereas template-directed polymerization is more similar to asexual cloning.

7.2.3 A note on notation

In the following, it will be advantageous to introduce a specific notation to describe the evolution of sequences according to these reactions. Monomer sequences are considered to be directional, as in the case of nucleic acids which have a distinct 5' and 3' end. Two subsequent sequences will be noted using a product notation $\omega\omega' = \omega_1\omega_2...\omega_l\omega'_1\omega'_2...\omega'_{l'}$, which is used for the addition of two chains. An inverse sequence is defined as a sequence that is removed, either from the front or from the back, by placing the inverse either in front or on the back of a sequence. $\omega\omega'^{-1} = \omega_1\omega_2...\omega_q$. We define a length operator as $|.|$, which counts the number of elements in a sequence.

With this notation, the Attack-Exchange may be written:

$$\omega_A\omega_B + \omega_C \rightleftarrows \omega_C\omega_B + \omega_A, \tag{7.1}$$

In a framework of mass-action kinetics for a single well-mixed reactor (Sec. SUBSEC: massacwellmix), reaction rates can be written as:

$$w^{\omega_A\omega_B}_{\omega_C} = k^{\omega_A\omega_B}_{\omega_C} c_{\omega_A\omega_B} c_{\omega_C}, \tag{7.2}$$
$$w^{\omega_C\omega_B}_{\omega_A} = k^{\omega_C\omega_B}_{\omega_A} c_{\omega_C\omega_B} c_{\omega_A}, \tag{7.3}$$

where w is the reaction rate and k the corresponding rate constant, which can, in principle, be sequence dependent. The concentration $c_\omega = N_\omega/V$ is the number of polymers of sequence ω, N_ω, per unit of volume V. In anticipation of the link we will establish with a microscopic fluctuating description, N_ω is used as a mean-field abundance, in contrast with the true number of molecules n_ω used in the rest of the manuscript.

Similarly, the Chain-Exchange reaction drawn in Fig. 7.2 can be written as:

$$\omega_A\omega_B + \omega_C\omega_D \rightleftarrows \omega_C\omega_B + \omega_A\omega_D, \tag{7.4}$$

to which we attribute mass-action rates:

$$\begin{aligned} w^{\omega_A\omega_B}_{\omega_C\omega_D} &= k^{\omega_A\omega_B}_{\omega_C\omega_D} c_{\omega_A\omega_B} c_{\omega_C\omega_D} \\ w^{\omega_A\omega_D}_{\omega_C\omega_B} &= k^{\omega_A\omega_D}_{\omega_C\omega_B} c_{\omega_A\omega_D} c_{\omega_C\omega_B}. \end{aligned} \tag{7.5}$$

When the forward and backward rate constants $k^{\omega_A\omega_B}_{\omega_C\omega_D}$ and $k^{\omega_A\omega_D}_{\omega_C\omega_B}$ are equal, the transformation is not accompanied by a net change in standard free energy ΔG°.

An important constraint for both reactions of Eq. (7.1) and Eq. (7.4) is that we exclude the formation of any species of zero length. This means that the total number of chains $N = \sum_\Omega N_\Omega$ is a conserved quantity for both dynamics. In other words, there is a minimum length of chains $l_{min} = 2$ for Chain-Exchange reactions while $l_{min} = 1$ for Attack-Exchange reactions. In addition, in both exchange reactions, the first monomer is never displaced, which leads to a conservation law for the composition of the first monomer. For Chain-exchange, such a law also exists for terminal monomers, because they always remain in a terminal position.

7.2.4 Equilibrium Thermodynamics of Recombination in closed systems

In the following sections, a thermodynamic framework will be developed to describe the dynamics of a polymer mixture undergoing Chain-Exchange reactions in a closed system. The calculations for

the Attack-Exchange reaction are equivalent in spirit, and any differences between their descriptions will be elaborated on when they come up.

We assume that the mixture contains m different monomer types $\{0,1,2...m\}$, with $m > 1$ so that polymer sequences can be defined. In our modeling of recombination thermodynamics, we do not explicitly describe the solvent. As will be shown, this is appropriate because recombination preserves the total number of chains: $N = \sum_\Omega N_\Omega$. In Sec. 7.5, hydrolysis reactions in irreversible polymerization will be considered, which requires an explicit description of the solvent. The case of reversible polymerization was treated in Ref. [36].

We define the polymer fraction of a sequence Ω as:

$$y_\Omega = \frac{N_\Omega}{N}, \tag{7.6}$$

which obeys the normalization condition $\sum_\Omega y_\Omega = 1$. We assume that the solution is sufficiently ideal and thus the chemical potentials of all present species follow the form:

$$\mu_\Omega = \mu_\Omega^\circ + k_B T \ln y_\Omega, \tag{7.7}$$

where T is the temperature. The enthalpy of the polymers in solution can be expressed in terms of their standard free enthalpies h_Ω° as

$$H = \sum_\Omega N_\Omega \, h_\Omega^\circ. \tag{7.8}$$

Likewise, the entropy can be defined in this manner,

$$S = \sum_\Omega N_\Omega (s_\Omega^\circ - k_B \ln y_\Omega), \tag{7.9}$$

where s_Ω° represents the internal contribution of the entropy associated with other degrees of freedom different from Ω and not described here. We will also use the system entropy per chain $\mathscr{S}$ defined as:

$$\mathscr{S} = \frac{S}{N} = \sum_\Omega y_\Omega (s_\Omega^\circ - k_B \ln y_\Omega), \tag{7.10}$$

Let us define $G = H - TS$ as the Gibbs free energy. Using $\mu_\Omega = h_\Omega - T s_\Omega$, we find

$$G = \sum_\Omega N_\Omega \mu_\Omega = \sum_\Omega N_\Omega (\mu_\Omega^\circ - k_B T \ln y_\Omega). \tag{7.11}$$

In the remainder of this section we set $k_B = 1$ to simplify the notation.

7.2.5 Non-equilibrium Thermodynamics of Recombination in Closed Systems

Let us now consider the non-equilibrium description of recombination. The kinetic rate equation for the concentration of chains with sequence Ω is:

$$\dot{N}_\Omega = \sum_{\omega_A = \Omega \omega_B^{-1}} \sum_{\omega_C} \sum_{\omega_D} \left[w^{\omega_A \omega_D}_{\omega_C \omega_B} - w^{\omega_A \omega_B}_{\omega_C \omega_D} \right]. \tag{7.12}$$

The kinetic constant can depend on the exact sequences and on the sites of splitting. The Chain-Exchange reaction exchanges chemical bonds between subsequences of nonzero length. As such, the set of subsequences we consider cannot be empty ($\omega \neq \emptyset$) and a total sequence is at least of length 2. For convenience, we choose to make this instruction implicit.

The second term is equivalent to the back reaction of the first term. When summing over all possible sequences Ω, the first sequence sum turns into a sum over subsequences ω_A and ω_B (all distinct ordered pairs (ω_A, ω_B) are generated):

$$\sum_{\Omega} \sum_{\omega_A = \Omega \omega_B^{-1}} = \sum_{\omega_A} \sum_{\omega_B}, \tag{7.13}$$

which generates the symmetry $\sum_\Omega \dot{N}_\Omega = -\sum_\Omega \dot{N}_\Omega$. This of course implies again the conservation of the number of chains $\sum_\Omega \dot{N}_\Omega = 0$. The entropy production rate Σ of an ensemble of chemical reactions, assumed to be elementary (there should be no hidden chemical reactions) takes the form [37]):

$$\Sigma = \sum_k (w_k^+ - w_k^-) \ln\left(\frac{w_k^+}{w_k^-}\right) \geq 0, \tag{7.14}$$

where w_k^+, w_k^- are respectively forward and backward reaction rates of the kth reaction. In the specific case of Chain-Exchange reactions, this becomes:

$$\Sigma = \frac{1}{4} \sum_\Lambda \left[w^{\omega_A \omega_D}_{\omega_C \omega_B} - w^{\omega_A \omega_B}_{\omega_C \omega_D} \right] \ln\left(\frac{w^{\omega_A \omega_B}_{\omega_C \omega_D}}{w^{\omega_A \omega_D}_{\omega_C \omega_B}} \right), \tag{7.15}$$

where the sum is carried out over Λ, which represents an arbitrary set of four sequences of the form $\{\omega_A, \omega_B, \omega_C, \omega_D\}$. The factor 4 can be understood as the cardinal of a discrete group $\mathscr{G}$ acting on elements of Λ. This group contains the following 4 elements: $\mathscr{G} = \{I, \chi, \pi, \rho\}$, where I is the identity, χ presents the exchange $\omega_A \to \omega_C$, π the exchange $\omega_B \to \omega_D$, and ρ the combined exchange $\omega_A \to \omega_C$ and $\omega_B \to \omega_D$. Similarly, for Attack-Exchange the relevant group $\mathscr{H}$ contains instead the elements: $\mathscr{H} = \{I, \chi\}$. Since the cardinal of $\mathscr{H}$ is 2 instead of 4 for $\mathscr{G}$, the equivalent of equation (7.15) for Attack-Exchange should contain a factor 2 in the place of the factor 4. In Sec. 7.3.1, such symmetry considerations will be derived more generally, using the stoichiometric matrix formalism.

Detailed balance should hold at equilibrium, which provides the following relation:

$$k^{\omega_A \omega_B}_{\omega_C \omega_D} y^{eq}_{\omega_A \omega_B} y^{eq}_{\omega_C \omega_D} = k^{\omega_A \omega_D}_{\omega_C \omega_B} y^{eq}_{\omega_A \omega_D} y^{eq}_{\omega_C \omega_B}. \tag{7.16}$$

Then, the condition $\Delta\mu = 0$ with the detailed balance condition (7.16) leads to:

$$T \ln\left(\frac{y^{eq}_{\omega_A \omega_B} y^{eq}_{\omega_C \omega_D}}{y^{eq}_{\omega_C \omega_B} y^{eq}_{\omega_A \omega_D}} \right) = -\Delta\mu^\circ_{\omega_A \omega_B, \omega_C \omega_D} = \mu^\circ_{\omega_A \omega_D} + \mu^\circ_{\omega_C \omega_B} - \mu^\circ_{\omega_C \omega_D} - \mu^\circ_{\omega_A \omega_B}. \tag{7.17}$$

Combining this relation with detailed balance (7.16), one obtains:

$$T \ln\left(\frac{k^{\omega_C \omega_D}_{\omega_A \omega_B}}{k^{\omega_A \omega_D}_{\omega_C \omega_B}} \right) = -\Delta\mu^\circ_{\omega_A \omega_B, \omega_C \omega_D}. \tag{7.18}$$

When the forward and backward rates are equal, the reaction is neutral in terms of free energy $\Delta\mu^\circ = 0$. This means standard entropy and enthalpy compensate each other, since $\Delta h^\circ = T\Delta s^\circ$.

If we now calculate the time evolution of the enthalpy H, we obtain:

$$\frac{dH}{dt} = \sum_\Lambda \left[w^{\omega_A \omega_D}_{\omega_C \omega_B} - w^{\omega_A \omega_B}_{\omega_C \omega_D} \right] h^\circ_{\omega_A \omega_B} = \frac{1}{4} \sum_\Lambda \left[w^{\omega_A \omega_D}_{\omega_C \omega_B} - w^{\omega_A \omega_B}_{\omega_C \omega_D} \right] \Delta h^\circ_{\omega_A \omega_B, \omega_C \omega_D}, \tag{7.19}$$

where we used the symmetry group $\mathscr{G}$ to write the evolution in single-reaction enthalpy changes:

$$\Delta h^\circ_{\omega_A \omega_B, \omega_C \omega_D} = h^\circ_{\omega_A \omega_B} + h^\circ_{\omega_C \omega_D} - h^\circ_{\omega_A \omega_D} - h^\circ_{\omega_C \omega_B}. \tag{7.20}$$

Similarly, for the entropy we obtain:

$$\frac{dS}{dt} = \frac{1}{4}\sum_{\Lambda}\left[w^{\omega_A\omega_D}_{\omega_C\omega_B} - w^{\omega_A\omega_B}_{\omega_C\omega_D}\right]\left[\Delta s^{\circ}_{\omega_A\omega_B,\omega_C\omega_D} - \ln\left(\frac{y_{\omega_A\omega_B}y_{\omega_C\omega_D}}{y_{\omega_A\omega_D}y_{\omega_C\omega_B}}\right)\right]. \tag{7.21}$$

We can combine the equations (7.17), (7.19) and (7.21) to get:

$$\frac{dG}{dt} = \frac{T}{4}\sum_{\Lambda}\left[w^{\omega_A\omega_B}_{\omega_C\omega_D} - w^{\omega_A\omega_D}_{\omega_C\omega_B}\right]\ln\left(\frac{y_{\omega_C\omega_B}y_{\omega_A\omega_D}y^{eq}_{\omega_A\omega_B}y^{eq}_{\omega_C\omega_D}}{y^{eq}_{\omega_C\omega_B}y^{eq}_{\omega_A\omega_D}y_{\omega_A\omega_B}y_{\omega_C\omega_D}}\right) \tag{7.22}$$

Using detailed balance (7.16) into Eq (7.22), one recovers the previous expression defined in Eq (7.15) for the entropy production rate Σ:

$$-\frac{1}{T}\frac{dG}{dt} = \Sigma = -\sum_{\Omega}\dot{N}_{\Omega}\ln\left(\frac{N_{\Omega}}{N^{eq}_{\Omega}}\right) \geq 0. \tag{7.23}$$

Since $G = H - TS$, this equation is equivalent to $\dot{S} = \Sigma + \dot{H}/T$, which expresses the second law of thermodynamics for a closed system. As expected, the heat released by the system into the environment Q is the change of enthalpy $Q = \Delta H$. Equation (7.23) is important to guarantee that the chemical system reaches a unique equilibrium state on long times [36].

7.2.6 Decomposition of the entropy production

Here, we split the entropy production of the polymer mixture into two contributions, where the first one represents the contribution of the various polymer lengths, while the second one represents that of their sequences. Using Eqs. (7.6), (7.23), we can rewrite the entropy production rate Σ in terms of polymer fractions:

$$\Sigma = -N\frac{d}{dt}\sum_{\Omega}y_{\Omega}\ln\left(\frac{y_{\Omega}}{y^{eq}_{\Omega}}\right) = -N\frac{d}{dt}\sum_{\Omega}y_{\Omega}\left(\frac{\mu^{\circ}_{\Omega}}{T} + \ln y_{\Omega}\right). \tag{7.24}$$

Since the polymer fractions y_{Ω} for all sequences Ω are normalized, y_{Ω} can be interpreted as the probability to observe a chain of sequence Ω when a polymer is drawn at random from the mixture. Furthermore, since the polymer of sequence Ω has only one possible length, namely $l = |\Omega|$, that probability to observe a polymer with sequence Ω can be denoted equivalently $P_{\Omega,l}(t)$ because the length is a redundant variable.

At any time t, we have therefore the identification

$$y_{\Omega}(t) = P_{\Omega,l}(t). \tag{7.25}$$

To proceed, we then factorize $P_{\Omega,l}(t)$ in the following way:

$$P_{\Omega,l}(t) = Y_l(t)\, U_{l,\Omega}(t), \tag{7.26}$$

with $Y_l(t)$ the probability distribution of polymer length at time t, and $U_{l,\Omega}(t)$ the conditional probability distribution of the sequence, conditional on the length l. The distributions Y_l and $U_{l,\Omega}$ are normalized: $\sum_l Y_l(t) = 1$, and $\sum_{\Omega} U_{l,\Omega}(t) = 1$ provided the sum is restricted to all chains which have a length l.

Using Eqs. (7.24)-(7.26), we deduce a splitting of the entropy production rate into three contributions:

$$\Sigma = -N\frac{d}{dt}\left[\sum_l Y_l \ln Y_l + \sum_{\Omega,l} Y_l U_{l,\Omega}\ln U_{l,\Omega} + \sum_{\Omega,l} Y_l U_{l,\Omega}\frac{\mu^{\circ}_{\Omega}}{T}\right]. \tag{7.27}$$

The various terms in this decomposition are:

- The first term: $\sum_l Y_l \ln Y_l$ represents the disorder in the length distribution Y_l (or length entropy).
- The second term: $\sum_{\Omega,l} Y_l U_{l,\Omega} \ln U_{l,\Omega}$ represents the disorder in the distribution of sequences (or sequence entropy). Importantly, this term is weighted by the length distribution Y_l and therefore introduces a coupling between length and sequence distributions. As a result, one expects that the dominant contribution to this sequence entropy will come from short sequences.
- The final contribution: $\sum_{\Omega,l} Y_l U_{l,\Omega}\ \mu^\circ_\Omega/T$ comes from the standard free energy change of each species. If we choose μ°_Ω such that our reactions are energetically neutral: $\Delta\mu^\circ = \mu^\circ_{\omega_A\omega_B} + \mu^\circ_{\omega_C\omega_D} - \mu^\circ_{\omega_C\omega_B} - \mu^\circ_{\omega_A\omega_D} = 0$, this term vanishes.

The latter term can be split further into two using $\mu^\circ = h^\circ - Ts^\circ$. Two terms will appear, $\sum_{\Omega,l} Y_l U_{l,\Omega}\ h^\circ_\Omega$, which corresponds to the heat exchanged with the surrounding medium and $\sum_{\Omega,l} Y_l U_{l,\Omega}\ s^\circ_\Omega$ which corresponds to an internal entropy contribution to Σ.

In the work of Andrieux and Gaspard [38], a sequence decomposition of entropy production was performed to model the thermodynamics of copolymerization of a single polymer. The single polymer grows by monomer addition, leading rapidly to steady-state growth statistics with only a small contribution to the entropy production.

Eq. (7.26) is similar in spirit, but it applies to ensembles of polymers performing recombination. There is no overall growth and here the length distribution is an important part of the entropy production.

Given an initial distribution $Y^I_l, U^I_{l,\Omega}$ and final distribution $Y^F_l, U^F_{l,\Omega}$, the total entropy production per chain $\Delta\mathscr{S}_{tot}$ in that transformation follows by integrating (7.27).

$$\Delta\mathscr{S}_{tot} = \sum_l \left(Y^I_l \ln Y^I_l - Y^F_l \ln Y^F_l\right) + \sum_{\Omega,l} \left(Y^I_l U^I_{l,\Omega} \ln U^I_{l,\Omega} - Y^F_l U^F_{l,\Omega} \ln U^F_{l,\Omega}\right) + \sum_{\Omega,l} \left(Y^I_l U^I_{l,\Omega} - Y^F_l U^F_{l,\Omega}\right) \frac{\mu^\circ_\Omega}{T}. \tag{7.28}$$

For arbitrary distributions, this can be interpreted as the integration of the entropy during a time-dependent protocol connecting them, or a particular kinetic realization.

7.2.7 Stochastic Thermodynamics Framework

The previous section relied on mass action laws and kinetic rate equations, which are appropriate in the thermodynamic limit when the number of chains $N \to \infty$. In a small system where fluctuations matter, a different approach is needed based on Stochastic Thermodynamics [39, 40, 41]. We define a state $\mathbf{n} = \{n_{\Omega_1}, n_{\Omega_2}, n_{\Omega_3}.....\}$, as a vector containing the numbers of each polymer (distinguished by their sequence and length) present in the system. The probability to be in a given state $\mathbf{n}$, $P(\mathbf{n})$, obeys a master equation [42] :

$$\frac{dP(\mathbf{n})}{dt} = \sum_{\mathbf{n}'} [W_{\mathbf{n}'\to\mathbf{n}} P(\mathbf{n}') - W_{\mathbf{n}\to\mathbf{n}'} P(\mathbf{n})], \tag{7.29}$$

where $W_{\mathbf{n}\to\mathbf{n}'}$ is the transition rate to jump from $\mathbf{n}$ to $\mathbf{n}'$. Given the size of the sequence space and the corresponding reaction network, this equation quickly becomes analytically intractable. Nevertheless, we can derive some useful results from it.

It is important to appreciate that the states $\mathbf{n}$ have an internal degeneracy $z(\mathbf{n})$, which follows from all the allowed permutations among species in that state:

$$z(\mathbf{n}) = \frac{N!}{n_{\Omega_1}! n_{\Omega_2}! .. n_{\Omega_n}! ...} = \frac{N!}{\prod_\Omega (n_\Omega!)}. \tag{7.30}$$

Fundamentally, this is a necessary procedure of coarse-graining[43] to acquire the entropy that is convenient for our description (see also Sec. 4.1.1), in which we treat the same types of molecules as indistinguishable.

The analogues of the ensemble averaged number of polymers of sequence Ω, N_Ω and of the entropy S introduced in the previous section are the stochastic particle number n_Ω and the stochastic entropy s. The connection between the two descriptions is that:

$$N_\Omega = \langle n_\Omega \rangle, \tag{7.31}$$
$$S = \langle s \rangle, \tag{7.32}$$

where the average is taken with respect to the distribution $P(\mathbf{n})$. Now, the expression of the stochastic entropy s is [44]

$$s(\mathbf{n}) = -\ln P(\mathbf{n}) + \ln z(\mathbf{n}) + s^\circ(\mathbf{n}), \tag{7.33}$$

where the first term on the right hand side gives after averaging over the distribution of $\mathbf{n}$ the Shannon entropy of that distribution, the second term is the contribution of the degeneracy while the last term is internal entropy coming from non-described molecular degrees of freedom. The precise definition of that last term is

$$s^\circ(\mathbf{n}) = \sum_\Omega n_\Omega s^\circ_\Omega, \tag{7.34}$$

in terms of s°_Ω, the intensive standard entropy of formation introduced in Eq. (7.10).

Assuming the reaction $\mathbf{n} \to \mathbf{n}'$ is elementary (i.e. the two vectors differ by only one recombination reaction among two of their components), the detailed balance condition is:

$$\frac{W_{\mathbf{n}\to\mathbf{n}'}}{W_{\mathbf{n}'\to\mathbf{n}}} = \frac{z(\mathbf{n}')}{z(\mathbf{n})} \exp(-\beta \Delta\mu^\circ), \tag{7.35}$$

where $\beta = 1/T$ and $\Delta\mu^\circ$ is the chemical potential difference of the elementary exchange reaction introduced in the previous section. We recall that the latter may be split into $\Delta\mu^\circ = \Delta h^\circ - T\Delta s^\circ$.

In the absence of degeneracy, the ratio $\ln W_{\mathbf{n}\to\mathbf{n}'}/W_{\mathbf{n}'\to\mathbf{n}}$ would correspond to the stochastic heat transferred from the system to the reservoir during that transition. However, in present case, due to the degeneracy, the correct definition of the stochastic heat, δq is:

$$-\beta\,\delta q = \ln \frac{W_{\mathbf{n}\to\mathbf{n}'}}{W_{\mathbf{n}'\to\mathbf{n}}} - \ln \frac{z(\mathbf{n}')}{z(\mathbf{n})} - \Delta s^\circ, \tag{7.36}$$

Using (7.35) and (7.36), it follows immediately that:

$$\delta q = \Delta h^\circ. \tag{7.37}$$

When summing (7.37) over all transitions, we obtain the total heat $q(t)$ exchanged with the heat bath, at time t, in the form of a sum over all past events index by j:

$$q(t) = \sum_j \delta q_j, \tag{7.38}$$

According to the second law of Stochastic Thermodynamics [41, 40], the total entropy production on this trajectory is:

$$\Delta s_{tot} = \Delta s + \Delta s_m, \tag{7.39}$$

where Δs is the change of system entropy between the final and initial states and Δs_m the change in medium entropy. The latter is fundamentally associated to the heat defined above by $\Delta s_m = -\beta q$.

Given Eq. (7.33), the difference of system entropy is:

$$\Delta s = \ln \frac{P(\mathbf{n}^I)}{P(\mathbf{n}^F)} + \ln \frac{z(\mathbf{n}^F)}{z(\mathbf{n}^I)} + s^\circ(\mathbf{n}^F) - s^\circ(\mathbf{n}^I), \tag{7.40}$$

which when combined with Eqs. (7.34)-(7.36), leads to the expected central result that the total entropy production is the ratio of the probability of forward paths to that of backward paths:

$$\Delta s_{tot} = \ln \frac{P(\mathbf{n}^I)\, W_{\mathbf{n^I}\to\mathbf{n^1}} \ldots W_{\mathbf{n^{F-1}}\to\mathbf{n^F}}}{P(\mathbf{n}^F)\, W_{\mathbf{n^1}\to\mathbf{n^I}} \ldots W_{\mathbf{n^F}\to\mathbf{n^{F-1}}}}. \tag{7.41}$$

The contribution due to degeneracy can be further split as

$$\frac{1}{N} \ln \frac{z(\mathbf{n^F})}{z(\mathbf{n^I})} = \frac{1}{N} \ln \frac{\prod_\Omega n_\Omega^I!}{\prod_\Omega n_\Omega^F!} = \Delta s_L + \Delta s_\omega, \tag{7.42}$$

with Δs_L the length entropy per chain and Δs_ω the weighted sequence entropy per chain of a finite system:

$$\Delta s_L = \frac{1}{N} \ln \frac{\prod_l n_l^I!}{\prod_l n_l^F!} \tag{7.43}$$

$$\Delta s_\omega = \frac{1}{N} \ln \frac{\prod_\Omega n_\Omega^I!}{\prod_\Omega n_\Omega^F!} - \frac{1}{N} \ln \frac{\prod_l n_l^I!}{\prod_l n_l^F!}, \tag{7.44}$$

7.2.8 Connection to the macroscopic approach

Let us now check how the stochastic and deterministic approaches are connected. We assume that there is no distribution of the initial condition, $P(\mathbf{n}^I) = 1$. This simplifies the first term in the change of stochastic system entropy in Eq. (7.40), since $\ln P(\mathbf{n}^I) = 0$. In order to evaluate $P(\mathbf{n}^F)$, let us assume that the system has reached equilibrium at the final time. For a macroscopic system, that probability distribution takes the equilibrium form *:

$$P(\mathbf{n}^F) = z(\mathbf{n}^F) \prod_\Omega (y_\Omega)^{n_\Omega^F}, \tag{7.45}$$

where we have used the definition of the degeneracy factor in Eq. (7.30) and the conservation law of the number of chains $\sum_\Omega n_\Omega = N$. To make the connection with the macroscopic description, we can show that the polymer fractions y_Ω previously defined in Eq. (7.6), must also be the ensemble average of n_Ω divided by N:

$$y_\Omega = \frac{\langle n_\Omega^F \rangle}{N}, \tag{7.46}$$

where the average is taken with respect to the equilibrium distribution of Eq. (7.45). Now, by reporting Eq. (7.45) into Eq. (7.33), one finds

$$s(\mathbf{n}^F) = -\sum_\Omega n_\Omega^F \ln y_\Omega + s^\circ(\mathbf{n}^F). \tag{7.47}$$

When this expression is averaged over the equilibrium distribution of Eq. (7.45), one recovers using Eqs. (7.31) and (7.34) the familiar expression of the entropy introduced in the equilibrium thermodynamics section, namely Eq. (7.9).

*Of course, the distribution only pertains to states $\mathbf{n}^F$ that are accessible within the scope of the conservation laws on sequences and chains

Let us discuss the connection to the macroscopic approach for the separate contributions of length and sequence. We start by using Stirling's approximation in Eq. (7.44), $\ln n! = n\ln n - n + O(\ln n)$. In this limit, one recovers the expected contributions to the entropy:

$$\begin{aligned}\Delta s_L &\approx \sum_l \left[\frac{n_l^I}{N}\ln\frac{n_l^I}{N} - \frac{n_l^F}{N}\ln\frac{n_l^F}{N}\right],\\ \Delta s_\omega &\approx \sum_{l,\Omega}\left[\frac{n_\Omega^I}{N}\ln\frac{n_\Omega^I}{N} - \frac{n_\Omega^F}{N}\ln\frac{n_\Omega^F}{N}\right] - \Delta s_L.\end{aligned} \tag{7.48}$$

In the thermodynamic limit, the probability distribution of n_Ω becomes peaked around the value $\langle n_\Omega\rangle = N_\Omega$. By replacing n_Ω by N_Ω and n_l by N_l and using the definitions: $N_\Omega = NY_l U_{l,\Omega}$ and $N_l = NY_l$, in Eq. (7.48), one recovers precisely the first two terms in (7.28). In this limit, the n_Ω becomes deterministic, therefore, the first term in Eq. (7.40) becomes negligible.

Finally, we note that the heat per polymer is:

$$\frac{q}{N} = \sum_{l,\Omega}\left[Y_l^F U_{l,\Omega}^F - Y_l^I U_{l,\Omega}^I\right]h_\Omega^\circ. \tag{7.49}$$

while the internal entropy part is similarly

$$\mathscr{S}^0 = \sum_{l,\Omega}\left[Y_l^F U_{l,\Omega}^F - Y_l^I U_{l,\Omega}^I\right]s_\Omega^\circ. \tag{7.50}$$

By combining Eqs. (7.48),(7.49) and (7.50), we see that we recover all the terms in the entropy production of Eq. (7.28) obtained in the macroscopic approach.

7.2.9 Equilibrium Length Distributions

Let us now study the length distribution at equilibrium for some simple systems. We define N_l the number of polymers of length l, such that $Y_l = N_l/N$. We will first consider an example where reactions are neutral, and then an example with nearest-neighbor interactions.

Neutral reactions

We consider all reactions to be neutral: $k_{\omega_A\omega_B,\omega_C\omega_D} = k_{\omega_A\omega_D,\omega_C\omega_B}$. Consequently, there is no energy landscape, and relaxation will be purely entropic.

We have two separate conservation law for the number of chains: $\sum_{l=l_{min}}^\infty N_l = N$ and for the number of monomers (mass conservation): $\sum_{l=l_{min}}^\infty lN_l = M$, with l_{min} the length of the shortest possible species. Now, detailed balance imposes $N_{l_A}N_{l_B} = N_{l_C}N_{l_D}$ with $l_A + l_B = l_C + l_D$, which leads to an exponential length distribution: $N_l = A(B)^{l-l_{min}}$, where A and B are constants depending on the mechanism. Solving the algebraic equations for N_l for Chain-Exchange where $l_{min} = 2$ yields:

$$N = \frac{A}{1-B}, \quad M = -\frac{AB^2(B-2)}{B^2(1-B)^2} = \frac{A(B-2)}{(1-B)^2}, \tag{7.51}$$

from which we find:

$$A = \left(\frac{N}{\frac{M}{N}-1}\right), \quad B = \left(\frac{\frac{M}{N}-2}{\frac{M}{N}-1}\right). \tag{7.52}$$

We thus have an expression for Y_l^{eq}:

$$Y_l^{eq} = \frac{1}{\frac{M}{N}-1}\left(\frac{\frac{M}{N}-2}{\frac{M}{N}-1}\right)^{l-2}. \tag{7.53}$$

For Attack-Exchange, $l_{min} = 1$ and a similar calculation leads to:

$$Y_l^{eq} = \frac{N}{M}\left(1 - \frac{N}{M}\right)^{l-1}. \tag{7.54}$$

Such exponential length distributions were already obtained long ago by Flory [45], [46] in their pioneering work on reversible polymerization. These equilibrium distributions also hold when the polymers contain different types of monomers (i.e. when $m \neq 1$).

Entropically, the number of allowed permutation operations on mobile monomers is not affected by the relaxation of the length distribution, which means the length distribution can freely tend to maximum disorder allowed by its conservation laws. This may no longer be the case when there is an energy function attached to the polymer sequences.

Nonneutral reactions with nearest neighbor interactions

Let us now consider a simple example of an energy landscape. We assign standard free energies to bonds as a function of the monomers that are held together by them. We denote with $\tilde{n}_\omega$ the total number of bonds ω among polymers in the system, where ω is a (sub)sequence of length 2 (e.g. 01). This number is:

$$\tilde{n}_\omega = \sum_{\omega_A,\omega_B} n_{\omega_A \omega \omega_B}. \tag{7.55}$$

When only two monomer types are present, the only relevant exchange reaction at the level of bonds is:

$$\omega_A 00 \omega_B + \omega_C 11 \omega_D \rightleftarrows \omega_C 10 \omega_B + \omega_A 01 \omega_D, \tag{7.56}$$

since the other reactions do not change bond composition.

Let us introduce the standard chemical potential of the various bonds: $\tilde{\mu}^\circ_{00}$, $\tilde{\mu}^\circ_{01}$, $\tilde{\mu}^\circ_{10}$ and $\tilde{\mu}^\circ_{11}$. Then the forward rate of reaction (7.56) is $k^+ \sim \exp(-\beta(\tilde{\mu}^\circ_{00} + \tilde{\mu}^\circ_{11}))$ while the backward rate is $k^- \sim \exp(-\beta(\tilde{\mu}^\circ_{01} + \tilde{\mu}^\circ_{10}))$. The detailed balance condition imposes

$$\frac{\tilde{n}^{eq}_{00}\,\tilde{n}^{eq}_{11}}{\tilde{n}^{eq}_{01}\,\tilde{n}^{eq}_{10}} = \frac{k^-}{k^+} = \exp(-\beta \Delta\tilde{\mu}^\circ), \tag{7.57}$$

in terms of the standard chemical potential change $\Delta\tilde{\mu}^\circ = \tilde{\mu}^\circ_{01} + \tilde{\mu}^\circ_{10} - \tilde{\mu}^\circ_{00} - \tilde{\mu}^\circ_{11}$.

Let us consider a symmetric initial condition, in the relative amount of subsequences 00 and 11, including terminal and initial positions. Since the only relevant reaction is given by Eq. 7.56, this symmetry will persist and we will have $\tilde{n}_{00} = \tilde{n}_{11}$ and $\tilde{n}_{01} = \tilde{n}_{10}$ at all times. As a result, Eq. (7.57) simplifies into:

$$\frac{\tilde{n}^{eq}_{00}}{\tilde{n}^{eq}_{01}} = \exp\left(\frac{-\beta\Delta\tilde{\mu}^\circ}{2}\right). \tag{7.58}$$

The free energy of the system can be written in terms of: (i) entropy of the length distribution (ii) standard free energy of the subsequences (iii) entropy of the subsequence distribution. Since (i) is not coupled to (ii) and (iii), we can maximize (i) independently. Consequently, we obtain the same length distribution as in the energetically neutral case: (7.53).

An explicit expression of the equilibrium sequence distribution for given length: $U^{eq}_{l,\Omega}$ can be found from the following argument. A given sequence Ω has an energy e_Ω corresponding to its bond composition. We define n_B as the number of bonds of the type 00 and 11 in Ω. Therefore:

$e_\Omega = n_B \Delta\tilde{\mu}^\circ/2$. There are $2\binom{l-1}{n_B}$ different sequences of length l with n_B of such bonds. We thus find for $U^{eq}_{l,\Omega}$:

$$\begin{aligned} U^{eq}_{l,\Omega} &= \frac{\exp(-\beta e_\Omega)}{\sum_{n_B=0}^{l-1} 2\binom{l-1}{n_B} \exp\left(-\frac{\beta n_B \Delta\tilde{\mu}^\circ}{2}\right)}, \\ &= \frac{\exp(-\beta e_\Omega)}{2\left(1+\exp\left(-\frac{\beta\Delta\tilde{\mu}^\circ}{2}\right)\right)^{l-1}}. \end{aligned} \tag{7.59}$$

[For less symmetric cases or more complex energy landscapes, Y^{eq}_l can deviate from an exponential. For example, let us consider a case in which i) the first and last monomer are 1, mobile monomers are 0 and ii) 01 and 10 bonds are strongly disfavored $-\beta\Delta\tilde{\mu}^\circ \gg 1$. To minimize energy, we then expect a large number of 11 sequences and a small number of 100...0001 sequences.]

7.2.10 Relaxation kinetics

In this system, a number of timescales are at play, as shown in Table 7.2:
(i) the mean reaction time is $1/k$,
(ii) the waiting time τ_r is the time it takes to perform the next chemical reaction. It follows from considering the combined rate of all possible pathways. For attack-exchange, a polymer terminus reacts with a bond. Since there are N termini and $M-N$ bonds, $\tau_r = 1/k(M-N)N$.
(iii) The relaxation time of the length, τ_l, follows from the detailed kinetic rate equations for number of polymers of length l, N_l. It can be shown that the N_l can be written as a sum of exponentially decaying perturbations, for which the slowest decays on a timescale τ_l.
(iv) The characteristic time for sequence relaxation, τ_ω, is defined as the longest relaxation time for subsequences of length 2 or larger. In the next section, calculations are provided for the expression of τ_l and τ_ω given in Table 7.2.

Reaction	τ_r	τ_l	τ_ω
Attack-Exchange	$\frac{1}{kN(M-N)}$	$\frac{1}{kM}$	$\frac{1}{kN}$
Chain-Exchange	$\frac{2}{k(M-N)^2}$	$\frac{1}{k(M-N)}$	$\frac{1}{k(M-N)}$

Table 7.1: Expressions of the various relaxation times: τ_r waiting time for a reaction to occur, τ_l relaxation time of the length, τ_ω relaxation time of the sequence.

Length Relaxation

We now derive the characteristic time of length relaxation, first for Chain-Exchange, then for Attack-Exchange. The amount of species of length l: N_l, evolves according to:

$$\begin{aligned} \frac{dN_l}{dt} &= k \sum_{l_A+l_B=l} \sum_{l_C,l_D}^{\infty} [N_{l_A+l_D} N_{l_C+l_B} - N_{l_A+l_B} N_{l_C+l_D}] \\ &= k \sum_{l_C,l_D}^{\infty} \sum_{l_B=1}^{l-1} N_{l_D+l-l_B} N_{l_C+l_B} - k(l-1) N_l \sum_{l_x=2}^{\infty} (l_x-1) N_{l_x} \\ &= k \sum_{l_B=1}^{l-1} \left(N - \sum_{l_x=2}^{l-l_B} N_{l_x}\right)\left(N - \sum_{l_y=2}^{l_B} N_{l_y}\right) - k(l-1)(M-N)N_l. \end{aligned} \tag{7.60}$$

Therefore, the homogeneous equation takes the form:

$$\frac{dN_l}{dt} = k(l-1)N^2 - k(l-1)(M-N)N_l, \tag{7.61}$$

which admits the solution:

$$N_l = \frac{N^2}{M-N} + \left(N_l^I - \frac{N^2}{M-N}\right)\exp(-k(l-1)(M-N)t), \tag{7.62}$$

Note that the exponential in the homogeneous solution is proportional to $l-1$. The highest possible order of additional terms introduced in the particular solution is $l-2$. These arise from terms such as

$$N_{l-l_B}N_{l_B} \propto \exp(-k(l-l_B-1)(M-N)t)\exp(-k(l_B-1)(M-N)t) = \exp(-k(l-2)(M-N)t) \tag{7.63}$$

We therefore have no resonant terms for any N_l, and Eq. 7.60 admits solutions of the form:

$$N_l = A_{0,l} + \sum_{n=2}^{l} A_{n,l}\exp(-k(n-1)(M-N)t), \tag{7.64}$$

where $A_{0,l}$ and $A_{n,l}$ are constants depending on initial concentrations of all species.

This expression confirms that the slowest relaxation time of the length for Chain-Exchange reaction is $\tau_l = 1/(k(M-N))$ as given in Table 7.2.

For Attack-Exchange, the kinetic equation for N_l is:

$$\frac{dN_l}{dt} = k\sum_{l_A,l_B=1}^{\infty}[N_{l_A}N_{l+l_B} - N_{l_A+l_B}N_l] + k\sum_{l_A}^{l-1}\sum_{l_B=1}^{\infty}[N_{l_A}N_{l_B+l-l_A} - N_lN_{l_B}] \tag{7.65}$$

$$= k[N(N - \sum_{l_B=1}^{l-1} N_{l_B}) + \sum_{l_B}^{l-1} N_{l-l_B}(N - \sum_{l_x=1}^{l_B} N_{l_x}) - (M + N(l-1))N_l] \tag{7.66}$$

Upon solving the homogeneous equations, the general solution for every N_l can be written as :

$$N_l = A_{0,l} + \sum_{n=1}^{l} A_{n,l}\exp(-k(M+N(l-1))t) \tag{7.67}$$

For which the relaxation time is: $\tau_l = \frac{1}{kM}$.

Length Relaxation in a CSTR

For a CSTR, the contribution of influx and outflux is directly taken into account via the term $(N_l^\circ - N_l)/\tau$, where N_l° is the number of polymers in the inflowing solution. For Chain-Exchange, this gives us the equation:

$$\frac{dN_l}{dt} = k\sum_{l_B=1}^{l-1}\left(N - \sum_{l_x=2}^{l-l_B} N_{l_x}\right)\left(N - \sum_{l_y=2}^{l_B} N_{l_y}\right) - k(l-1)(M-N)N_l + \frac{1}{\tau}(N_l^\circ - N_l). \tag{7.68}$$

As before, we can first solve the homogeneous equation, and the equation can be solved by solving a hierarchy of equations. The homogeneous equation is

$$\frac{dN_l}{dt} = k(l-1)N^2 + \frac{1}{\tau}N_l^\circ - \left(k(l-1)(M-N) + \frac{1}{\tau}\right)N_l, \tag{7.69}$$

for which we can write the solution

$$N_l = \frac{k(l-1)N^2 + \frac{1}{\tau}N_l^\circ}{k(l-1)(M-N)+\frac{1}{\tau}} + \left(N_l^I - \frac{k(l-1)N^2 + \frac{1}{\tau}N_l^\circ}{k(l-1)(M-N)+\frac{1}{\tau}} \right) \exp\left(-\left(k(l-1)(M-N) + \frac{1}{\tau} \right) t \right). \tag{7.70}$$

Where N_l^I is the initial number of polymers of length l. The particular solution therefore has the form

$$N_l = A_{0,l} + \sum_{n=2}^{l} A_{n,l} \exp\left(-\left(k(l-1)(M-N) + \frac{1}{\tau} \right) t \right), \tag{7.71}$$

For l=2, the homogeneous solution is the full solution. Compared to the closed system, the relaxation timescale τ_l for is reduced to $\tau_l = 1/(k(M-N)+1/\tau)$.

For Attack-Exchange, the modified equation becomes

$$\frac{dN_l}{dt} = k\left[N\left(N - \sum_{l_B=1}^{l-1} N_{l_B} \right) + \sum_{l_B}^{l-1} N_{l-l_B}\left(N - \sum_{l_x=1}^{l_B} N_{l_x} \right) \right] + \frac{1}{\tau}N_l^\circ - \left(k(M+N(l-1)) + \frac{1}{\tau} \right) N_l. \tag{7.72}$$

Which admits a particular solution of the form

$$N_l = A_{0,l} + \sum_{n=1}^{l} A_{n,l} \exp\left(-\left(k(M+N(l-1)) + \frac{1}{\tau} \right) t \right), \tag{7.73}$$

which modifies the timescale of length relaxation to $\tau_l = 1/(kM+1/\tau)$.

Sequence Relaxation

We will now consider the relaxation of the bond composition in the chain-exchange reaction. Let us assume that the initial condition is symmetric with respect to the content of 0 and 1 monomers in the pool. As a result, this symmetry will remain at all times, and we can introduce x and y variables such that $\tilde{n}_{00} = \tilde{n}_{11} = x$ and $\tilde{n}_{01} = \tilde{n}_{10} = y$. In the mean-field approximation, the evolution of equations of these variables are

$$\frac{dx}{dt} = k^- y^2 - k^+ x^2, \tag{7.74}$$

$$\frac{dy}{dt} = k^+ x^2 - k^- y^2, \tag{7.75}$$

where k^+ is a forward rate and k^- a backward rate. By summing the two equations above, one recovers the conservation law that the sum of x and y is constant. The constant is fixed by the initial number of bonds: $2x+2y = M-N$. Therefore, we end up with the equation:

$$\frac{dx}{dt} = k^-\left(\frac{M-N}{2} - x \right)^2 - k^+ x^2 \tag{7.76}$$

For neutral reactions, $k^+ = k^- = k$, the equation simplifies into:

$$\frac{dx}{dt} = -k\left[(M-N)x - \left(\frac{M-N}{2} \right)^2 \right]. \tag{7.77}$$

This linear ODE has a simple exponential as solution with the characteristic relaxation time $\tau_\omega = 1/k(M-N)$, which was given in Table 7.2.

Let us now extend the above results to the case that transitions are affected by an energy landscape. We start with the detailed balance condition: $k^+ = k^- \exp(-\beta\Delta\tilde{\mu}^\circ)$. We now go back to Eq (7.76) when $k^- \neq k^+$. We obtain a nonlinear ODE of the form

$$\frac{dx}{dt} = ax^2 + bx + c \tag{7.78}$$

With a, b and c constants, given by:

$$a = k^- - k^+, \quad b = k^-(M-N), \quad c = k^- \left(\frac{M-N}{2}\right)^2. \tag{7.79}$$

We note that $\sqrt{b^2-4ac} = \sqrt{k^+k^-}(M-N) > 0$. Therefore, we can make use of the integral:

$$\int_0^t dt = \int_{x(0)}^{x(t)} \frac{dx}{ax^2+bx+c} = \frac{-2}{\sqrt{b^2-4ac}} \tanh^{-1}\left(\frac{2ax(t)+b}{\sqrt{b^2-4ac}}\right) + C. \tag{7.80}$$

Therefore, the solution is of the form:

$$x(t) \propto \tanh\left[\frac{-\sqrt{b^2-4ac}}{2}(t-C)\right] + D \tag{7.81}$$

where C and D are constants. As $\tanh(t) = (1-\exp(-2t))/(1+\exp(-2t))$, we can identify $1/\sqrt{b^2-4ac}$ as a characteristic sequence relaxation time τ_ω equal to:

$$\tau_\omega = \frac{\exp\left(-\frac{\beta\Delta\tilde{\mu}^\circ}{2}\right)}{k^+(M-N)}, \tag{7.82}$$

For Attack-Exchange, let us consider a system with 2 monomers. let us denote by [0],[1] the concentration of terminal monomers and let us consider the bonds $[00], [01], [10]$ and $[11]$.

$$\frac{d[0]}{dt} = -k[0]([10]+[11]) + k[1]([01]+[00]), \tag{7.83}$$
$$\frac{d[1]}{dt} = -k[1]([01]+[00]) + k[0]([10]+[11]), \tag{7.84}$$
$$\frac{d[00]}{dt} = -k[1][00] + k[0][10], \tag{7.85}$$
$$\frac{d[10]}{dt} = -k[0][10] + k[1][00], \tag{7.86}$$
$$\frac{d[01]}{dt} = -k[0][11] + k[1][01], \tag{7.87}$$
$$\frac{d[11]}{dt} = -k[1][01] + k[0][11]. \tag{7.88}$$

There are a number of useful conserved quantities

$$N = [0]+[1], \tag{7.89}$$
$$n_1 = [11]+[10], \tag{7.90}$$
$$n_0 = [00]+[01], \tag{7.91}$$
$$M-N = [00]+[01]+[10]+[11]. \tag{7.92}$$

Which allows to write for the terminal monomer

$$\frac{d[0]}{dt} = kNn_0 - k[0](M-N), \tag{7.93}$$
$$\frac{d[1]}{dt} = kNn_1 - k[1](M-N). \tag{7.94}$$

which have exact solutions

$$[0] = \left([0]_0 - \frac{n_0 N}{M-N}\right)\exp\left(-k(M-N)t\right) + \frac{n_0 N}{M-N} \tag{7.95}$$

$$[1] = \left([1]_0 - \frac{n_1 N}{M-N}\right)\exp\left(-k(M-N)t\right) + \frac{n_1 N}{M-N} \tag{7.96}$$

Where $[0]_0$ and $[1]_0$ denote the monomer concentration at $t = 0$.

$$\frac{d[00]-[10]}{dt} = -kN([00]-[10]) + k([1]-[0])n_0. \tag{7.97}$$

7.3 Polymer recombination in open systems

In this section, copolymer systems are coupled to reservoirs with which they exchange copolymers with particular monomer sequences. The exact composition of reservoirs has profound consequences on conservation laws, chemical currents and growth. To study them systematically, we will use the stoichiometry matrix formalism introduced in Ch.2 (see also Refs[47, 48]), combined with the sequence notation introduced in Sec. 7.2.3.

7.3.1 Stoichiometric Matrices for Polymer Chemistries

Addition-Fragmentation

The Addition-Fragmentation reaction is a coupling reaction of the type $[\boldsymbol{i}] + [\boldsymbol{j}] \rightleftarrows [\boldsymbol{i}+\boldsymbol{j}]$. It is the most popular and straightforward reaction in models used to study origins-of-life scenarios [49, 19, 20, 50, 51]. For RNA, this reaction is strongly unfavorable. As will be shown in Sec. 7.5, many theoretical models in the literature operate in regimes where reversible polymerization is not appropriate for RNA, their dynamics implicitly requires a chemical regeneration step that happens on a fast timescale to fuel an activated pathway. For other polymers (e.g. polyethene), polymerization can be reversible, or even irreversible in the forward direction.

For the present discussion, we will consider the addition-fragmentation scheme without activation or modification for copolymer species. Given the reaction ($\omega_A\omega_B \rightleftarrows \omega_A + \omega_B$), we can write the following stoichiometric matrix:

$$\boldsymbol{\nu}^{\Omega}_{\omega_A,\omega_B} = -\delta^{\Omega}_{\omega_A} - \delta^{\Omega}_{\omega_B} + \delta^{\Omega}_{\omega_A\omega_B} \tag{7.98}$$

For the reaction between ω_A, ω_B, we then write the molecular current (or rate) :

$$J^{\omega_A\omega_B} = k^{+}_{\omega_A,\omega_B} n^{\omega_A} n^{\omega_B} - k^{-}_{\omega_A\omega_B} n^{\omega_A\omega_B} \tag{7.99}$$

The evolution equation for a species with sequence Ω is then:

$$\dot{n}^{\Omega} = \sum_{\omega_A,\omega_B} \nu^{\Omega}_{\omega_A\omega_B} J^{\omega_A\omega_B} \tag{7.100}$$

Note that, alternatively, we could have used an Einstein summation convention for sequence space, by the repeated $\omega_A\omega_B$ indices. Such a notation considerably shortens the upcoming equations. To preserve a general coherence with other chapters, this convention is not adopted here, but equations will be constructed such that this structure can be easily extracted.

The conservation laws that appear in this reaction can be found using the cokernel of the stoichiometry matrix, which yields the following relation for ℓ:

$$\ell^{(k)}_{\omega_A\omega_B} - \ell^{(k)}_{\omega_A} = \ell^{(k)}_{\omega_B}. \tag{7.101}$$

Here, (k) is an index for the conservation law. Intuitively, we expect a conservation law for the total amount of each monomer type III in the alphabet $[m] = \{1, 2, ..j, ..m\}$, such that $\text{III} \in [m]$. To see this, we introduce the counting function

$$\chi^{\text{X}}_{\Omega} \equiv \sum_{k=1}^{|\Omega|} \alpha \delta^{\text{III}}_{\Omega_k}, \tag{7.102}$$

which compares all $|\Omega|$ monomers with III and counts the number of matches. We thus obtain $\ell^{(\text{III})}_{\omega} = \alpha \chi^{\text{III}}_{\omega}$, with α an arbitrary constant. We therefore have m independent conservation laws, corresponding to the conservation of the total number of monomers for all the monomer types.

Attack-Exchange

This mode of exchange was introduced in Sec 7.2.1. It appears in splicing pathways of RNA and in some prebiotic scenarios [35].

The reaction conserves the amount of chemical bonds, as well as their chemical nature. The stoichiometry matrix for this reaction ($\omega_A \omega_B + \omega_C \rightleftarrows \omega_C \omega_B + \omega_A$) is

$$\boldsymbol{\nu}^{\Omega}_{\omega_A \omega_B, \omega_C} = -\delta^{\Omega}_{\omega_A \omega_B} - \delta^{\Omega}_{\omega_C} + \delta^{\Omega}_{\omega_C \omega_B} + \delta^{\Omega}_{\omega_A} \tag{7.103}$$

and the current

$$J^{\omega_A \omega_B, \omega_C} = k^{+}_{\omega_A \omega_B, \omega_C} n^{\omega_A \omega_B} n^{\omega_C} - k^{-}_{\omega_B \omega_C, \omega_A} n^{\omega_B \omega_C} n^{\omega_A} \tag{7.104}$$

As shown in Sec. 7.2.5 , the symmetry group of the reaction is of order 2. $\boldsymbol{\nu}^{\Omega}_{\omega_A \omega_B, \omega_C}$ admits exchange: $\omega_A \leftrightarrow \omega_C$ and the identity operation: I, to yield the equivalent stoichiometry matrix. This means we need a prefactor $\frac{1}{2}$ to avoid overcounting in the kinetic equations

$$\dot{n}^{\Omega} = \frac{1}{2} \sum_{\omega_A, \omega_B, \omega_C} \boldsymbol{\nu}^{\Omega}_{\omega_A \omega_B, \omega_C} J^{\omega_A \omega_B, \omega_C} \tag{7.105}$$

To find the conservation laws, we examine the cokernel

$$\boldsymbol{\ell} \cdot \boldsymbol{\nu} = \mathbf{0} \tag{7.106}$$

Which on the level of a single reaction yields the scalar equation

$$\sum_{\Omega} \ell_{\Omega} \cdot \boldsymbol{\nu}^{\Omega}_{\omega_A \omega_B, \omega_C} = 0. \tag{7.107}$$

From this relation follows that all conservation laws must solve

$$\ell^{(k)}_{\omega_A \omega_B} - \ell^{(k)}_{\omega_A} = \ell^{(k)}_{\omega_C \omega_B} - \ell^{(k)}_{\omega_C}, \tag{7.108}$$

where k is an index designating the kth conservation law. Eq. (7.108) admits a first set of m solutions of the form

$$\ell^{(\text{III},1)}_{\Omega} = \alpha \delta^{\text{III}}_{\Omega_1} \tag{7.109}$$

with α an arbitrary constant. Here, the delta function compares the first monomer of the sequence, Ω_1, with the chosen monomer $\text{III} \in [m]$. This reflects the fact that the first monomer is conserved in an Attack-Exchange reaction: there is no bond to break or exchange in front of the first monomer. The superscript 1 after III indexes the solution in the set of solutions.

There is a second set of m solutions for Eq. (7.108) of the form

$$\chi^{\mathrm{X}}_{\Omega} = \sum_{k=1}^{|\Omega|} \alpha \delta^{\mathrm{III}}_{\Omega_k} \tag{7.110}$$

which we first encountered for addition-fragmentation. We write it as $\ell^{\mathrm{III},2}_{\Omega} = \alpha\chi^{\mathrm{III}}_{\Omega}$. This corresponds to mass conservation for each monomer type.

When all conservation laws for the first monomer in the sequence ($\ell^{k,1}$) are taken together, we obtain the conservation of the total number of chains

$$\sum_{\mathrm{III}=1}^{m} \ell^{\mathrm{III},1}_{\Omega} n^{\Omega} = N, \tag{7.111}$$

where N denotes the total number of chains. The conservation of chains is thus a linearly dependent conservation law that follows from m other ones we already have. Similarly, taking all $\ell^{\mathrm{III},2}$ together, we sum the masses of all polymers

$$\sum_{\mathrm{III}=1}^{m} \ell^{(\mathrm{III},2)}_{\Omega} n^{\Omega} = M \tag{7.112}$$

Therefore, the conservation of total mass [†] M is also a linearly dependent conservation law that follows from m independent ones.

In total, there are $2m$ independent conservation laws. For $m = 1$, this reproduces a known result for a monomer-exchange version of Attack-Exchange [52]: $\ell^{(1)} = \alpha$ and $\ell^{(2)} = \alpha|\Omega|$).

Chain-Exchange

The Chain-Exchange reaction involves two chains, which exchange part of their chains (See also Sec. 7.2.1), and is also known as metathesis. The number of bonds are conserved as well as the chains. The energetics of the Chain-Exchange are therefore similar to Attack-Exchange. The dynamics is quite different, however.

The stoichiometry matrix for this reaction ($\omega_A\omega_B + \omega_C\omega_D \rightleftarrows \omega_A\omega_D + \omega_C\omega_B$) is

$$\boldsymbol{\nu}^{\Omega}_{\omega_A\omega_B,\omega_C\omega_D} = -\delta^{\Omega}_{\omega_A\omega_B} - \delta^{\Omega}_{\omega_C\omega_D} + \delta^{\Omega}_{\omega_A\omega_D} + \delta^{\Omega}_{\omega_C\omega_B} \tag{7.113}$$

and the current corresponds to

$$J^{\omega_A\omega_B,\omega_C\omega_D} = k_{\omega_A\omega_B,\omega_C\omega_D} n^{\omega_A\omega_B} n^{\omega_C\omega_D} - k_{\omega_A\omega_D,\omega_C\omega_B} n^{\omega_A\omega_D} n^{\omega_C\omega_B} \tag{7.114}$$

From the above we can infer that the symmetry group of the reaction is of order 4. We therefore have a prefactor of $\frac{1}{4}$ in the kinetic rate equation:

$$\dot{n}^{\Omega} = \frac{1}{4} \sum_{\omega_A,\omega_B,\omega_C,\omega_D} \boldsymbol{\nu}^{\Omega}_{\omega_A\omega_B,\omega_C\omega_D} J^{\omega_A\omega_B,\omega_C\omega_D} \tag{7.115}$$

Using the cokernel, as in (7.106), we find the equation

$$\ell^{(k)}_{\omega_A\omega_B} - \ell^{(k)}_{\omega_A\omega_D} = \ell^{(k)}_{\omega_C\omega_B} - \ell^{(k)}_{\omega_C\omega_D}. \tag{7.116}$$

This equation shares two solutions with the Attack-Exchange reaction

$$\ell^{(\mathrm{III},1)}_{\Omega} = \alpha\delta^{\mathrm{III}}_{\Omega_1}, \tag{7.117}$$

$$\ell^{(\mathrm{III},2)}_{\Omega} = \alpha\chi^{\mathrm{III}}_{\Omega}. \tag{7.118}$$

[†] More exactly, the conservation monomers

There is, however, an additional set of conservation laws that is unique to Chain-Exchange

$$\ell_{\Omega}^{(\mathrm{III},3)} = \alpha\delta_{\Omega_F}^{\mathrm{III}}. \tag{7.119}$$

where $\Omega_F = \Omega_{|\Omega|}$, the final monomer composing the chain Ω. In Chain-Exchange all terminal monomers remain in terminal positions, which is the origin of this conservation law.

The conservation of terminal monomers also generates the conservation of N :

$$\sum_{\Omega}\sum_{\mathrm{III}=1}^{m}\ell_{\Omega}^{(\mathrm{III},3)}n^{\Omega} = \sum_{\mathrm{III}=1}^{m}\boldsymbol{\ell}^{(\mathrm{III},3)}\cdot\boldsymbol{n} = N. \tag{7.120}$$

This means that we do not have $3m$, but $3m-1$ independent conservation laws, as we can pick any $\ell^{(\mathrm{III},3)}$ (or $\ell^{(\mathrm{III},1)}$ for that matter) and express it in terms of the rest. E.g. $\ell^{(1,3)}$ can then be written as $\ell^{(1,3)} = \sum_{\mathrm{III}'=1}^{m}\ell^{(\mathrm{III}',1)} - \sum_{\mathrm{III}=2}^{m}\ell^{(\mathrm{III},3)}$.

7.3.2 Open Chemical systems

In this section, the chemical system will be allowed to exchange matter with chemostats. As discussed in Sec. 2.5 and 3.1, a chemostat is a reservoir that fixes the chemical potential of a species, which can (often) be equated to fixing the concentration. As discussed in Sec. 2.5, the stoichiometry matrix can be decomposed in an external chemostatted part $\boldsymbol{\nu}^Y$ and an internal part $\boldsymbol{\nu}^X$, such that $\boldsymbol{\nu} = (\boldsymbol{\nu}^Y, \boldsymbol{\nu}^X)^T$. We have s_X internal species and s_Y chemostats. The rank nullity theorem $s-\ell = r-c$ can now applied for every submatrix. No reactions were added or removed, so the number of reactions r is unchanged. Decreasing the number of species[47] s by one should therefore decrease the number of conservation laws ℓ by one, or increase the number of cycles c by one. Equivalently, the number of chemostats s_Y corresponds to the number of affinities a and broken conservation laws b

$$s_Y = a+b. \tag{7.121}$$

If we add a new chemostat, and the chemical network allows a path (a sequence of reactions with no local buildup or exhaustion) towards other chemostats already present, then chemical potentials have been put in contact: a new affinity ensues. If a new chemostat does not have such a path, its addition allows to inject matter in novel ways, which breaks a conservation law. The number of conservation laws thus provides an upper bound for the amount of chemostats one can introduce before currents are generated.

A current arising from chemostatting is associated with an emergent cycle $\boldsymbol{c}^*$. Where $\boldsymbol{c}^*$ is a right nullvector for $\boldsymbol{\nu}^X$, but not for $\boldsymbol{\nu}^Y$

$$\boldsymbol{\nu}^X\cdot\boldsymbol{c}^* = \boldsymbol{0} \tag{7.122}$$

$$\boldsymbol{\nu}^Y\cdot\boldsymbol{c}^* \neq \boldsymbol{0} \tag{7.123}$$

$\boldsymbol{c}^*$ is a linear combination of chemical reactions that leads to no compositional change of the system and a net exchange of matter between chemostats.

It has been shown, for a dynamics with the same conservation laws as Attack-Exchange[52] and $m=1$, that for $s_Y \leq 1$, the system relaxes to equilibrium. For $s_Y = 2$, the system either i) relaxes to equilibrium or ii) grows in unbounded fashion. For $s_Y \geq 3$, the system either i) relaxes to a nonequilibrium steady-state (NESS) or ii) grows in unbounded fashion.

For $m > 1$, we do not need to break all conservation laws to obtain a NESS. However, having $s_Y = \ell+1$ ensures that a current arises. For Attack-Exchange and Chain-Exchange, the minimum number of chemostats necessary to get a NESS is 3 for $m=1$, but we will show that for higher m this becomes 2 chemostats. We have a region for s_Y from 2 to ℓ in which both a NESS and equilibrium are possible outcomes. Table 7.2 provides the bounds for the number of chemostats in which both types of behavior can be observed

Reaction	$s_{Y,min}$	l
Addition-Fragmentation	2	m
Attack-Exchange	2	2m
Chain-Exchange	2	3m-1

Table 7.2: bounds for chemostats to be either NESS or Equilibrium. If we have less than $s_{Y,min}$ chemostats we always have equilibrium, above $s_{Y,max}$ we always obtain affinities.

Broken conservation laws and emergent cycles

We denote chemostatted species, with elements $\{\Omega_Y\}$. We will use Ω_X to explicitly denote non-chemostatted species.

For Attack-Exchange, an emergent cycle ($c^{\omega_A\omega_B,\omega_C}$) is defined by:

$$\frac{1}{2}\sum_{\Omega_X} \nu^{\Omega_X}_{\omega_A\omega_B,\omega_C} c^{\omega_A\omega_B,\omega_C} = 0, \tag{7.124}$$

$$\frac{1}{2}\sum_{\Omega_Y} \nu^{\Omega_Y}_{\omega_A\omega_B,\omega_C} c^{\omega_A\omega_B,\omega_C} \neq 0. \tag{7.125}$$

Applying the conservation laws $\ell^{(\mathrm{III},k)}_{\Omega_Y}$ to the emergent cycle $\boldsymbol{c}^*$, we find that the conservation law still constrains overall currents

$$\frac{1}{2}\sum_{\omega_A,\omega_B,\omega_C,\Omega_Y} \ell^{(\mathrm{III},k)}_{\Omega_Y} \nu^{\Omega_Y}_{\omega_A\omega_B,\omega_C} c^{\omega_A\omega_B,\omega_C} = 0 \tag{7.126}$$

Equation (7.126) shows that the emergent cycles obey all conservation laws. Let us denote

$$\square^{\Omega_Y} = \sum_{\omega_A,\omega_B,\omega_C} \ell^{(\mathrm{III},k)}_{\Omega_Y} \nu^{\Omega_Y}_{\omega_A\omega_B,\omega_C} c^{\omega_A\omega_B,\omega_C}, \tag{7.127}$$

This provides a method of finding the emergent cycles, using $\ell^{(\mathrm{III},1)}_{\Omega} = \alpha\delta^{\mathrm{III}}_{\Omega_1}$ and $l^{(\mathrm{III},2)}_{\Omega} = \alpha\chi^{\mathrm{III}}_{\Omega}$, emergent cycles arise as non-trivial solutions to Eq. (7.126). As an example, let us chemostat the sequences 010 and 001, so that $\Omega_Y\{010,001\}$ and $m = 2$. This gives us:

$$\begin{aligned}
&\delta^0_{010}\square^{010} + \delta^0_{001}\square^{001} = \square^{010} + \square^{001} = 0\\
&\delta^1_{010}\square^{010} + \delta^1_{001}\square^{001} = 0 + 0 = 0\\
&\chi^0_{010}\square^{010} + \chi^0_{001}\square^{001} = 2\square^{010} + 2\square^{001} = 0\\
&\chi^1_{010}\square^{010} + \chi^1_{001}\square^{001} = \square^{010} + \square^{001} = 0
\end{aligned} \tag{7.128}$$

We see that a non-trivial solution appears: $\left\{\square^{010},\square^{001}\right\} = \{1,-1\}$. As a consequence, we have an independent affinity: $a = 1$, and one net broken conservation law: $b = 1$. We verify this by doing a computer simulation of this system using Gillespie's algorithm [53], choosing all rate constants equal (the reactions can be considered neutral) and starting with an empty system in contact with chemostats. The system reaches a steady-state, Figure 7.3 shows the constant injection and ejection of sequences.

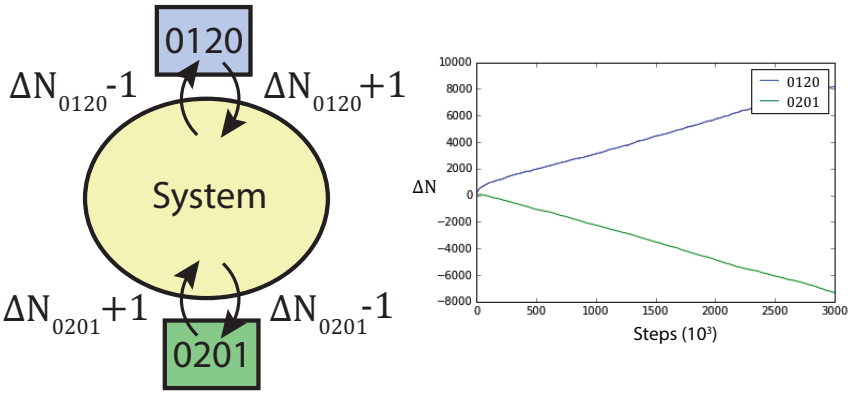


Figure 7.3: left: schematic representation of two sequence reservoirs performing exchange with the system. A bookkeeping of the net exchange is kept through the variables ΔN_{0120} and ΔN_{0201}. right: $\Delta N_{0120}, \Delta N_{0201}$ in the steady state regime, as function of time, expressed in simulation steps in Gillespie's algorithm. Reservoirs are maintained at $\bar{c}^{0120} = 5, \bar{c}^{0201} = 2$.

As we inject sequences containing 0's, 1's, and having a 0 at their first position, we lose the conservation laws $\ell^{(0,1)} = \alpha\delta^0_{\Omega_1}$, $\ell^{(1,2)} = \alpha\chi^1_\Omega$ and $\ell^{(0,2)} = \alpha\chi^0_\Omega$. However, we obtain modified (denoted with an asterisk) conservation laws, corresponding to the nullvectors

$$\ell^{(0,2*)}_\Omega = \alpha(\chi^0_\Omega - \chi^0_{010}), \tag{7.129}$$
$$\ell^{(1,2*)}_\Omega = \alpha(\chi^1_\Omega - \chi^1_{010}). \tag{7.130}$$

The corresponding conservation laws are:

$$L_{0*} = M_0 - \chi^0_{010}N = M_0 - 2N, \qquad L_{1*} = M_1 - \chi^1_{010}N = M_1 - N, \tag{7.131}$$

where L_{0*} and L_{1*} are constants, N the total number of chains, and M_{III} the total amount of monomer III in sequences. One might be tempted to also try this trick for $l^{(0,1)}$, which would yield:

$$\ell^{(0,1*)}_\Omega = \alpha(\delta^0_{\Omega_1} - \delta^0_{010}) = -\alpha\delta^1_{\Omega_1} \equiv l^{(1,1)}, \tag{7.132}$$

where we use that α is an arbitrary constant. We see that $\ell^{(0,1)}$ is no longer linearly independent. We are only left with three linearly independent conservation laws: $\ell^{(0,2*)}, \ell^{(1,2*)}$ and $\ell^{(1,1)}$.

An example with 6 chemostats

Let us place an attack-exchange system in contact with the following reservoir species: $\Omega_y = \{0, 010, 1, 121, 2, 202\}$, chemostatted in the amounts $\bar{c}_{\Omega_y} = [400, 20, 400, 40, 400, 20]$ (These values are chosen such that unbalanced growth does not occur, which will be detailed in the next section). As can be verified with the algebra from the last section, these 6 species cannot give an emergent cycle, despite being numerous and diverse in composition.

Figure 7.4 shows ΔN (the amount injected minus the amount ejected) for every species during the simulation. We see that initially, the environment injects longer sequences and the system ejects the shorter ones. This leads to more mass per chain. Clearly, the currents vanish. For emergent cycles between chemostats, we would expect linear asymptotes as observed in Fig 7.3.

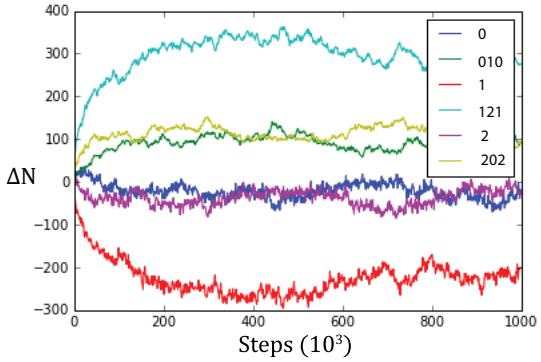


Figure 7.4: Number of species injected minus ejected during simulation for an Attack-Exchange reaction in contact with 6 chemostats. The system relaxes to an equilibrium.

An example with 7 chemostats

Let us now consider a system with 7 chemostatted species: $\Omega_y = \{00, 0, 010, 1, 121, 2, 202\}$ where we set the concentrations at $\bar{c}_{\Omega_y} = [1, 20, 2, 20, 1, 20, 2]$. As predicted, this number of chemostats must yield at least one affinity. Upon inspecting the simulation result in Fig. 7.5, a net current is indeed observed.

The species introduced allow for an emergent cycle which involves all 7 species: $\boldsymbol{c}^* = \{\nu_\gamma^{00}, \nu_\gamma^{0}, \nu_\gamma^{010}, \nu_\gamma^{1}, \nu_\gamma^{121}, \nu_\gamma^{2}, \nu_\gamma^{202}\} = \{-2, 1, 1, 1, -1, -1, 1\}$, which is the largest number of different species an emergent cycle can obtain for Attack-Exchange and $m = 3$. Written as an overall reaction, the emergent cycle corresponds to

$$2\ \ 00 + 121 + 2 \underset{-\boldsymbol{c}^*}{\overset{+\boldsymbol{c}^*}{\rightleftharpoons}} 0 + 010 + 1 + 202. \tag{7.133}$$

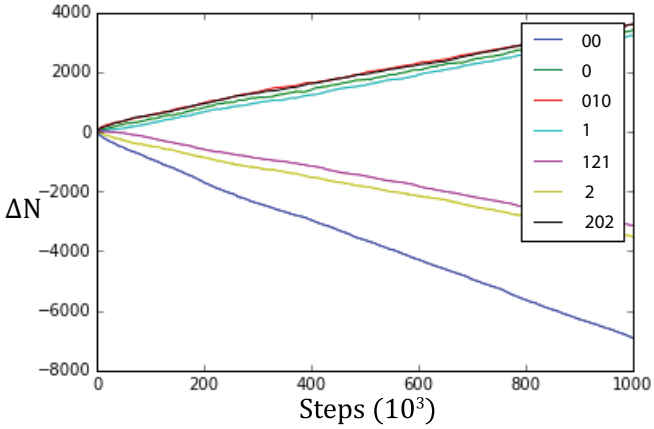


Figure 7.5: Number of species injected minus ejected during simulation, 7 chemostat Atack-exchange. The nonequilibrium steady state has an affinity corresponding to the net reaction7.133.

7.3.3 Net polymerization driven by reservoirs

So far, we have discussed systems reaching a steady-state, corresponding either to an equilibrium state or a NESS. In a recent investigation of monomer exchange reactions [54], a regime has been identified in which the system does not reach a steady state, but perpetually grows in time. In this section, we will first discuss the sequenceless version of this phenomenon for the most easy case (2 chemostats). We recall that then, $b = 2$ and $a = 0$, we do not have emergent cycles.

Sequenceless unbalanced growth

For a sequenceless exchange mechanism, we only consider length exchange

$$[i]+[j] \rightleftharpoons [i-k]+[j+k] \tag{7.134}$$

Assuming all rate constants equal, the following equation is satisfied when detailed balance holds:

$$N_i N_j = N_{i-k} N_{j+k} \tag{7.135}$$

Two of these abundances are fixed by the chemostats. We search for solution with the property $f(i)f(j) = f(i-k)f(j+k)$, which hints towards exponential functions, as $\exp(x)\exp(y) = \exp(x+y) = \exp(x-a)\exp(x+a)$. Denoting the length of a species with k, the smallest length chemostatted k_1 and the largest length k_2, the exponential solution under these constraints is found to be:

$$N_k = N_{k_1}\left(\frac{N_{k_2}}{N_{k_1}}\right)^{\frac{k-k_1}{k_2-k_1}}. \tag{7.136}$$

This distribution is controlled by the ratio N_{k_2}/N_{k_1}. As long as $N_{k_2}/N_{k_1} < 1$, we get a decreasing exponential. As derived in Sec. 7.2.9, this is a very typical equilibrium distribution.

As soon as $N_{k_{y2}}/N_{k_{y1}} \geq 1$, however, we obtain an increasing exponential. In the absence of some imposed cutoff length, such an equilibrium solution becomes unphysical: its realization requires an infinite accumulation of mass.

This regime is referred to as the 'unbalanced' regime[54]. (7.136) ceases to be a physical solution, and no finite-mass stationary solution can be found. The use of (7.136) is that it gives a criterion, $N_{k_{y2}}/N_{k_{y1}} \geq 1$, for when this happens. When the system accumulates sufficient mass in its volume, we can no longer treat it as an ideal dilute system, but would rather have to treat it using e.g. Flory-Huggins solution theory[55, 45].

2 Chemostats, no monomer bias

We will now add different types of monomers to the mix. In terms of the length growth, we still have aforementioned dynamics: $[i]+[j] \rightleftharpoons [i-k]+[j+k]$. The main difference comes from the chemostats, which chemostat a sequence concentration. The concentration of all species of the corresponding length, however, can become quite a bit larger. The conclusions and approach of this section are general for all exchange mechanisms, here we will treat the Attack-Exchange mechanism: $\omega_A\omega_B + \omega_C \rightleftharpoons \omega_C\omega_B + \omega_A$. For detailed balance, we then have:

$$N^{\omega_A\omega_B} N^{\omega_C} = N^{\omega_A} N^{\omega_C\omega_B} \tag{7.137}$$

To illustrate the effect of sequence chemostats, we will study a simple example, where $\{\Omega_Y\} = \{0, 010\}$. We recall that for Chain-Exchange reactions, the first monomer is conserved, the monomers that come after it can be scrambled. In this example, the mobile sequence is 10. Every time we inject 010, we inject one mobile 1 and one mobile 0. Now imagine the following sequence of Attack-Exchange reactions

$$010+010 \rightleftharpoons 01010+0 \rightleftharpoons 0101+00 \rightleftharpoons 010+001. \tag{7.138}$$

If the system reaches detailed balance, this means

$$\bar{N}^{010}\bar{N}^{010} = N^{001}\bar{N}^{010}, \tag{7.139}$$

where $\bar{N}_\omega$ denotes a sequence with fixed concentration. Thus $\bar{N}^{010} = N^{001}$. Similarly, we have

$$001+010 \rightleftharpoons 0010+01 \rightleftharpoons 00+0110 \rightleftharpoons 000+011, \tag{7.140}$$

which leads to

$$\bar{N}^{010}N^{001} = \left(\bar{N}^{010}\right)^2 = N^{011}N^{000}. \tag{7.141}$$

As we have by no means biased 011 over 000 (equal rates, no energy landscape, equal injection of mobile 0 and 1), by symmetry we require

$$N^{011} = N^{000} = \bar{N}^{010}. \tag{7.142}$$

We have chemostatted one species of length 3 at $\bar{N}^{010}$, but the occupation of length 3 species $N_3 = N^{000} + N^{001} + N^{010} + N^{011} = 4\,\bar{N}^{010}$. The only species of length one possible in the system is 0, the concentration of this length is not amplified: $N_1 = \bar{N}^0$. This amplification can dramatically lower the threshold for the transition to unbalanced. If the mobile sequence has equal abundance of every monomer, the amplification becomes m^{L-1} for Attack-Exchange, where $L-1$ is the length of the mobile sequence.

Using the same arguments as before, we find Eq. (7.136), on the level of sequences. For our example, this gives:

$$N^{\Omega} = \bar{N}^0 \left(\frac{\bar{N}^{010}}{\bar{N}^0}\right)^{\frac{|\Omega|-|0|}{|010|-|0|}} = \bar{N}^0 \left(\frac{\bar{N}^{010}}{\bar{N}^0}\right)^{\frac{|\omega|-1}{2}} \tag{7.143}$$

where $|\omega|$ denotes a sequence length, e.g. $|010| = 3$. To now go the level of chain lengths, we take into account that a species of length k can have m^{k-1} different sequences (the first monomer is fixed)

$$N_k = \bar{N}^0\, m^{k-1} \left(\frac{\bar{N}^{010}}{\bar{N}^0}\right)^{\frac{k-|0|}{|010|-|0|}} = \bar{N}^0 \left(\frac{m^2\,\bar{N}^{010}}{\bar{N}^0}\right)^{\frac{k-1}{2}} = \bar{N}^0 \left(\frac{4\,\bar{N}^{010}}{\bar{N}^0}\right)^{\frac{k-1}{2}} \tag{7.144}$$

This detailed balance solution is very reminiscent of Eq 7.136. However, the effect of alternative sequences provide an entropic amplification, which dramatically lowers the barrier for entering the unbalanced regime. If the mobile sequences have no more than one of each monomer, we enter the unbalanced regime for $\frac{m^{k_2-k_1}\,\bar{N}_{\omega_2}}{\bar{N}_{\omega_1}} \geq 1$, where k_2 and k_1 are the lengths of sequences ω_2 and ω_1.

It should be stressed that Eq. (7.144) is not unphysical per se. If we introduce a maximum length that species can attain by removing any chemistry leading to larger species, we are left with a detailed-balance solution for a finite-sized system. This is illustrated in Fig 7.6 , showing the stationary distribution of a Gillespie[53] simulation of the $\Omega_Y \in \{0,010\}$ system with a length cutoff and which obeys (7.143).

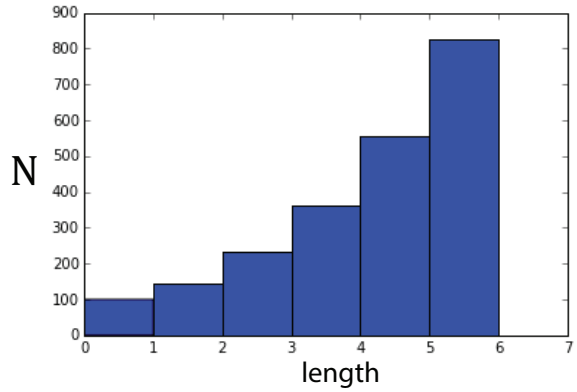


Figure 7.6: Equilibrium polymer length distribution for an Attack-Exchange reaction with chemostats set at $\Omega_Y = \{0,010\}$. Polymers length cannot exceed 6 (hexamers), reservoir concentrations are $\bar{c}_0 = 100, \bar{c}_{010} = 60$, which is sufficient for unbalanced growth.

A system with a length cutoff can have a perfectly physical solution. When looking at its dynamics, the polymer length distribution does not see its own cutoff on short timescales (provided we start with small enough oligomers). On this timescale there is no distinction between the dynamics of a system with a cutoff and a system without one. On longer timescales, they may no longer be captured by the same approximations. The unbalanced regime without a cutoff has an equilibrium state that is not described properly by the dilute ideal solution limit.

2 chemostats, biased monomer composition

To extend the entropic amplification to more general sequences, we need to take into account the abundance of monomers within the chemostatted sequence. E.g, if we were to chemostat 000000010, a sequence rich in 1's would be much more rare than a sequence rich in 0's. We will proceed by deriving a general formula, while illustrating it with a simple example.

Suppose we chemostat 0 and 0100 and have them perform the Attack-Exchange reaction. The mobile sequence is 100, which contains more 0's than 1's. We can form 2^3 different sequences of length 3 using 0's and 1'. At equilibrium, only the abundance of a monomer is important to determine the concentration of a sequence. Therefore 0100, 0010 and 0001 become equally abundant. We will refer to a mobile sequence to designate the part of the sequence that can be scrambled, as opposed to the first monomer that is fixed. E.g. in 0ω, ω is the mobile sequence and 0 is fixed. We find the concentration of all monomers, by assigning a probability p_i of incorporation to every monomer, and then factorize these to calculate the probability of a mobile sequence ω of given length L.

$$p_\omega = \prod_{\omega_j=\omega_1}^{\omega_L} p_i \delta^i_{\omega_j}. \tag{7.145}$$

This probability per monomer position is normalized ($\sum_i p_i = 1$), and this also normalizes the multinomial containing all mobile sequences of length L: $(\sum_i p_i)^L = 1$, with $p_i = \frac{\chi^i_{\tilde{\omega}}}{\sum_j \chi^j_{\tilde{\omega}}}$. For example, the mobile sequence 100 gives us: $p_1 = \frac{1}{2+1} = \frac{1}{3}$ and $p_0 = \frac{2}{3}$. If we were given a random sequence from the population, the probability to observe 0110 out of all mobile sequences of length 3 (mobile sequence ω =110) would be $p_1 p_1 p_0 = \frac{2}{27}$.

We now use that at equilibrium

$$\frac{\bar{N}^{0100}}{N_4} = p_{100} \tag{7.146}$$

and thus:

$$N_4 = \frac{\bar{N}^{0100}}{p_{100}} = \frac{27}{4}\bar{N}^{0100}. \tag{7.147}$$

The entropic amplification is now $\frac{27}{4}$. Let us denote the mobile sequence as: $\tilde{\omega}$, the entropic amplification A_0 becomes

$$A_0 = \frac{1}{p_{\tilde{\omega}}} = \frac{(\sum_j \chi^j_{\tilde{\omega}})^L}{\prod_i (\chi^i_{\tilde{\omega}})^{\chi^i_{\tilde{\omega}}}} = \frac{|\tilde{\omega}|^{|\tilde{\omega}|}}{\prod_i (\chi^i_{\tilde{\omega}})^{\chi^i_{\tilde{\omega}}}}. \tag{7.148}$$

This reduces to a factor m^m if $\tilde{\omega}$ is composed of m distinct monomers that appear each 1 single time. For $m = 1$, or empty sequences, $A_0 = 1$.

We can now write:

$$N_k = N^0 \left(A_0 \frac{\bar{N}^{0\bar{\omega}}}{\bar{N}^0} \right)^{\frac{k-1}{|\bar{\omega}|}} = N^0 \left(\frac{\bar{N}^{0\bar{\omega}} \, |\tilde{\omega}|^{|\bar{\omega}|}}{\bar{N}^0 \, \prod_i (\chi^i_{\tilde{\omega}})^{\chi^i_{\tilde{\omega}}}} \right)^{\frac{k-1}{|\bar{\omega}|}} \tag{7.149}$$

$$N^{0\omega} = N_{|\omega|+1} p_\omega = N^0 \left(\frac{\bar{N}^{0\bar{\omega}} \, |\bar{\omega}|^{|\bar{\omega}|}}{\bar{N}^0 \, \prod_i (\chi^i_{\bar{\omega}})^{\chi^i_{\bar{\omega}}}} \right)^{\frac{k-1}{|\bar{\omega}|}} \left(\frac{\prod_i (\chi^i_{\bar{\omega}})^{\chi^i_{\bar{\omega}}}}{|\bar{\omega}|^{|\bar{\omega}|}} \right) \tag{7.150}$$

In this simple treatment of unbalanced systems we chose 0 as the smallest sequence for computational convenience. The analytical treatment becomes a lot more laborious for longer sequences, especially if the chemostats show biases towards different monomer types.

More chemostats

The above derivation of an unbalanced regime supposes a detailed-balance solution. Upon increase of the number of chemostats, we may introduce affinities, which hamper the validity of detailed balance. The analysis of such systems is more complicated, but the intuitions derived in the previous sections remain valid: fixing larger lengths at higher amplitude makes the system unbalanced.

Unbalanced growth occurs because the smallest sequences are fixed, leaving the system little freedom to store mass in small species. If we are to chemostat $0, 101$ (Attack-Exchange dynamics), then this limitation does not present itself since mass can be stored in other sequences. We then expect a balanced system.

Some implications

By opening systems up to an environment, interesting new nonequilibrium behavior becomes accessible. Of particular interest is the observation that arbitrarily long polymers can be made through exchange reactions, without the need of chemical activation. The generation of long polymers is a key ingredient in many prebiotic scenarios, and coupling oligomer recombination to reservoir exchange may provide an elegant means to polymerize species that inherently oppose elongation, such as RNA.

7.4 Polymer Adsorption on Minerals

Minerals occupy an important place in the OOL literature and prebiotic polymer scenarios. Cairns-Smith considered them as the first templates[56], Wächtershauser considered them as a substrate to concentrate reactants for a surface metabolism[57] and the recent Mica-first scenario[58] highlights how porous sheets may have supplied compartments. New pioneering experiments use serial transfer of mineral particles to support an ecological mechanism of chemical evolution [59]. For RNA, it has been shown that minerals can protect against degradation[60, 61], promote the synthesis of building blocks[62, 63] and the ligation of activated RNA[64, 65].

In 1965, Bernardi reported[66] how hydroxyapatite could be used for the chromatographic separation of RNA strands. In 1980, Gibbs, Lohrmann and Orgel[67] used the same mineral to selectively adsorb the high molecular weight products of a template-directed synthesis polyadenylates, and posited that mineral surfaces could play a major role in RNA world scenarios, by sequestering long RNA and releasing small RNA for reactivation. Later studies have shown a more favorable picture for adsorption of short oligomer species, partly owing to the fact that a surface can accommodate more small species than large ones[68]. Since these works studied adsorption for one polymer length at a time, however, they do not capture the effect of a mixture of polymers competing for adsorption sites.

In this section, a simple toy model for adsorption of polymers from a polymer mixture on mineral surfaces is considered. This model was derived to rationalize experimental results on RNA adsorption and its temperature dependence for 5 different mineral species, in a joint study

with the labs of D. Baum and N. Lehman. Some of these results are reproduced here for the purpose of illustration. The corresponding publication (Ref. [27]) goes further, notably by showing how recombination can be combined with mineral adsorption to promote the formation of large polymers.

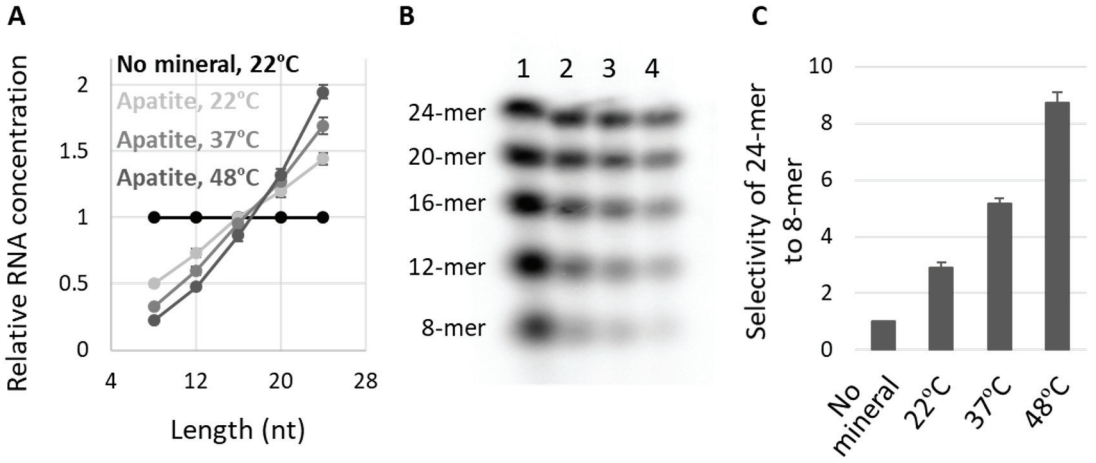


Figure 7.7: Figure taken from Ref. [27]. RNA adsorption on minerals as function of temperature. The adsorption experiment was performed by incubating a mixture of $8-, 12-, 16-, 20-, 24-$mer fully random RNAs (0.6 μM each) and 0.2 mg apatite in 10 μl for 2h at 22C, 37C, or 48C. (A) Concentration of each length RNA on the surfaces, normalized to the levels of a control reaction performed at RT, in the absence of minerals. (B) An example of an analyzed gel image. Lanes 1, no mineral, 22C; 2, + apatite, 22C; 3, + apatite, 37C; 4, + apatite, 48 C. (C) Selectivity of $24-$mer RNAs to $8-$mers, calculated as the concentration of $24-$mers relative to $8-$mers. The selectivity was set to 1 for the in absence of mineral. In all panels, the error bars indicate standard errors ($n = 3$).

7.4.1 Adsorption of oligomers on a lattice

We first derive a model for the simultaneous adsorption of different oligomers on a 1D surface, to obtain exact expressions for the surface fraction covered by each oligomer in a low-dimensional case. We then consider some extensions of the model to higher dimensions for stiff and flexible oligomers. These approaches to extend the dimension are derived from the works of Ramirez-Pastor and colleagues [69, 70], where they have been applied to the case of a single adsorbent.

Adsorption on a line

We start by considering a large solution of RNA oligomers, each maintained at a fixed dimensionless concentration $\bar{c}$. In addition, the solution contains a mineral, with an exposed surface on which oligomers can adsorb. We consider the exposed surface to have M adsorption sites, with a size comparable to a single monomer. Correspondingly, to fully adsorb a k-mer, k adsorption sites need to be occupied. For our purposes, the RNA solution contains $8-, 12-, 16-, 20-$ and $24-$mers, and whenever we take a sum (i.e. $\sum$) it will denote a sum over these values. We will start the simplest case, for which the surface can be modeled as a line, which is depicted in Fig. 7.8. Let us denote with N_k the number of adsorbed RNA oligomers of length k. In total, these oligomers occupy kN_k mineral sites. Consequently, we find that the number of unoccupied mineral sites $N_\emptyset$ can be written as

$$N_\emptyset = M - \sum_i iN_i \tag{7.151}$$

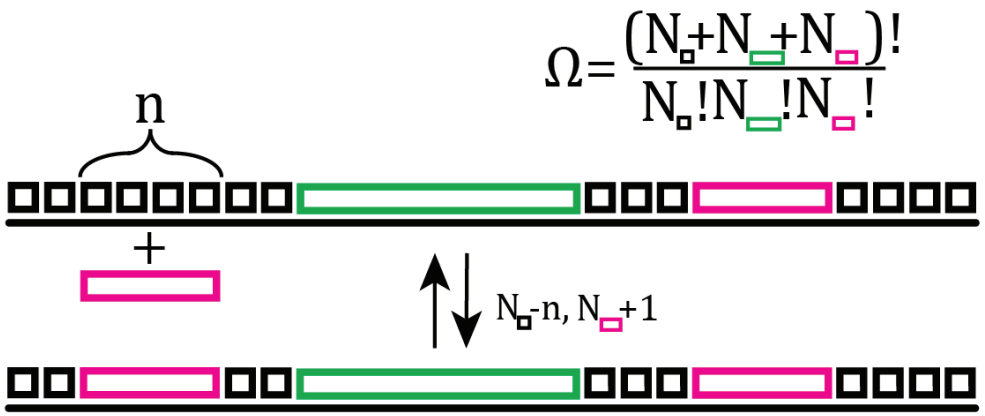


Figure 7.8: A 1D mineral surface. Black squares represent empty sites, green and pink bars represent oligomer types. Ω is given by the number of permutations among oligomers and empty sites. The reversible adsorption of one pink oligomer replaces k empty sites.

We denote by W the number of empty mineral sites plus the number of adsorbed species

$$W = N_\emptyset + \sum_i N_i \tag{7.152}$$

A surface state is completely described by the exact sequence in which the surface bound molecules and empty sites appear. The number of states is consequently given by all their possible permutations

$$\Omega(\{N_i\},M) = \binom{W}{N_8,N_{12},N_{16},N_{20},N_{24},N_\emptyset} = \frac{W!}{N_8!N_{12}!N_{16}!N_{20}!N_{24}!N_\emptyset!} \tag{7.153}$$

We now introduce a reference standard free energy for a k-mer adsorbed on a mineral surface, $\mu^\circ_{k,min}$, for which we expect that (most) thermodynamic contributions are either constant or linear (free energy of formation, adsorption on k sites), especially since these are determined by the local chemical environment which changes little along the oligomer. Our starting point is then an affine function of k

$$\mu^\circ_{k,min} = a_0 + a_1 k, \tag{7.154}$$

where we suppose that $a_1 < 0$. We write the canonical ensemble $Q(\{N_i\},M,T)$ for a covered mineral via

$$Q(\{N_i\},M,T) = \Omega(\{N_i\},M)\exp\left(-\beta\sum_i N_i\mu^\circ_{i,min}\right), \tag{7.155}$$

where $\beta = 1/k_bT$, with k_b Boltzmann's constant and T the absolute temperature. We can then extract the Helmholtz free energy F by

$$\beta F(\{N_i\},M,T) = -\ln Q(\{N_i\},M,T) = -\ln\Omega(\{N_i\},M) + \beta\sum_i N_i\mu^\circ_{i,min} \tag{7.156}$$

We can rewrite $\ln\Omega(\{N_i\},M)$ by performing a Stirling approximation $\ln N! = N\ln N - N + O(\ln N)$

$$\ln\Omega(N_i,M) = W\ln W - \sum_i N_i\ln N_i - N_\emptyset\ln N_\emptyset \tag{7.157}$$

Eq. (7.157) can be rewritten to

$$\ln\Omega(N_i,M) = -\sum_i N_i \ln\frac{N_i}{W} - N_\emptyset \ln\frac{N_\emptyset}{W}, \tag{7.158}$$

which has the functional form of a mixing entropy (but taken relative to W instead of M), which gives a clear interpretation of this object. Because the quantity W depends on the surface coverage, it is more practical to define (7.159) relative to the fixed number of sites M. Let us therefore write

$$\begin{aligned}\ln\Omega(N_i,M) &= W\ln\frac{W}{M} - \sum_i N_i \ln\frac{N_i}{M} - N_\emptyset \ln\frac{N_\emptyset}{M}, \\ &= M\left[\theta_W \ln\theta_W - \sum_i \frac{\theta_i}{i}\ln\frac{\theta_i}{i} - \theta_0\ln\theta_0\right].\end{aligned} \tag{7.159}$$

Where we have introduced $\theta_k = kN_k/M$, (fraction of sites covered by k-mers), $\theta_0 = N_\emptyset/M$ (fraction of empty sites), $\theta_W = W/M$ (fraction of empty sites and oligomers). The $\theta_W \ln\theta_W$ term in Eq. (9) is a consequence of multisite adsorption. We can now extract the chemical potential of adsorbed polymers of length k, using

$$\mu_{k,min} = \left(\frac{\partial F}{\partial N_k}\right)_{T,M,N_i\neq N_k} \tag{7.160}$$

which after taking the appropriate derivatives affords the expression

$$\mu_{k,min} = \mu^\circ_{k,min} + k_bT\left[\ln\frac{\theta_k}{k} + (k-1)\ln\frac{\theta_W}{\theta_0} - \ln\theta_0\right] \tag{7.161}$$

If our mixture would contain only a single type of oligomer of length k, we recover the isotherm in [69].

Let us now put our system in contact with a large solution of oligomers, maintained at a dimensionless concentration $\bar{c} = c_k/c^\circ$. Where c° is a standard concentration (1 M). The oligomers in solution have a chemical potential

$$\mu_k = \mu^\circ_k + k_bT\ln\bar{c} \tag{7.162}$$

where μ°_k is a standard free energy of formation at concentration c°. We consider μ°_k to be an affine function of k

$$\mu^\circ_k = b_0 + b_1 k \tag{7.163}$$

A reversible adsorption process will lead to chemical equilibrium, at which point $\mu_{k,min} = \mu_k$. Let us now substitute Eq. (13) in Eq. (11), and write

$$\Delta\mu^*_{k,ads} + k_bT\left[\ln\frac{\theta_k}{k} + (k-1)\ln\frac{\theta_W}{\theta_0} - \ln\theta_0\right] = 0 \tag{7.164}$$

From Eqs. (7.154), (7.163), (7.162), it follows that $\Delta\mu^*_{k,ads}$ is again an affine function. We write

$$\Delta\mu^*_{k,ads} = \varepsilon + \delta k \tag{7.165}$$

where $\varepsilon = a_0 + b_0 - k_bT\ln\bar{c}$ and $\delta = a_1 - b_1$. Eq. (7.164) then gives the equilibrium surface coverage for k-mers

$$\theta_k = k\exp\left(-\beta(\varepsilon+\delta k)\right)\left(\frac{\theta_0}{\theta_W}\right)^{k-1}\theta_0 \tag{7.166}$$

Let us define

$$r_k = \frac{\theta_k/k}{\sum_i \theta_i/i} \tag{7.167}$$

as the relative surface concentration of k-mers. The ratio of k-mers to j-mers is then

$$\frac{r_k}{r_j} = \frac{j\theta_k}{k\theta_j} = \left(\frac{\theta_0}{\theta_W}\right)^{k-j} \exp\left(-\beta\delta(k-j)\right). \tag{7.168}$$

As $W \geq N_\emptyset$, $\theta_0/\theta_W \leq 1$. If we consider $k > j$, we see that the entropic term $\left(\frac{\theta_0}{\theta_W}\right)^{k-j}$ favors shorter oligomers. This is to be expected, as smaller oligomers allow for more possible surface configurations. Since $\delta < 0$, the factor $\exp\left(-\beta\delta(k-j)\right)$ favors longer oligomers. From Eq. (7.168) we deduce that the relative concentrations of adsorbed oligomers follow an exponential trend. We will now extend the model by relaxing the 1D assumption. Interestingly, this can largely be taken account by simply shifting the constants ε and δ. Consequently, we can proceed with an approximative model by using Eq. (7.167).

7.4.2 Extensions towards 2D

Connectivity ansatz

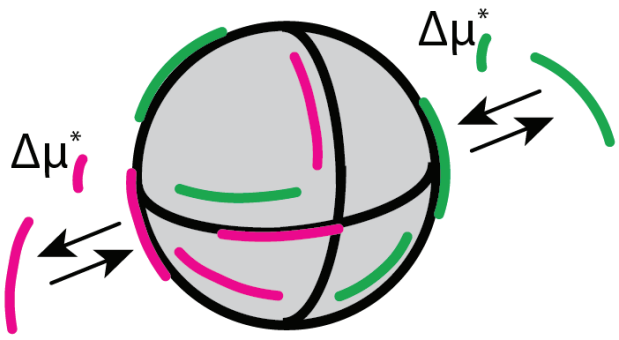


Figure 7.9: Reversible adsorption of two types of oligomers (pink and green) from bulk to mineral surface. Associated with this transition is a free energy change $\Delta\mu^*_{k,ads}$ containing all contributions other than mineral configurations.

As shown by Ramirez-Paster et al in Ref. [69], an effective way to describe stiff oligomers on a 2D lattice is by introducing a connectivity ansatz. Let c be the number of connections of a lattice point (for a 2D square lattice: 4, on the line: 2). By supposing Ω scales with dimension c as in the Flory model[45], the argument by Ramirez-Pastor et al leads to

$$\frac{\Omega(M,\{N_i\},c)}{\Omega(M,\{N_i\},c')} = \left[\frac{c-1}{c'-1}\right]^{\sum_i N_i(i-1)}. \tag{7.169}$$

Note that Ref. [69] considered the adsorption of single type of oligomer, whereas Eq. (7.169) concerns a mixture.

Performing our previous calculation and setting $c' = 2$, we find for arbitrary c

$$\mu_{k,min,c} = \mu_{k,min,2} - kT(k-1)\ln(c-1), \tag{7.170}$$

which means we can incorporate it in $\Delta\mu^*_{ads}$ by defining

$$\varepsilon' = \varepsilon + k_bT\ln(c-1), \tag{7.171}$$

$$\delta' = \delta + k_bT\ln(c-1). \tag{7.172}$$

2D: dilute lattice placements

Another extension, put forward in Ref. [70], is to study a dilute limit and consider the number of ways an oligomer can be placed on a lattice. For stiff oligomers, which can only be placed on square lattices, we can then consider every 1D placement and place them along all $c/2$ directions. The microcanonical partition function then grows as:

$$\frac{\Omega(M,N,c)}{\Omega(M,N,c')} = \left[\frac{c}{c'}\right]^{\Sigma_i N_i} \tag{7.173}$$

Which yields a constant correction to μ_k

$$\mu_{k,min,c} = \mu_{k,min,2} - k_b T \ln(c/2) \tag{7.174}$$

It can be absorbed in the expression for $\Delta\mu^*_{ads}$, by defining $\varepsilon' = \varepsilon + k_b T \ln(c/2)$.

For dilute flexible oligomers, the number of single-oligomer configurations $\gamma(c,k)$ is the number of self-avoiding random walks of length k. On a square lattice, this quantity behaves as $\gamma(c,k) = u^k k^v$, with $u_{2d} \approx 2.62$, $v_{2d} = 11/32$. On a 3D lattice, we have $v_{3d} \approx 0.16$. The correction for dilute systems is then

$$\mu_{k,min,4} = \mu_{k,min,2} - k_b T \left[k \ln\left(\frac{u_{2d}}{u_{3d}}\right) + (v_{2d} - v_{3d}) \ln k\right]. \tag{7.175}$$

Just as with the connectivity ansatz, we can absorb a contribution proportional to k, since we can write $\delta = \delta' - k_b T \ln u$. The $\ln k$ contribution gives a new factor $(\frac{k}{j})^{v_{2d}-v_{3d}}$, and we now obtain

$$\frac{r_{k,c}}{r_{j,c}} = \frac{j\theta_k}{k\theta_j} = \left(\frac{\theta_0}{\theta_W}\right)^{k-j} \exp(-\beta\delta'(k-j)) \left(\frac{k}{j}\right)^{v_{2d}-v_{3d}} \tag{7.176}$$

However, since $v_{2d} - v_{3d} \approx 0.18$, this contribution is relatively small, and we will neglect it in our further derivation. Overall, we see that the extension of the model to 2D for stiff and flexible polymers can be accounted for by shifting the parameters in the adsorption energy. In the subsequent sections, we will solve the model for Eq. (7.166).

7.4.3 Solving for Θ_i

Since the solutions are expressed in terms of θ_0/θ_W, we do not have a full solution. Expressed in terms of θ_i, we find

$$\theta_W = \sum_i \frac{\theta_i}{i} + \theta_0 \tag{7.177}$$

$$\theta_0 = 1 - \sum_i \theta_i \tag{7.178}$$

We can then write

$$\theta_W = \sum_i \exp(-\beta(\varepsilon + \delta i))(\theta_0/\theta_W)^{i-1}\theta_0 + \theta_0 \tag{7.179}$$

Let us define $\zeta = \frac{\theta_0}{\theta_W}$. From Eq. we then find

$$\sum_i \exp(-\beta(\varepsilon + \delta i))\zeta^i + \zeta - 1 = 0 \tag{7.180}$$

Which is a nonlinear polynomial equation from which we need a real root $\zeta < 1$. We can then express θ_0 as

$$\theta_0 = \frac{1}{1 + \sum_i i \exp(-\beta(\varepsilon + \delta i))\zeta^{i-1}}, \tag{7.181}$$

and thus we can numerically solve the system of equations by finding ζ.

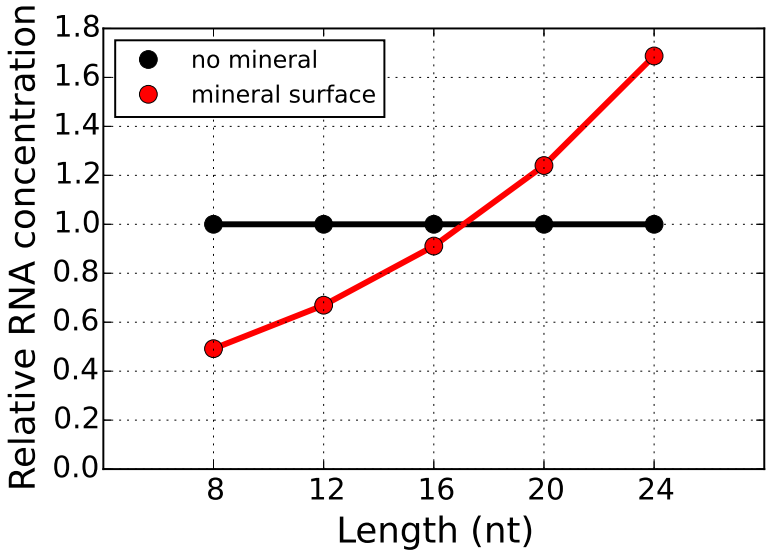


Figure 7.10: Distribution of adsorbed RNA on a mineral surface $\varepsilon = 4k_bT^*, \delta = -0.5k_bT^*, T^* = 293K$.

7.4.4 Temperature dependence

It was observed [27] that the relative abundance of longer RNAs increases at higher temperature. To investigate this effect, let us again consider the quantity r_k/r_j in Eq. (17) where $k > j$, and take its derivative with respect to temperature T

$$\frac{d\left(\frac{r_k}{r_j}\right)}{dT} = \left(\frac{r_k}{r_j}\right)\left[\frac{\delta - T\frac{d\delta}{dT}}{k_bT^2}(k-j) + (k-j)\frac{1}{\zeta}\frac{d\zeta}{dT}\right]. \tag{7.182}$$

As δ and ε correspond to a Gibb's free energy change, we can write them as enthalpies $\Delta h_\delta, \Delta h_\varepsilon$ and entropies $\Delta s_\delta, \Delta s_\varepsilon$

$$\delta = \Delta h_\delta - T\Delta s_\delta, \tag{7.183}$$
$$\varepsilon = \Delta h_\varepsilon - T\Delta s_\varepsilon, \tag{7.184}$$

and thus

$$\delta - T\frac{d\delta}{dT} = \Delta h_\delta. \tag{7.185}$$

Taking the derivative with respect to T of Eq. (7.178), we find

$$\left(\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\, i\, \zeta^{i-1} + 1\right)\frac{d\zeta}{dT} = -\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)(\Delta h_\varepsilon + i\Delta h_\delta)/k_bT^2\zeta^i, \tag{7.186}$$

which can be rewritten give

$$\frac{d\zeta}{dT} = \frac{\frac{\Delta h_\delta}{k_bT^2}\zeta - \frac{\Delta h_\varepsilon}{k_bT^2}\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\zeta^i}{\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\, i\, \zeta^{i-1} + 1} - \frac{\Delta h_\delta}{k_bT^2}\zeta \tag{7.187}$$

Plugging this back in Eq. we then have

$$\frac{d\left(\frac{r_k}{r_j}\right)}{dT} = \left(\frac{r_k}{r_j}\right)(k-j)\left[\frac{\frac{\Delta h_\delta}{k_bT^2} - \frac{\Delta h_\varepsilon}{k_bT_b^2}\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\zeta^{i-1}}{\sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\, i\, \zeta^{i-1} + 1}\right] \tag{7.188}$$

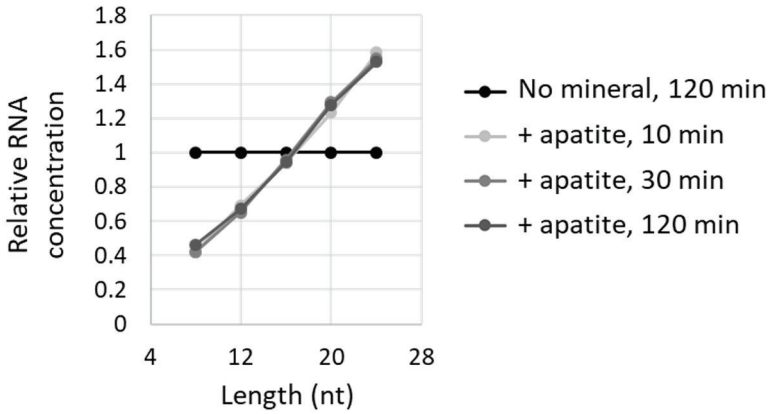


Figure 7.11: Experimental time course of the change of RNA concentrations on mineral surfaces. Adsorption experiment was performed by incubating a mixture of $8-,12-,16-,20-,24-$mer fully random RNAs (0.6 μM each) and 0.2 mg apatite in 10 μl at $293K$ for 10 min, 30 min, or 120 min. The concentrations were determined by radioactivity of 32P-labeled RNA and normalized to the levels of the control reaction (120 min) performed without minerals. Figure taken from Ref. [27]

It follows that selectivity can increase with temperature, provided that the enthalpic contributions obey

$$\Delta h_\delta - \Delta h_\varepsilon \sum_i \exp\left(-\beta(\varepsilon+\delta i)\right)\zeta^{i-1} > 0 \tag{7.189}$$

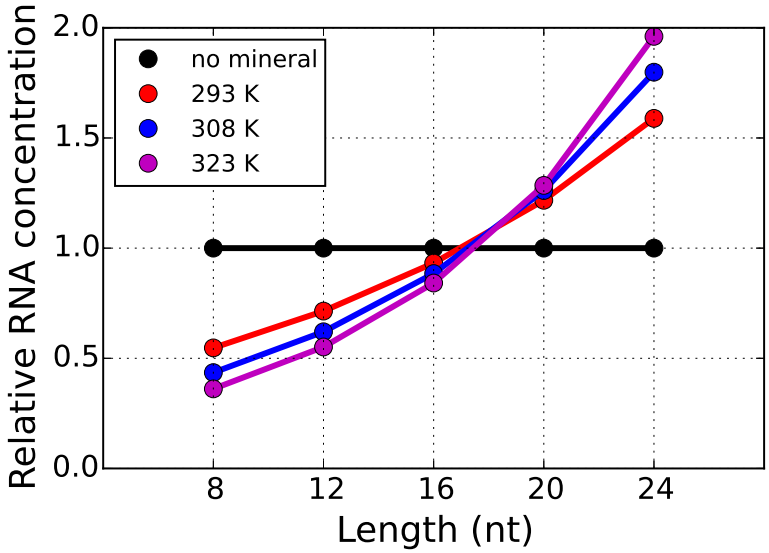


Figure 7.12: $\varepsilon = 3.4k_bT^*$,$\Delta_\varepsilon = -11.5k_bT^*$, $\delta = -0.8k_bT^*$,$\Delta_\delta = -2.5k_bT^* = 293K$

Example: consider Fig. 7.12 , for which $\varepsilon = 3.4k_bT^*$, $\delta = -0.8k_bT^*$ and $r_{24}/r_8 = 2.90$. Numerically, we find that selectivity would increase with T in this case if $-0.884\Delta h_\varepsilon > -\Delta h_\delta$. As an illustration, let us choose $\Delta h_\varepsilon = -11.5k_bT^*$, $h_\delta = -2.5k_bT^*$. Augmenting the temperature with

$\Delta T = 10K$, we would then have

$$\frac{d\left(\frac{r_k}{r_j}\right)}{dT}\Delta T \approx 0.072 \cdot 10 = 0.72 \tag{7.190}$$

Which corresponds well with the order of magnitude observed in the experiment depicted in Fig. 7.7.

The minimal model employed in this section provides a simple explanation why longer polymers can adsorb preferentially with relative ease. By construction, it yields increasing exponential distributions. In practice, distributions, such as in Fig. 7.7 can deviate from an exponential. An equilibrium argument for such deviations can be found when nonaffine contributions are introduced, such as interactions between adsorbed polymers.

Thermodynamic equilibrium may not always be an appropriate assumption for mineral adsorption. For example, the phosphodiester bonds of RNA are degraded by hydrolysis, which disproportionally acts on long polymers. On the timescale of the experiment, however, hydrolysis was relatively slow and adsorption profiles measured after 10, 30 and 120 minutes showed little change.

7.5 Activation, ligation and fragmentation

Many scenarios for the origins of life share a key step: the elongation of the first (bio)monomers through abiotic means. In this section, we revisit a large class of these scenarios, involving the use of ligation-fragmentation models, by explicitly including chemical activation. The steady state polymer length distribution can be described with two dimensionless quantities, which quantify the relative rates of activation and ligation with respect to hydrolysis. When activation is the slowest step, the addition of template-assisted ligation does not lead to a further increase of the average polymer length. On the other hand, folding and hybridization become more effective in increasing the size of polymers in this regime. The inclusion of the activation step also makes it possible to use nonequilibrium thermodynamics to make general, model-independent statements about e.g. dissipation. In particular, we find that there is a minimum cost associated with sequence exploration, irrespective of catalysis. This cost is a property of the activation chemistry and is absent for nondissipative reaction schemes that perform such exploration, such as recombination reactions. Our statements lead to a number of quantitative requirements for prebiotic scenarios and novel plausibility criteria based on the second law of thermodynamics.

7.5.1 Chemical Activation

When a polymer is chemically activated, it contains a leaving group that favors a subsequent ligation. Such group-transfer reactions are pivotal to life and are often considered to be essential to prebiotic chemistry [71]. This should come as no surprise when one considers that the reversible assembly of many biopolymers through condensation reactions is not favorable in water. For RNA, one typically only needs to consider the reverse of condensation: hydrolysis, which degrades the polymer.

An important problem that a putative RNA world must solve is getting its monomers to polymerize, which modern biology achieves through chemical activation. Some proposed candidates for a prebiotic activation in the case of RNA are triphosphate [63], 2',3'-cyclic phosphate [72] (both possibly generated by diamidophosphate [73]), 2-aminoimidazole [74], 2-methylimidazole [75]. The activation step generalizes these different schemes.

7.5.2 Model Setup

In this section, we will develop a general scheme for elongation in the presence of activation. The model is composed of three reactions: i) activation, ii) ligation, iii) hydrolysis (fragmentation).

Step i) is implicit in virtually all of the literature models. Wu and Higgs [21] dedicated attention to activation in polymer models. In later work, this step became unnecessary for the problems under consideration [22]. We wil show here that an explicit description of activation is quite valuable: it allows to use nonequilibrium thermodynamics and places essential constraints on the prebiotic conditions. We will refer to this scheme as activation-ligation-fragmentation (ALF).

Activation

We will denote an unactivated polymer of length n as $\mathbf{n}$. To denote an activated species, we will add an asterisk: $\mathbf{n}^*$. When it is convenient to use a sum, we will use brackets for single species with length $n+m$: $[\mathbf{n}+\mathbf{m}]$.

The chemical activation step proceeds through the species XY, which transfers the activating group Y to a polymer, as shown in Fig. 7.13

$$\mathbf{n} + \mathrm{XY} \rightleftarrows \mathbf{n}^* + \mathrm{XH}. \tag{7.191}$$

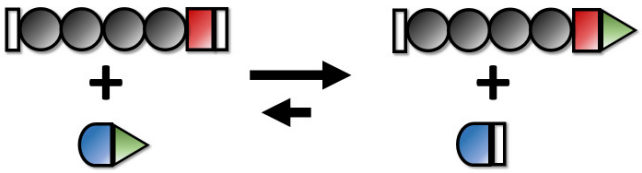


Figure 7.13: Schematic picture of activation. Green triangles: activating group Y, blue hemisphere: X.

Ligation

An activated polymer $\mathbf{n}^*$ can perform one ligation reaction, at the expense of its activating group Y, as shown in Fig. 7.14

$$\mathbf{n}^* + \mathbf{m} \rightleftarrows [\mathbf{n}+\mathbf{m}] + \mathrm{YOH}, \tag{7.192}$$
$$\mathbf{n}^* + \mathbf{m}^* \rightleftarrows [\mathbf{n}+\mathbf{m}]^* + \mathrm{YOH}. \tag{7.193}$$

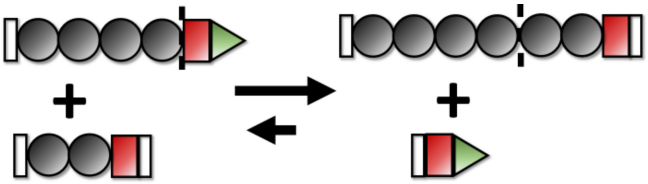


Figure 7.14: Schematic picture of ligation.

Fragmentation

As the solvent under consideration is typically water, condensation preferably runs in reverse for most condensation polymers, and this gives hydrolysis (fragmentation, see Fig. 7.15).

$$[\mathbf{n}+\mathbf{m}] + \mathrm{H_2O} \rightleftarrows \mathbf{n}+\mathbf{m}, \tag{7.194}$$
$$[\mathbf{n}+\mathbf{m}]^* + \mathrm{H_2O} \rightleftarrows \mathbf{n}+\mathbf{m}^*. \tag{7.195}$$

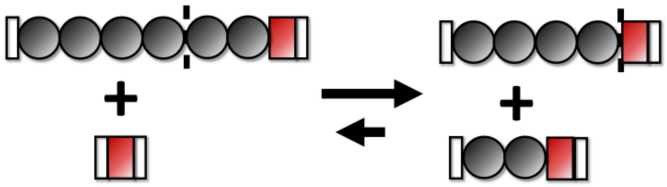


Figure 7.15: Schematic picture of hydrolysis. Gray circles: monomers (**1**), Red squares: oxygen (O), white rectangles: hydrogen (H).

7.5.3 ALF Cycles

By putting the aforementioned reactions together, we can construct a cycle of subsequent activation, ligation and hydrolysis (Fig. 7.16). After one such cycle, the net reaction we have performed is

$$XY + H_2O \rightleftharpoons HX + YOH \tag{7.196}$$

In order to maintain a steady state polymer distribution, we need a constant influx of XY and H_2O and an outflux of HX and YOH, which can e.g. be achieved by chemostatting these compounds. From the works of the group of M. Esposito [47, 76] we know that we can perform a cycle decomposition to describe all steady state currents within a chemical system, where a cycle is a collection of reactions that leaves the system unchanged (chemostatted species may be converted and exchanged, however). For our present purposes, it suffices to note that at steady state, every ligation is accompanied by a hydrolysis and an activation. In Appendix 10.4. the cycle decomposition is discussed in more detail.

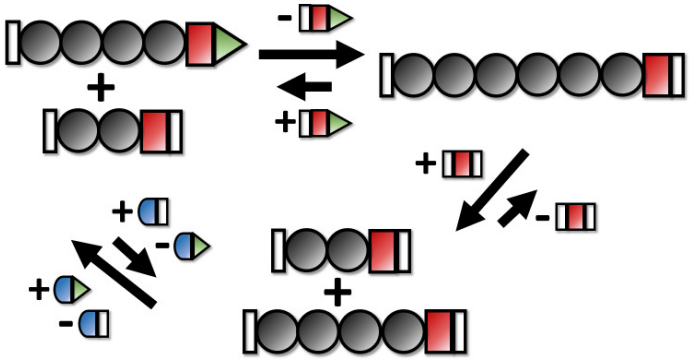


Figure 7.16: Schematic picture of an ALF cycle.

Kinetics and cycles

We introduce the following quantities:
N_n^D, N_n^A: number of unactivated, activated polymers of length n, respectively.
N_n: number of polymers of length n, such that $N_n = N_n^D + N_n^A$.
$N_{XY}, N_{H_2O}, N_{HX}, N_{YOH}$: number of XY, H_2O, HX, YOH molecules.
k^+, k^-: forward, backward rate constant respectively. We assume a fixed system volume V and absorb all volume dependence in the rate constants. The net rates of the ligation reactions $\mathbf{n}^* + \mathbf{m}$ and $\mathbf{n}^* + \mathbf{m}^*$ become

$$J_{\mathbf{n}^*+\mathbf{m}} = k_{lig}^+ N_n^A N_m^D - k_{lig}^- N_{n+m}^D N_{YOH}, \tag{7.197}$$

$$J_{\mathbf{n}^*+\mathbf{m}^*} = k_{lig}^+ N_n^A N_m^A - k_{lig}^- N_{n+m}^A N_{YOH}. \tag{7.198}$$

For hydrolysis reactions $[\mathbf{n}+\mathbf{m}] \rightarrow \mathbf{n}+\mathbf{m}$ and $[\mathbf{n}+\mathbf{m}]^* \rightarrow \mathbf{n}+\mathbf{m}^*$ we have

$$J_{[\mathbf{n}+\mathbf{m}]} = k^+_{hyd} N^D_{n+m} N_{\mathrm{H_2O}} - k^-_{hyd} N^D_n N^D_m, \tag{7.199}$$

$$J_{[\mathbf{n}+\mathbf{m}]^*} = k^+_{hyd} N^A_{n+m} N_{\mathrm{H_2O}} - k^-_{hyd} N^D_n N^A_m. \tag{7.200}$$

Finally, for activation $\mathbf{n} \rightarrow \mathbf{n}^*$ we can write

$$J_{\mathbf{n}} = k^+_{act} N^D_n N_{\mathrm{XY}} - k^-_{act} N^A_n N_{\mathrm{HX}} \ . \tag{7.201}$$

Taking these reactions together, the time evolution of N^A_n becomes

$$\begin{aligned} \frac{dN^A_n}{dt} &= J_{\mathbf{n}} + J_{\mathbf{k}^*+\mathbf{l}^*}\delta^n_{k+l} + J_{[\mathbf{k}+\mathbf{n}]^*} e_k - J_{[\mathbf{k}+\mathbf{l}]^*}\delta^n_{k+l} \\ &- J_{\mathbf{n}^*+\mathbf{k}} e_k - J_{\mathbf{n}^*+\mathbf{k}^*} e_k - J_{\mathbf{k}^*+\mathbf{n}^*} e_k, \end{aligned} \tag{7.202}$$

where any repeated index is summed according to the Einstein summation convention. Therefore, we have

$$J_{[\mathbf{k}+\mathbf{n}]} e_k = \sum_{k=1}^{\infty} J_{[\mathbf{k}+\mathbf{n}]}, \tag{7.203}$$

$$J_{\mathbf{k}^*+\mathbf{l}^*}\delta^n_{k+l} = \sum_{k+l=n} J_{\mathbf{k}^*+\mathbf{l}^*}, \tag{7.204}$$

in which we have introduced the Kronecker delta δ^b_a, ($\delta^b_a = 1$ if $a = b, \delta^b_a = 0$ if $a \neq b$) and the unit vector e_k. For the time evolution of N^D_n, we can write

$$\begin{aligned} \frac{dN^D_n}{dt} = -J_{\mathbf{n}} &+ J_{\mathbf{k}^*+\mathbf{l}}\delta^n_{k+l} + J_{[\mathbf{k}+\mathbf{n}]} e_k + J_{[\mathbf{n}+\mathbf{k}]^*} e_k \\ &+ J_{[\mathbf{n}+\mathbf{k}]} e_k - J_{[\mathbf{k}+\mathbf{l}]}\delta^n_{k+l} - J_{\mathbf{k}^*+\mathbf{n}} e_k. \end{aligned} \tag{7.205}$$

By combining Eq.(7.202) and (7.205), we obtain the rate of change in N_n. In doing so, we eliminate the reactivation contribution $J_{\mathbf{n}}$

$$\begin{aligned} \frac{dN_n}{dt} = [J_{[\mathbf{n}+\mathbf{k}]} &+ J_{[\mathbf{k}+\mathbf{n}]} + J_{[\mathbf{n}+\mathbf{k}]^*} + J_{[\mathbf{k}+\mathbf{n}]^*}] e_k \\ &+ [J_{\mathbf{k}^*+\mathbf{l}} + J_{\mathbf{k}^*+\mathbf{l}^*}]\delta^n_{k+l} - [J_{[\mathbf{k}+\mathbf{l}]} + J_{[\mathbf{k}+\mathbf{l}]^*}]\delta^n_{k+l} \\ &- [J_{\mathbf{n}^*+\mathbf{k}} + J_{\mathbf{k}^*+\mathbf{n}} + J_{\mathbf{n}^*+\mathbf{k}^*} + J_{\mathbf{k}^*+\mathbf{n}^*}] e_k \end{aligned} \tag{7.206}$$

7.5.4 Steady state currents and distributions

At steady state, the current across all cycles becomes fixed (see Appendix 10.4), as well as all concentrations. A solution for (7.206) that verifies $dN_n/dt = 0$ is

$$J_{\mathbf{n}^*+\mathbf{m}^*} + J_{\mathbf{n}^*+\mathbf{m}} = J_{[\mathbf{n}+\mathbf{m}]^*} + J_{[\mathbf{n}+\mathbf{m}]} \ , \tag{7.207}$$

Upon substitution of Eq. (7.197)-(7.200), (7.207) becomes

$$k^+_{lig} N^A_n N_m - k^-_{lig} N_{n+m} N_{\mathrm{YOH}} = k^+_{hyd} N_{n+m} N_{\mathrm{H_2O}} - k^-_{hyd} N^D_n N_m$$

We can infer from Eq. (7.208), by choosing $n = 1$, that

$$N_{m+1} = \left(\frac{k^-_{hyd} N^D_1 + k^+_{lig} N^A_1}{k^+_{hyd} N_{\mathrm{H_2O}} + k^-_{lig} N_{\mathrm{YOH}}} \right) N_m. \tag{7.208}$$

Via induction this immediately gives

$$N_m = \left(\frac{k^-_{hyd} N^D_1 + k^+_{lig} N^A_1}{k^+_{hyd} N_{\mathrm{H_2O}} + k^-_{lig} N_{\mathrm{YOH}}} \right)^{m-1} N_1. \tag{7.209}$$

Evaluating Eq. (7.208) for $m = 1$, we obtain

$$N_{n+1} = \Big(\frac{k^-_{hyd}N^D_n + k^+_{lig}N^A_n}{k^+_{hyd}N_{\mathrm{H_2O}} + k^-_{lig}N_{\mathrm{YOH}}}\Big)N_1. \tag{7.210}$$

From which one can show that

$$N^D_m = \Big(\frac{k^-_{hyd}N^D_1 + k^+_{lig}N^A_1}{k^+_{hyd}N_{\mathrm{H_2O}} + k^-_{lig}N_{\mathrm{YOH}}}\Big)^{m-1}N^D_1, \tag{7.211}$$

$$N^A_m = \Big(\frac{k^-_{hyd}N^D_1 + k^+_{lig}N^A_1}{k^+_{hyd}N_{\mathrm{H_2O}} + k^-_{lig}N_{\mathrm{YOH}}}\Big)^{m-1}N^A_1. \tag{7.212}$$

Which implies that we can split Eq. (7.207) in two equations

$$J_{\mathbf{n}^*+\mathbf{m}^*} = J_{[\mathbf{n}+\mathbf{m}]^*}, \tag{7.213}$$

$$J_{\mathbf{n}^*+\mathbf{m}} = J_{[\mathbf{n}+\mathbf{m}]}. \tag{7.214}$$

If we plug Eqs. (7.213)-(7.214) in (7.202) for $dN^A_n/dt = 0$, we obtain

$$J_{\mathbf{n}} = [J_{[\mathbf{n}+\mathbf{k}]^*} + J_{[\mathbf{n}+\mathbf{k}]}]e_k. \tag{7.215}$$

Steady-state distribution

The length distribution in ALF is exponential (see Eq. (7.210)), just as in (undecorated) ligation-fragmentation models [19, 49, 20]. Now, we introduce M, the total number of monomer units in all polymers: $M = \sum_n nN_n$ and N, the total number of polymers: $N = \sum_n N_n$. Using Eqs. (7.211) and (7.212) one finds

$$N = \frac{N_1}{1 - \frac{k^-_{hyd}N^D_1 + k^+_{lig}N^A_1}{k^+_{hyd}N_{\mathrm{H_2O}} + k^-_{lig}N_{\mathrm{YOH}}}}, \tag{7.216}$$

$$M = \frac{N_1}{\Big(1 - \frac{k^-_{hyd}N^D_1 + k^+_{lig}N^A_1}{k^+_{hyd}N_{\mathrm{H_2O}} + k^-_{lig}N_{\mathrm{YOH}}}\Big)^2}. \tag{7.217}$$

For the total number of activated and unactivated polymers (N^A, N^D) one can find similar expressions. Combining (7.216) and (7.217) we find the number of monomers

$$N_1 = \frac{N^2}{M} \tag{7.218}$$

For activation of monomers, we use Eq. (7.215) for $n = 1$

$$k^+_{act}N^D_1 N_{XY} - k^-_{act}N^A_1 N_{HX} = k^+_{hyd}(N - N_1)N_{H_2O} - k^-_{hyd}N^D_1 N \tag{7.219}$$

The activation of all species obeys

$$k^+_{act}N^D N_{XY} - k^-_{act}N^A N_{HX} = k^+_{hyd}(M - N)N_{H_2O} - k^-_{hyd}N^D N. \tag{7.220}$$

Let us denote by α the fraction of activated species, such that

$$N^A = N\alpha = \frac{M\alpha}{\chi}, \tag{7.221}$$

$$N^D = N(1-\alpha) = \frac{M(1-\alpha)}{\chi}. \tag{7.222}$$

We can then solve the full problem by finding α and χ, since

$$N_n^A = \frac{M\alpha}{\chi^2}\left(1-\frac{1}{\chi}\right)^n \tag{7.223}$$

$$N_n^D = \frac{M(1-\alpha)}{\chi^2}\left(1-\frac{1}{\chi}\right)^n \tag{7.224}$$

From Eq. (7.220), we find that

$$k_{act}^+(1-\alpha)N_{XY} - k_{act}^-\alpha N_{HX} = k_{hyd}^+(\chi-1)N_{H_2O} - k_{hyd}^-(1-\alpha)\frac{M}{\chi}. \tag{7.225}$$

which upon rewriting gives

$$\alpha = \frac{k_{act}^+ N_{XY} + k_{hyd}^+(1-\chi)N_{H_2O} + k_{hyd}^-\frac{M}{\chi}}{k_{act}^+ N_{XY} + k_{act}^- N_{HX} + k_{hyd}^-\frac{M}{\chi}} \tag{7.226}$$

By combining Eqs. (7.216) and (7.217), we find

$$\chi = \frac{1}{1-\frac{k_{hyd}^-\frac{M}{\chi}(1-\alpha)+k_{lig}^+\frac{M}{\chi}\alpha}{k_{hyd}^+ N_{H_2O}+k_{lig}^- N_{YOH}}}, \tag{7.227}$$

from which we can extract α

$$\alpha = \frac{(k_{hyd}^+ N_{H_2O} + k_{lig}^- N_{YOH})(1-\chi) + k_{hyd}^- M}{(k_{hyd}^- - k_{lig}^+)M}. \tag{7.228}$$

By combining Eqs. (7.226) and (7.228), χ can be found as the solution of a quadratic equation.

Strongly irreversible reactions

For our purposes, we are interested in the case where all forward reactions are much faster than backward ones (as otherwise no appreciable elongation can be expected). In this limit, we can neglect all k^- terms for the kinetics (but not for the thermodynamics).

We will now transform the following quantities w.r.t. M: $M \to mM, N \to nM, N_n \to n_n M$. We can then introduce two dimensionless quantities, the ligation ratio r_l and the activation ratio r_a, defined as

$$r_l = \frac{k_{lig}^+ M}{k_{hyd}^+ N_{H_2O}} \tag{7.229}$$

$$r_a = \frac{k_{act}^+ N_{XY}}{k_{hyd}^+ N_{H_2O}} \tag{7.230}$$

In this limit, we solve the system of equations Eq. (7.216)-(7.220) and write the solution in terms of dimensionless quantities

$$z = \frac{r_l - r_a + \sqrt{r_a^2(4r_l+1)+2r_a r_l + r_l^2}}{2r_l(r_a+1)} \tag{7.231}$$

$$n_1^A = \frac{r_a(2r_l+1)+r_l-\sqrt{r_a^2(4r_l+1)+2r_a r_l+r_l^2}}{2r_l^2(r_a+1)} \tag{7.232}$$

$$n_1^D = \frac{1-(r_l+1)n_1^A}{(r_a+1)} \tag{7.233}$$

as $z, n_1^A, n_1^D \leq m = 1.0$, all of these quantities take values in the interval $[0,1]$. The length distributions can then be written as

$$n_n^A = n_1^A (r_l\, n_1^A)^{n-1}, \tag{7.234}$$
$$n_n^D = n_1^D (r_l\, n_1^A)^{n-1}, \tag{7.235}$$

which is shown in Fig. 7.17 (see Appendix 10.4.1 for an assesment of the stability of this fixed point).

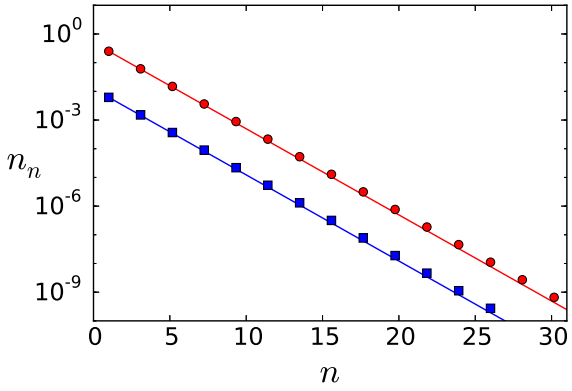


Figure 7.17: Length distribution for activated polymers (blue) and deactivated polymers (red), obtained by simulation of the ODEs. Solid lines correspond to Eqs. (7.234) and (7.235). We chose $r_l = 80$, $r_a = 1$.

Concatenation regimes

We define the degree of concatenation χ as $\chi = M/N = 1/n$. From Eq. (7.231), one can find two limiting regimes for χ:

$$r_l \gg r_a^2, \chi \to r_a + 1 \tag{7.236}$$
$$r_a^2 \gg r_l, \chi \to \sqrt{r_l} \tag{7.237}$$

Where Eq. (7.236) corresponds to an activation-limited regime, and Eq. (7.237) to a ligation-limited regime. By plotting χ versus r_a for various values of r_l, these regimes become apparent: In Fig.7.18,

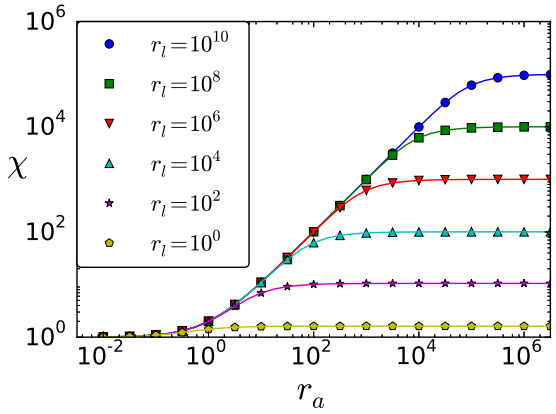


Figure 7.18: Degree of concatenation χ as function of r_l and r_a.

the plateaus on the right correspond to the ligation-limited regime. In this regime, increasing the

k^+_{act} or N_{XY} has no effect on the polymer distribution. The sloped region left to the plateau is the activation-limited regime. Ligation-Fragmentation models are in the former regime and therefore their corresponding scenarios need to verify $r_a^2 \gg r_l$. This means that we can write

$$k^+_{act} N_{XY} >> \sqrt{k^+_{lig} k^+_{hyd} M N_{H_2O}} \tag{7.238}$$

At the crossover between the two regimes, when $r_l = r_a^2$, the activation ratio $\zeta = N_1^A / N_1^D$ approaches 1, as can be seen in Fig.7.19

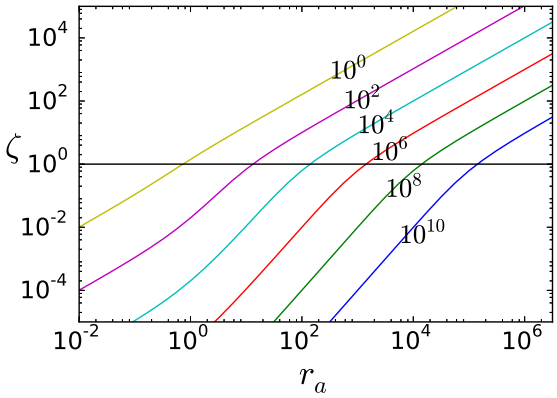


Figure 7.19: Activation ratio ζ, numbers on curves give the value of ligation ratio r_l.

7.5.5 Recovery of ligation-fragmentation model

The ligation-fragmentation model uses the nonelementary reaction

$$[\mathbf{n}+\mathbf{m}] \underset{\boldsymbol{g}_2}{\overset{\boldsymbol{g}_1}{\rightleftharpoons}} \mathbf{n}+\mathbf{m}, \tag{7.239}$$

where $\boldsymbol{g}_1$ denotes a reaction vector composed of a hydrolysis and a reactivation, and $\boldsymbol{g}_2$ a ligation. In early work [21], this point was made clear, and the use of the reversible reaction arrow $\rightleftharpoons$ was chosen as a convenient shorthand due to dynamical similarity. In later work, this notation has remained common, but its subtle interpretation has largely been lost in translation, leading to confusion and thermodynamically inconsistent statements, e.g. by asserting equilibrium polymerization for systems that explicitly require activation.

Since we are interested in sufficiently strong ligation and given that hydrolysis is strongly irreversible, one normally has

$$k^+_{lig} N^A_n N_m \gg k^-_{lig} N_{n+m} N_{\mathrm{YOH}}, \tag{7.240}$$

$$k^+_{hyd} N_{n+m} N_{\mathrm{H_2O}} \gg k^-_{hyd} N^D_n N_m, \tag{7.241}$$

$$k^+_{lig} N^A_n N_m \approx k^+_{hyd} N_{n+m} N_{\mathrm{H_2O}}. \tag{7.242}$$

This means that the elongation and fragmentation are not guided by the same chemical reaction, elongation is due to ligation whereas fragmentation is due to hydrolysis. This distinction is important: if Eq. (7.239) were an elementary reaction, we would have equilibrium polymerization obeying equilibrium thermodynamics. In our system, we have nonequilibrium polymerization driven by chemical currents between reservoirs.

In the absence of such currents, we would have obtained an equilibrium state. This equilibrium state is unique, and the acceleration of one reaction via catalysis cannot alter this equilibrium state.

However, in a non-equilibrium steady state, the system is less constrained (e.g. detailed balance can be broken), and the introduction of a catalyst can alter the steady-state.

Let us consider the three ALF reactions forming a cycle, such as shown in Fig. 7.20. For the kinetics of such a system, backwards reactions can be neglected, and catalysis simply increases a reaction rate. If activation is fast enough to be neglected, we end up with the Ligation-Fragmentation scheme. Contrary to an equilibrium polymerization, it is composed of two separate irreversible reactions, which can be catalyzed separately.

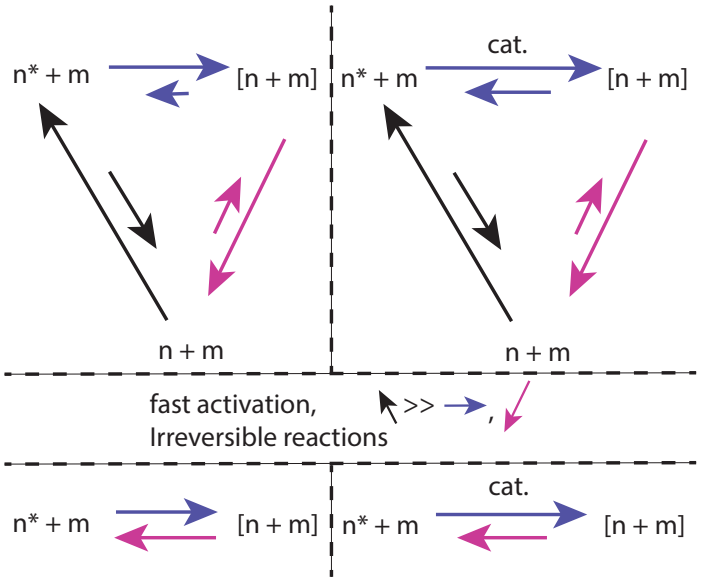


Figure 7.20: Schematic derivation of the ligation-fragmentation model and the effect of catalysis. While catalysis would not alter the net degree of polymerization at equilibrium, the nonequilibrium steady-state admits higher net degree of polymerization.

The main assumption is that activation (Eq. (7.191)) is rapid enough, such that $N^a \to N$, (or $\zeta \gg 1$, see Fig.7.19). This is true in the ligation-limited regime, which implies that Eq.(7.238) must hold. In prebiotic chemistry, accessing such a regime is far from trivial.

7.5.6 Nonequilibrium Thermodynamics

In this section we will derive thermodynamic costs for maintaining a length distribution at steady state, regardless of its exact shape and underlying catalytic schemes.

Dissipation

We can define the total rates of every reaction type in forward (+) and backward (-) direction

$$R^+_{lig} = k^+_{lig} N^A N, \tag{7.243}$$
$$R^-_{lig} = k^-_{lig} (M-N) N_{YOH}, \tag{7.244}$$
$$R^+_{hyd} = k^+_{hyd} (M-N) N_{H_2O}, \tag{7.245}$$
$$R^-_{hyd} = k^-_{hyd} N^D N, \tag{7.246}$$
$$R^+_{act} = k^+_{act} N^D N_{XY}, \tag{7.247}$$
$$R^-_{act} = k^-_{act} N^A N_{HY}. \tag{7.248}$$

The entropy production for performing a single cycle (assuming the chemical reactions are elementary) can be written as [37]

$$\begin{aligned}\Delta S_{tot} &= k\ln\left(\frac{R^+_{lig}R^+_{hyd}R^+_{reg}}{R^-_{lig}R^-_{hyd}R^-_{reg}}\right) \\ &= k\ln\left(\frac{k^+_{lig}k^+_{hyd}k^+_{reg}}{k^-_{lig}k^-_{hyd}k^-_{reg}}\right)+k\ln\left(\frac{N_{XY}N_{H_2O}}{N_{YOH}N_{HY}}\right),\end{aligned} \tag{7.249}$$

If we set $\Delta S_{tot}=0$ (chemical equilibrium), then

$$\ln\left(\frac{k^+_{lig}k^+_{hyd}k^+_{act}}{k^-_{lig}k^-_{hyd}k^-_{act}}\right) = \ln\left(\frac{N^{eq}_{YOH}N^{eq}_{HY}}{N^{eq}_{H_2O}N^{eq}_{XY}}\right) \tag{7.250}$$

Which is the standard free energy change $\Delta G^\circ/kT$ of the net reaction (7.196). We typically expect $k^+ >> k^-$ for all reactions and $\Delta S_{tot} >> 0$. The latter is achieved by having chemostat concentrations that are different from the equilibrium concentration. The free energy change per cycle corresponds to $T\Delta S_{tot}$. At steady state, we need to have

$$R^+_{lig}-R^-_{lig}=R^+_{hyd}-R^-_{hyd}=R^+_{act}-R^-_{act} \tag{7.251}$$

Therefore, one finds the dissipation rate to be

$$\Sigma=(R^+_{hyd}-R^-_{hyd})\Delta S_{tot}\approx R^+_{hyd}\Delta S_{tot}. \tag{7.252}$$

As R^+_{hyd} scales with the number of chemical bonds $M-N$, we need to pay an energetic upkeep cost of $T\Delta S_{tot}$ for every chemical bond, on a timescale $\tau_{hyd}=1/k^+_{hyd}N_{H_2O}$, in order to maintain the steady state.

In general, at steady state, all dissipation can be expressed in terms of currents of chemostatted species [47] by

$$\Sigma=\sum_k I_k\frac{\mu_k}{T}. \tag{7.253}$$

Here, I_k is the current of chemostatted species k between the reservoir and the system and μ_k is its chemical potential. In our process, it is the net reaction (7.196) for which we calculate this dissipation. The influx rate I^* of activating groups is balanced by the outflux of leaving groups, such that we have

$$\Sigma=\frac{I^*}{T}(\mu_{HX}+\mu_{YOH}-\mu_{XY}-\mu_{H_2O}) \tag{7.254}$$

This result is independent of the exact process happening in our system and therefore it remains valid regardless of the length distribution or the catalysis of certain steps. If all bonds are equally prone to hydrolysis, the energy cost for maintenance depends only on the number of polymer bonds, not the exact length distribution.

7.5.7 Exploration and the search for sequences

While one can safely assume a steady state distribution for the length, the sequence space available is so large that many sequences are not present at all times. If a mixture contains ν types of monomers, there are ν^l polymer sequences of length l. Simultaneously, the length distribution decreases exponentially as shown in Eq. (7.209), so that for any given system, we can find a cutoff length l^*. At this cutoff length, the number of possible sequences is larger than the number of polymer species of that length: $\nu^{l^*}>N_{l^*}$.

An essential part of prebiotic polymer scenarios is that the sequence space is being 'explored': the system composition of species longer than l^* changes constantly. This ultimately may lead to functional species being 'found' among explored sequences. Indeed, in a schematic sketch for timeline of increasing 'aliveness', J. Sutherland considers 'Energy-dissipative exploration of macromolecular sequence and composition space' as the process that precedes the first major innovation and bump towards lifelike matter[77]. We will now make one example of a sequence exploration process explicit and consider some of the things it can and cannot do.

Dissipative random search and its efficiency

A cycle of hydrolysis, activation and ligation forms a species that holds the potential of being a sequence that was not encountered before. Let us now suppose that we repeat this process continue until a specified sequence ω is found. Our strategy can then be called a 'dissipative random search', with a cycle of hydrolysis, activation and ligation corresponding to an 'attempt' to sample this sequence.

To sample a sequence ω, the species formed in the ligation reaction must have length $|\omega|$. Many cycles generate a species with $l \neq |\omega|$, however. At steady state, polymers of a given length l are replenished as fast as they are degraded (see Eq. (7.207)). The rate of sampling for length l, is then the corresponding rate of hydrolysis (R_l^{hyd}) of species of length l,

$$J_l^{samp} = R_l^{hyd}. \tag{7.255}$$

If the hydrolysis rate is proportional to the number of hydrolyzable bonds, this leads to

$$J_l^{samp} = k_{hyd}^{+}(l-1)N_l N_{\mathrm{H_2O}}. \tag{7.256}$$

Only a fraction of these reactions occur at $l = |\omega|$. Let us denote this fraction ε_l

$$\varepsilon_l = \frac{J_l^{exp}}{R_{hyd}^{+}} = \frac{(l-1)N_l}{M-N}. \tag{7.257}$$

We can interpret ε_l as an efficiency for this sampling process, it characterizes the fraction of cycles that are sampling the length of the target species.

This efficiency ε_l is a function of the polymer length distribution. Ideally, one would like to have peaked distributions, as observed in a theoretical work on autocatalytic polymer sets [78], such that $\varepsilon_l \to 1$. For an exponential distribution, the optimal ε_l is obtained by having an average length of $l/2$, for which $\varepsilon_l \approx 4/(e^2 l)$ (See appendix 10.4.2).

Trials to to find a sequence

Every time we sample a random sequence of length $l = |\omega|$, we have a probability p_ω to encounter our target species, and with probability $(1-p_\omega)$ we encounter another one. The number trials n_a^l to find ω among species of length $|\omega|$ is then distributed according to

$$\mathscr{P}_{n_a^l}^{\omega} = p_\omega (1-p_\omega)^{n_a^l - 1}, \;\; n_a^l \geq 1 \tag{7.258}$$

for which

$$\langle n_a^l \rangle = \frac{1}{p_\omega}. \tag{7.259}$$

Taking into account the efficiency ε_l, the average total number number of cycles perfomed, n_{cyc}, is then

$$n_{cyc} = \frac{\langle n_a^l \rangle}{\varepsilon_l} \tag{7.260}$$

This calculation can be extended to more general cases, such as a search for any of a particular set of sequences $\Omega = \{\omega_1, .., \omega_s\}$. If these sequences have the same length, we can use

$$p_\Omega = \sum_{k=1}^{s} p_{\omega_k} \tag{7.261}$$

for Eq. (7.258).

If these sequences occur at different lengths, we may consider the general exploration process

$$\mathscr{P}^{\Omega}_{n_a} = P_\Omega (1 - P_\Omega)^{n_a - 1}, \;\; n_a \geq 1 \tag{7.262}$$

where P_Ω is the probability to find the elements in Ω among all possible lengths and sequences, such that

$$P_\omega = p_\omega Y_l \tag{7.263}$$

with $Y_l = N_l/N$ the length fraction.

Let l_Ω be the set containing the lengths of species in Ω. Using Eq. (7.261), we can express the total number of attempts in terms of a parallel search in separate length categories:

$$n_{cyc} = \frac{1}{P_\Omega} = \frac{1}{\sum_{k \in l_\Omega} \frac{\varepsilon_k}{\langle n_a^k \rangle}}. \tag{7.264}$$

Note that the present calculation makes no assumptions on the length distribution or sequence distribution. We do, however, suppose that sampling events can be treated as sufficiently independent.

A tradeoff between power, time and complexity

At steady state, the time t_{cyc} taken to undergo n_{cyc} cycles can be found from the rate of a single cycle, which is equivalent to the total hydrolysis rate R_{hyd}

$$\langle t_{cyc} \rangle = \frac{n_{cyc}}{R_{hyd}}. \tag{7.265}$$

At steady state, the free energy dissipation associated with a single ALF cycle is equal to $T\Delta S_{tot}$, which means the free energy dissipation for n_{cyc} cycles becomes

$$T\Delta S_{cyc} = n_{cyc} T\Delta S_{tot}. \tag{7.266}$$

For a given steady state, we can thus attribute a typical cost for a search process of $T\Delta S_{cyc}$. If we provide n copies of our system to perform the search in parallel, we hydrolyze n times faster, but on average we still require n_{cyc} independent trials. The search time will then verify

$$\langle t_{cyc,n} \rangle = \frac{\langle t_{cyc} \rangle}{n}. \tag{7.267}$$

The total cost $T\Delta S_{cyc}$ has remained unchanged, but more free energy must be expended per unit time: a higher dissipation rate Σ (Eq. (7.253)) must be maintained.

Put together, we can write t_{cyc} in entropic terms

$$\langle t_{cyc} \rangle = \frac{n_{cyc} \Delta S_{tot}}{\Sigma}. \tag{7.268}$$

The average required number of cycles n_{cyc} was shown to be the inverse of the probability P_Ω of randomly 'finding' one of the sequences $\omega_k \in \Omega$. We then find

$$\langle t_{cyc} \rangle = \frac{\Delta S_{tot}}{P_\Omega \Sigma}, \tag{7.269}$$

which relates search time, dissipation rate and how 'hard' an object is to find. An important determinant for how hard it is to find an object of a given length, is the efficiency ε_l. If all sequences in Ω have the same length $l^\bullet$, we can then write

$$\langle t_{cyc} \rangle = \frac{\Delta S_{tot}}{p_\Omega \varepsilon_{l^\bullet} \Sigma}. \tag{7.270}$$

For exponential length distributions that are optimal for searching $l^\bullet$, we have $\varepsilon_{l^\bullet} \approx 4/(e^2 l)$. A null model for the probability p_ω to sample ω among species of length $|\omega|$ would weigh the the contribution of each sequence by its monomer abundances as in Eq. (7.145). If each monomer is equally abundant, this simplifies to the number of sequences

$$p_\omega = \frac{1}{\nu^{l^\bullet}}, \tag{7.271}$$

where ν is the number of distinct monomers in the mixture.

Oftentimes, we are not looking for a particular sequence, but any of a large number of sequences that fills a certain requirement. This requirement may e.g. be that a sequence has a certain secondary structure. Let us denote $\square_l$ the number of possible secondary structures of length l. In a work by Schuster et al[79], $\square_l$ was estimated for RNA through folding algorithms, which yielded a scaling

$$\square_l \approx 1.4848 l^{-3/2} 1.8488^l. \tag{7.272}$$

Note that $\square_l \ll 4^l$, the number of possible sequences drastically outnumbers the number of secondary structures. It was further found that the frequency of these secondary structures is distributed according to a generalized form of Zipf's law:

$$f(x) = a(b+x)^{-c}, \tag{7.273}$$

with x the rank of the secondary structure, and a, b and c fitting parameters.

From Eqs. (7.271),(7.272) and (7.273) it becomes clear that dissipative random search has clear limitations. As l increases, the search for a particular sequence or secondary structure will rapidly lead to escalating costs in terms of Eq. (7.270), even for an optimal efficiency ε_l. Within a finite energy budget, a search may afford to look for specific small objects, or aspecific large objects (i.e. it should accept an overwhelming variety of sequences and secondary structures).

A first application: thermodynamically forbidden scenarios

There are fundamental limitations (Eq. (7.270)) to what we can expect to find through random ligations and fragmentations. We will now, through a deliberately naive calculation, establish an upper bound for how hard dissipative random exploration may search with a planetary energy budget. We stress that this naive bound is only useful to demonstrate an absurdity: if a scenario invokes dissipative random exploration and comes anywhere near this upper bound, then its energy requirements make it unfeasible.

Using todays total influx of solar energy, we find a planetwide power of $P = 1.73\ 10^{17} W$. Let us approximate $T\Delta S_{tot}$ with the $\Delta\mu^\circ$ for the hydrolysis of RNA, 8.4 kT (taking $O(10-100)kT$ one can repeat the argument for an arbitrary copolymer). Applying the full energy influx for dissipative search, we have $T\Sigma = P$. Then, the total influx of solar energy would allow at most for $3.16\ 10^{37}/8.4 = 3.8\ 10^{36}$ trials per second. Per m^2, this amounts to an upper bound of $3\ 10^{23}$ sequences.

Suppose now a scenario which includes dissipative random exploration, to yield a catalytic species of length 200 (a number often evoked in RNA scenarios) with some very particular activity. For simplicity, this activity will be attributed to a secondary structure, to which we ascribe an

optimistic abundance of $p_\Omega = 1/\square_{200}$. Using an exponential length distribution that is tailored to this particular length ($1/\varepsilon_{200} \approx 369$), from Eq. (7.270), we find a value of 3.8 $10^8 y$, slightly longer than the 200 million years provided by the timeline in Ref. [25].

What is considerably harder still, is when a scenario requires more than one object to be formed, especially if these objects need to meet each other. Producing an RNA-based replicase (of thought to be longer than 200 nt long) plus its complementary strand (or another polymerase) requires two prohibitively rare events, occurring in each other's vicinity, in a timeframe that is short enough so that hydrolysis may not have degraded either species yet.

By the generosity of the bound, we can qualify such scenarios as absurd. Note, however, that this absurdity resides in our manner of acquiring complex objects, not in the objects themselves. It highlights the need for a plausible trajectory towards complex species, a problem that inevitably needs to be addressed when we wish to account for modern biopolymers.

Some other search strategies

Through multiple rounds of mutations, recombinations and selection, highly competent objects with a desired function can be acquired (e.g. through in-vitro evolution). Such a strategy can gradually perfect a large sequence, while only sampling a small part of the sequence space. An essential problem for the origins of life is that such a process would seem to already require sophisticated objects like a replicase and error-correction machinery (Eigen's paradox[18]). An interesting question is whether such tasks may be achieved without such machinery as proposed in Refs.[80, 81].

A provocative idea can be distilled from the findings of Ref. [82]: complex molecules today may have started out short functional sequences as short as trimers. In Ref[82] thiol-containing tripeptides react with Fe^{2+} under UV irradiation, to form an iron-sulfur cluster, which subsequently is converted into a stable dodecapeptide complex functioning as a redox catalyst. Such clusters are the defining feature of ferredoxins, which fulfill a key role metabolism. This suggestion still requires use to come up with ways to complexify the sequence and subsequently maintain it, but it sidesteps the challenge of finding a large functional sequence.

Another interesting strategy involves self-assembling functional objects from smaller sequences. This was pioneered with the azoarcus ribozyme, a catalyst for recombination chemistry whose fragments can self-assemble in a noncovalent complex, which subsequently links these fragments through recombination (small oligomers are released in the linking of the fragments)[34, 12]. Such an approach has recently been applied for the assembly of a ribozyme that catalyzes the ligation of triplets [83]. As an origin for complex structures, such an approach requires that these smaller sequences are formed abundantly and locally, so that self-assembly can occur (in spite of e.g. degradation, misassembly, etc.).

These are only some of the 'search strategies' that may be imagined. The strategies may be combined, along with other strategies such as dissipative random exploration. Nevertheless, these considerations have not yet led to a convincing solution for the origins of complex sequences. Although we do not yet understand through what mechanisms this will come to be, it is thought to be part of prebiotic chemical evolution. What this latter term lacks in rigorous definition, it makes up for in anticipation: it has long been expected that there are physical mechanisms that gradually convert a soup of chemicals to a biosphere.

Dissipation-free exploration

The cycles produced by the ALF scheme couple sequence exploration to the expenditure of $T\Delta S_{tot}$. This cost is fundamental to activation chemistry, but not to sequence exploration. As an example, we can consider a recombination reaction

$$[\mathbf{n}+\mathbf{k}]+\mathbf{m} \rightleftarrows [\mathbf{m}+\mathbf{k}]+\mathbf{n}, \tag{7.274}$$

with $1 < k < n$, which was treated in detail in Sec. 7.2.

Again, we will obtain an exponential length distribution. Here, the steady state corresponds to equilibrium and it can be obtained for a closed system. The exact shape of the exponential distribution is unique, and depends on M and N

$$N_l = \frac{N^2}{M}\left(1-\frac{N}{M}\right)^{l-1}. \tag{7.275}$$

For exchange reactions, M and N are conserved quantities. For ALF, only M is conserved, N depends on the particular values of r_l and r_a. If both reactions happen in the same system, we will retain the same steady state, but exploration will be amplified by a cost-free mechanism.

For a given energy budget, ALF reactions cannot explore better than the limits discussed in Sec. 7.5.6, even if new catalysts are formed in the process. Recombination reactions are decoupled from such costs. If exploration would yield a species that catalyzes recombination, the system would explore more sequences with the same energy budget. The potential of recombination for sequence exploration has been explored by other authors [84, 85], but thermodynamic implications have been less explicit.

7.5.8 Decorated ALF models

Our discussion used a rather simple model, in which no additional features like surfaces, folding, templates etc. were introduced. As detailed in Sec. 7.5.5, our system is in a nonequilbrium steady state, and as such modifications of certain rates can modify the exponential length distribution to yield a different length distribution. Since such ingredients are an important part of many scenarios, we will briefly touch on some qualitative aspects of 'decorated ALF models'. For these decorated models, we can no longer use our exact solutions from Sec. 7.5.4, but one can still introduce dimensionless numbers r_l and r_a as measures of the relative strength of ligation and activation processes. Since ligation-fragmentation models in the literature are equivalent to ALF if $r_a \to \infty$, it is instructive to see how these models respond to an activation-limited regime.

Template-assisted ligation

To illustrate the effect of template-assisted ligation, we use an adapted toy-model that captures the main features from other models[49, 19]. Typically, some typical length scale l_1 is chosen for which templated ligation proceeds optimally. We will consider the ALF model, where species with length $l \geq 6$ or larger are considered to be templates, and species with length $l \geq 3$ can engage in template replication. We write the forward rate of this reaction as

$$R^+_{temp} = k^+_T N_T (N^A - N^A_1 - N^A_2)(N - N_1 - N_2) \tag{7.276}$$

where k^+_T is a rate constant and N_T the number of templates as defined above: $N_T = \sum_{l=6}^{\infty} N_l$.

The polymer length distribution in the ligation-limited regime resembles a double-exponential distribution[49, 19]. In this nonequilibrium steady-state we cannot directly employ the results from Sec. 7.2.9. However, if the system is in the activation-limited regime, it can be shown that Eq. (7.236) still holds. For this, consider the ratio of the activation rate and the hydrolysis rate, which can be written as

$$\frac{R^+_{act}}{R^+_{hyd}} = r_a \frac{N^D}{M-N}. \tag{7.277}$$

At steady state, we have $R^+_{act} = R^+_{hyd}$ and in the activation-limited regime we have $N^D \to N$, which leads to

$$r_a \to \chi - 1, \tag{7.278}$$

which implies that if activation becomes too slow with respect to ligation, templates cannot further improve χ. This point is illustrated in Fig. 7.21, where $\chi - 1$ is plotted for a system with and without template-assisted ligation.

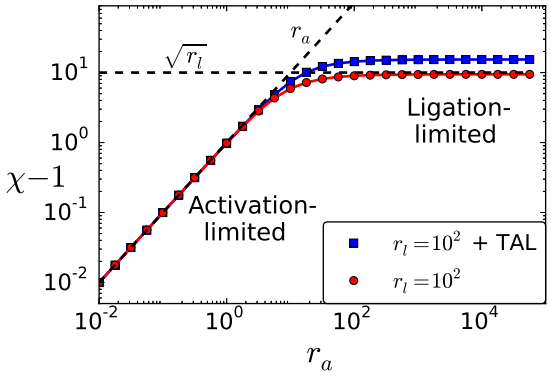


Figure 7.21: $\chi - 1$ plotted against activation ratio, for a system with templated-assisted ligation (TAL, $k_T^+ = 100$), and a reference system with the same parameters but without templates. Dotted lines correspond to asymptotes for the two regimes.

Polymer folding

One can also consider folding of polymers. E.g. when RNA is base-paired in a helical fashion, bonds in the helix are better protected, which can reduce the rate of hydrolysis of this bond with a factor of up to 10 [86, 87]. This effect enriches the steady-state concentration of species that are more folded or hybridized, due to their slower hydrolysis. Some implications of this effect on the sequence space have been considered theoretically for an open prebiotic reactor [51]. Folding is associated with an interesting paradox in replication scenarios: folded species are good candidates for potent catalysts but form poor templates. In Ref. [88] it was suggested that wobble pairs might be a key ingredient to overcoming this obstacle. In the present case, we will only study the effect of folding on χ, through the use of a heuristic toy model.

Let us denote the fraction of double-stranded nucleotides at steady-state with ρ and the hydrolysis rate for dsRNA with $k^+_{hyd,F}$. Note that ρ is not a parameter like e.g. k^+_{hyd}, but an observable that reaches a steady-state value for a given system state. Now, we can write for the hydrolysis rate

$$R^+_{hyd} = (k^+_{hyd}(1-\rho) + k^+_{hyd,F}\rho)(M-N), \tag{7.279}$$

which can be recast in the form of the unmodified overall hydrolysis rate (Eq. (7.245)), if we define a new effective hydrolysis rate constant $\bar{k}^+_{hyd}$, such that

$$\bar{k}^+_{hyd} = k^+_{hyd}(1-\rho) + \bar{k}^+_{hyd,F}\rho. \tag{7.280}$$

Using this effective rate constant, we can write an effective ligation ratio and an effective activation ratio

$$\bar{r}_l = \frac{k^+_{lig}M}{\bar{k}^+_{hyd}N_{H_2O}} = r_l\frac{k^+_{hyd}}{\bar{k}^+_{hyd}}, \tag{7.281}$$

$$\bar{r}_a = \frac{k^+_{act}N_{XY}}{\bar{k}^+_{hyd}N_{H_2O}} = r_a\frac{k^+_{hyd}}{\bar{k}^+_{hyd}}. \tag{7.282}$$

Since the overall rates of activation and hydrolysis have the same form as before, Eq. (7.278) can be derived in the same fashion as in Sec. 7.5.8 so that in the activation-limited regime we can write

$$\chi - 1 \to \bar{r}_a. \tag{7.283}$$

Similarly, we can write the ratio of overall ligation and hydrolysis rates

$$\frac{R^+_{lig}}{R^+_{hyd}} = \frac{k^+_{lig} N^A N}{\bar{k}^+_{hyd} N_{H_2O}(M-N)}. \tag{7.284}$$

At steady-state, $R^+_{lig} = R^+_{hyd}$ and in the ligation-limited regime $N^A \to N$, which yields

$$\bar{r}_{lig} = \frac{M(M-N)}{N^2} = \chi^2 - \chi \tag{7.285}$$

which for large χ becomes equivalent to Eq. (7.237), as

$$\chi \to \sqrt{\bar{r}_{lig}} \tag{7.286}$$

Since folding is assumed to slow down hydrolysis, we have that $\bar{k}^+_{hyd} \leq k^+_{hyd}$, so folding yields a relative gain for the effective ligation ratio and activation ratio. As shown by Eqs. (7.281) and (7.282), both are increased by the same factor $k^+_{hyd}/\bar{k}^+_{hyd}$.

The scaling of χ is quite different however. In the activation-limited regime (Eq. (7.283)), the increase in χ with respect to a nonfolding scenario is linear in $k^+_{hyd}/\bar{k}^+_{hyd}$. In the ligation-limited regime (Eq. (7.286)), the asymptotic scaling of χ is proportional to $\sqrt{k^+_{hyd}/\bar{k}^+_{hyd}}$. Consequently, for a given value of ρ, the gain in χ due to folding is larger when activation is slow.

The enrichment of folded species can also happen through energetic means, if the system is allowed to explore sequences in a reversible fashion. This can e.g. be achieved through exchange reactions as discussed in Sec. 7.5.7. If such exchange reactions are rapid with respect to ALF, we approach a thermodynamic equilibrium distribution for folding. Then, folded species are enriched according to the Boltzmann weight of their folding energy: $\exp\left(-\Delta G_{fold}/kT\right)$ This increase in dsRNA fraction ρ then decreases $\bar{k}^+_{hyd}$, thus increasing the ratio $k^+_{hyd}/\bar{k}^+_{hyd}$. Thus, exchange reactions allow to increase the degree of concatenation χ, irrespective of the concatenation regime. This conclusion is reminiscent of Ref. [84], where it was proposed that 'distributions of lengths can shift upwards through recombination'.

7.6 Some general aspects of length distributions

Some polymer length distributions are encountered particularly often in the OOL literature: the exponential and (approximate) multi-exponential distribution. In this section, we will provide some basic arguments that describe what kind of length distributions we should expect for some common kinetic equations in polymer models.

Single Exponentials

As discussed in Sec. 7.2.9, an exponential distribution maximizes the entropy associated with a length distribution. It Is the most random distribution of a finite number of elements, so one should expect that a system naturally tends to move towards such a distribution. Another argument for exponentials comes when we look at the evolution equations. E.g. for Chain exchange:

$$\frac{dN_l}{dt} = k \sum_{l_A+l_B=l} \sum_{l_C,l_D}^{\infty} [N_{l_A+l_D} N_{l_C+l_B} - N_{l_A+l_B} N_{l_C+l_D}]. \tag{7.287}$$

We see directly that the total length for the pairs involved is fixed. Therefore, a stationary solution of the form

$$N_l = A \exp(-al), \tag{7.288}$$

with $a > 0$, is a very natural choice. Coupled with the conserved quantities

$$\sum N_l = N, \tag{7.289}$$
$$\sum lN_l = M. \tag{7.290}$$

This then provides a complete equilibrium solution. The same argument applies for Attack-exchange, for which the kinetic equation is

$$\frac{dN_l}{dt} = k \sum_{l_A,l_B=1}^{\infty} [N_{l_A} N_{l+l_B} - N_{l_A+l_B} N_l] + k \sum_{l_A}^{l-1} \sum_{l_B=1}^{\infty} [N_{l_A} N_{l_B+l-l_A} - N_l N_{l_B}]. \tag{7.291}$$

With the exception that $l \geq 2$. For systems performing ligation-fragmentation, the most basic description is an equation of the form

$$\frac{dN_l}{dt} = \sum_{k=1}^{l-1} k^+ N_k N_{l-k} - k^- (l-1) N_l - \sum_{k=1}^{\infty} [k^+ N_l N_k - k^- N_{l+k}] \tag{7.292}$$

Which is still solved by the exponential ansatz, but the we no longer have a conservation law for N, only for M. The stationary solution of Eq. (7.292) provides an extra constraint:

$$k^+ A^2 = k^- A \tag{7.293}$$

Which describe the thermodynamic contribution of binding vs fragmentation. An elegant approach is to consider an equilibrium constant to form an n-mer

$$K = \frac{N_n}{N_{n-1} N_1} = \frac{k^+}{k^-} \tag{7.294}$$

From which an equilibrium solution can be found of the form

$$N_1 = \frac{(2KM-1) - \sqrt{(2KM-1)^2 - 4K^2M^2}}{2KB^2M}, \tag{7.295}$$
$$N_n = N_1 K^{n-1}. \tag{7.296}$$

The degree of polymerization M/N that follows is a compromise between the configurational freedom of having more species and the free energy contribution of binding them together. Since our null model has a fixed binding energy (no length dependence), these bonds are expected to be distributed as arbitrary as possible, which would yield an exponential.

In the ALF model, we must be more careful with the aforementioned arguments, as we are not at equilibrium. However, for rapid activation, the equations do simplify to those of an equilibrium polymerization, and an exponential steady-state solution is applicable. However, the object 'K' can no longer be interpreted as an equilibrium constant. In the theoretical treatment of ligation-fragmentation in a thermal trap[20], the fast-reaction approximation implies that at any given height the system has an exponential length distribution. As mentioned before, this result is derived for dilute systems and leads to unphysical behavior when this condition is no longer verified. In Ref. [20] a form of mass divergence was found in the dilute limit, referred to as an escalation, due to the indefinite accumulation of external material. While this description will then no longer be able to provide a correct description of the length distribution, its divergence provides a clear signature of why that is.

Multiple exponentials

In the search for longer polymers, one typically aims to do better than the single exponential. Many models do this by considering a length-dependent mechanism that improves a production rate [19, 49] or reduces a degradation rate[51] starting at a certain length, such as templates or folding. What happens in many of these ligation models then, boils down to the introduction of a second timescale, which does not act below a threshold length $l*$. An instructive set of equations is the following, in which degradation is occurs at a rate k_1^- for $l < l*$:

$$\begin{aligned}\frac{dN_l}{dt} &= k^+\sum_{k=1}^{l-1} N_k N_{l-k} - k_1^-(l-1)N_l - \sum_{k=1}^{l*-l}[k^+N_lN_k - k_1^-N_{l+k}] - \sum_{k=l*-l+1}^{\infty}[k^+N_lN_k - k_2^-N_{l+k}] \\ &= k^+\sum_{k=1}^{l-1} N_k N_{l-k} - k_1^-(l-1)N_l - \sum_{k=1}^{l*-l}[k^+N_lN_k - k_1^-N_{l+k}] - \sum_{k=l*-l+1}^{\infty}[k^+N_lN_k - k_1^-N_{l+k}] \\ &+ (k_2^- - k_1^-)(N-N*) \quad l < l* \end{aligned} \tag{7.297}$$

and is slowed down to a rate k_2^- (e.g. due to folding) for $l \geq l*$

$$\frac{dN_l}{dt} = k^+\sum_{k=1}^{l-1} N_k N_{l-k} - k_2^-(l-1)N_l - \sum_{k=1}^{\infty}[k^+N_lN_k - k_2^-N_{l+k}] \quad l \geq l* \tag{7.298}$$

For $l < l*$, the equation can be written as the ligation-fragmentation model with a correction term. For $l \gg l*$, we have the ligation fragmentation model, but with contributions up to N_l* modified by the first set of equations, and with increasing l, the relative contribution of terms up to N_l* diminishes. Qualitatively, the length distribution resembles a double exponential, especially asymptotically. Many length distributions in the literature (e.g. in Ref. [19, 49]] can be readily understood as approximate double exponentials.

Power laws

Power laws are more exotic, but the equations that yield them are not. An instructive example is a system with a dominant, length-independent degradation process, such as happens in a CSTR or if a (highly mobile) degradative agent attacks polymers at their endpoints.[‡] We can then consider an equation of the form

$$\frac{dN_l}{dt} = k^+\sum_{k=1}^{l-1} N_k N_{l-k} - k_\emptyset N_l - \sum_{k=1}^{\infty} k^+N_lN_k = k^+\sum_{k=1}^{l-1} N_k N_{l-k} - k'_\emptyset N_l \tag{7.299}$$

Where $k'_\emptyset = Nk^+ + k_\emptyset$. Let us furthermore chemostat a particular species, e.g. N_1. Compared to ligation-fragmentation, longer species are more favored, as the increasing prefactor $(l-1)$ is now absent. We should therefore expect the length distribution to fall of slower than an exponential. Indeed, ansatz Eq. (7.288) is no longer valid. However, a modified ansatz

$$N_l = Af(l)\exp(-al), \tag{7.300}$$

can be used, which can be solved by obeying the constraints

$$k^+A^2 = k'_\emptyset A, \tag{7.301}$$

$$\sum_{k=1}^{l-1} f(k)f(l-k) = f(l). \tag{7.302}$$

[‡]I am indebted to Joachim H. Rosenberger and Tobias Göppel, who showed me their experimental results and simulations, which prompted me to brush up this calculation.

Eq. (7.302) can be related to the Catalan numbers, defined by

$$C_0 = 1, \quad C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}. \tag{7.303}$$

From which we find that $f(l) = C_{l-1}$. The asymptotic growth of the Catalan numbers goes as

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}} \tag{7.304}$$

Which means that asymptotically, N_l behaves as a power law

$$N_{l+1} \sim \frac{B}{l^{3/2}} \exp(-bl). \tag{7.305}$$

In deriving this expression, we supposed that degradation (e.g. by leaving the reactor) is faster than e.g. the reverse of ligation or the forward fragmentation. Such a situation of irreversible polymerization can occur in various instances, but we should bear in mind that the degradation does not inherently 'help' to get longer polymers. In the regime where fragmentation can be neglected, ligation fragmentation without extra degradation would have led to even longer polymers.

7.7 Polymer scenarios and the search for long polymers

We have discussed some new ways in which longer polymers can be assembled, through coupling with reservoirs in new ways. This is but a small exposé of the options that have been proposed.

New approaches are showing that a large array of nonequilibrium situations can be imagined to polymerize species that inherently have trouble doing so. While we have no proof that any of these mechanisms were actively exploited in prebiotic chemistry, their steady increase in number makes it increasingly conceivable that prebiotic polymers that do not spontaneously polymerize, may have been provided early on.

Narratives for what prelife did with those polymers, however, either remain vague, or explicitly use mechanisms that are demonstrably problematic from the point of view of thermodynamics, chemistry and evolution. In this section, some of these concerns are discussed. We propose that we should consider the early role of prebiotic polymers in terms of more simple, general behavior of polymers, of which the coacervates afforded by polyelectrolyte condensation is taken as an example. Like in modern evolution, we may need to consider a gradual progression, in which intermediate features may have provided a scaffold that became redundant and was later discarded.

Fundamental concerns: implausible ingredients in polymer scenarios

Provided that some mechanism existed that made prebiotic polymers early on, we may then ask: what happens then? As shown in Sec. 7.5.7, a dissipative random search to find a specific long sequence is prohibitively expensive, so we may at best hope for something aspecific. The picture of 'sequence exploration finding complex components' remains invariably tempting, but at present no solid mechanistic picture has been advanced to plausibly make this concrete.

We should equally well be careful with the view that chemical networks of polymer sequences autocatalytically self-assemble, replicate and evolve as soon as diversity is large enough. As outlined in Sec. 6.1.1, side reactions, degradation, trapping stapes, transport limitations and other undesirable side-effects grow rapidly with increasing local diversity. The typical growth in autocatalytic pathways is expected to be relatively meager compared to these contributions. In this context, autocatalytic network motifs alone are not a guarantee for feasibility, since a growing cumulative rate of side-reactions and degradation[89, 90] already leads to a precipitous drop in the probability to finish any autocatalytic cycle. Experimentally, autocatalytic networks of polymers

are made with an extremely narrow subset of long sequences. The prebiotic establishment of such a specific food set is a fundamental problem in and off itself.

Our concerns with polymer scenarios seem considerable. We wish to formulate scenarios that avoid too rare rare events or too high thermodynamic cost and that are in line with the chemical principles that govern autocatalysis.

An example of a plausible ingredient: coacervates

These concerns only pertain to specific (polymer) mechanisms, not polymers themselves. The prebiotic relevance of polymeric species should not hinge on the validity of a subset of ideas, but on what polymers can bring to the table in general. A lot of interesting polymer behavior does not require strict control over copolymer sequence.

As an example, charged polymers (such as RNA and peptides with ionizable side groups) can, upon achieving a sufficient size and concentration, undergo a polyelectrolyte condensation, forming liquid microcompartments called coacervates. The local environment in the droplet is considerably different from a bulk water phase, as testified e.g. by partition coefficients and local catalytic activity [91]. More and more of these droplets are now being identified in biological systems (e.g. Cajal bodies[92], P-granules[93], Stress granules [94]), where they have been shown to perform various various roles, such as localizing enzymatic reactions, amplifying signals, nucleating the growth of acting filaments and storage of RNA and proteins[95]. In prebiotic scenarios, their interest was evoked by Oparin[3] in the early 20th century. They are now garnering renewed interest, spurred by rapid developments in biology and artificial coacervates. These coacervates are not limited to the single polymer droplet case: in complex coacervates alternative morphologies are elaborated, such as aggregates of micelles held together by a charged polymer [96].

Polyelectrolyte condensation is a phase change induced (roughly speaking) when a combination of length, abundance and charge (due to polymers and small ions) overcome a critical threshold. Complementary sequences may facilitate assembly and shift this threshold, but in absence of sequence control coacervates remain accessible species. They play a role in various prebiotic scenarios [3, 97] and protocell models [91]. In our general framework for stoichiometric autocatalysis (Sec. Ch.5), forming chemically distinct environments (such as coacervates) provides an elegant pathway towards further autocatalytic evolution in a spatial setting (Sec. Ch.6). In this manner, a more gradual chemical evolution can be formulated.

Coacervates are an example of the interesting features we may readily introduce in a prebiotic polymer scenario. It is such features that justify our continued search for prebiotic polymerization. In conceptualizing how biological sophistication arose (e.g. enzymes, genetic machinery, organelles, cells), we may need a more detailed picture of plausible (and arguably, simpler) features that were in place before. Features like coacervates may not only have preceded such sophistication, but quite likely have enabled their formation.

Another interesting perspective comes from considering nonenzymatic, template-directed ligation [98]. Leu et al showed that a template of either RNA,DNA or LNA, supplied with RNA or DNA primers and activated nucleotides exhibit replication that strongly favors correct copying: after incorrect incorporation further extension is slowed by a factor of 10-100, attributed to a lack of cooperativity due to mismatch. Incorrect incorporation also favored more subsequent errors, leading to more extensive sequence exploration. Such systems invite us to rethink the evolutionary mechanisms that may act on copolymers, and reconsider what such systems really need. One interesting addition might be the non-enzymatic backbone proofreading system by Mariani and Sutherland[81].

7.7.1 Scaffolds and gradualism

For many, the key argument for an early (pre-)RNA world is the functional repertoire of catalytic polymers and the prospect of digital evolution on heritable sequences, akin to modern evolution.

The transmission of a heritable state, however, does not imply polymer sequences. By definition, prebiotic chemistry does not do things the way life does it, and we strive to explain how they gradually became identical.

A powerful concept to address this question is the notion of a scaffold, a supportive structure necessary in the initial construction of e.g. a building, that is removed afterward. Cairns-Smith used this analogy to argue how mineral surfaces could have had major roles in prebiotic chemistry, that were readily abandoned in more advanced stages of prelife[56]. It is also a useful analogy in explaining how evolution may produce large functional organs by small cumulative changes. For such a process, we don't want to take one large implausible leap, but rather small, gradual steps that yield intermediates that are competent enough to preserve themselves (e.g. by conferring a simple, but significant function).

Features for which polymers are invoked in prebiotic scenarios, are often achievable by small molecules. Compartments can be made with relatively short surfactant (e.g. lipids, terpenes). Catalysis can be performed by the smallest of ions. Proofreading is a network feature equally accessible to small molecules. Species as small as glycolaldehyde have self-replication pathways. Single amino acids can self-assemble with counterions, e.g. to form long fibrils [99].

Even prebiotic chemical evolution can be formulated plausibly and coherently in terms of multicompartment autocatalytic evolution (Ch.6). Certain modes of (multicompartment) autocatalysis may be more readily performed by small molecules (Sec. 5.4). A scenario that requires such functions from polymers can therefore rely on a rich collective of plausible, simple scaffolds, that were eventually replaced.

An elegant recent example of a small gradualistic proposal came from S.S. Mansy's group. By having thiol-containing tripeptides react with Fe^{2+} under UV irradiation, an iron-sulfur cluster was formed, complexed by four tripeptides. These tripeptides can subsequently ligate to form a more stable dodecapeptide complex that functions as a redox catalyst. Today, such clusters are the defining feature of enzymes referred to as ferredoxins, which play a key role metabolism. In Ref. [82] it was proposed that such clusters may have formed early on to form the basis of an iron-sulfur metabolism. The small length (trimers) of the peptides and the limited sequence requirements (at least one cysteine), make it an attractive prebiotic species that is functional from the start. This 'protoferredoxin' requires neither sequence control nor long polymers, whereas modern ferredoxins do. This then begs the question: can we start with a set of self-assembling functional species and gradually have genetic polymers take over their production and evolution? And if so, through what kind of stages can something like that plausibly come about?

Whether we owe our ferredoxins to such a chemistry or not should not distract us from the more general point: impressive functional species may be readily accessible without being long or having a particular fixed sequence. Their sequence need not be stored in genes and translated; the species may assemble themselves (in the example: aided by UV radiation). Such mechanisms are part of the daily routine in the young fields of supramolecular and dynamic combinatorial chemistry[100]. In dynamic combinatorial chemistry, a large diversity of rapidly interconverting species is generated. These species can subsequently be screened for a desired activity, e.g. by adding a preferred substrate to bind to. Such a preferred binding leads to a strong and rapid accumulation of the favorable complexed species, allowing for quick identification of small species with exceptional activity. Such a simple mechanism may temporarily sidestep the thermodynamic challenges of synthesizing long polymers with specific sequences, while still supplying exceptionally well-adapted species and exploring new ones.

Bibliography

Articles

[1] Walter Gilbert. “The RNA world”. In: *Nature* 319 (1986), p. 618.

[2] Paul G Higgs and Niles Lehman. “The RNA World: molecular cooperation at the origins of life”. In: *Nat. Rev. Genet.* 16 (2015), pp. 7–17.

[4] Stuart A. Kauffman. “Autocatalytic Sets of Proteins”. In: *J. Theor. Biol.* 119 (1986), pp. 1–24.

[5] Jayanta Nanda et al. “Emergence of native peptide sequences in prebiotic replication networks”. In: *Nat. Comm.* 8.434 (2017), pp. 1–8.

[6] Peter E Nielsen. “Peptide Nucleic Acids and the Origin of Life”. In: *Chem. Biodivers.* 4 (2007), pp. 1996–2002.

[7] Kevin E Nelson, Matthew Levy, and Stanley L Miller. “Peptide nucleic acids rather than RNA may have been the first genetic molecule”. In: *Proc. Natl. Acad. Sci. U.S.A.* 97.8 (2000), pp. 3868–3871.

[8] Peter Egil Nielsen. “Peptide nucleic acid (pna): a model structure for the primordial genetic material?” In: *Orig. Life Evol. Biosph.* 23 (1993), pp. 323–327.

[9] Vitor B Pinheiro et al. “Synthetic Genetic Polymers Capable of Heredity and Evolution”. In: *Science (80-.).* 336 (2012), pp. 341–344.

[10] Ramanarayanan Krishnamurthy. “Life’s Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry?” In: *Chem. Eur. J.* 24 (2018), p. 16708.

[11] Daniel Segré et al. “The lipid world”. In: *Orig. Life Evol. Biosph.* 31 (2001), pp. 119–145.

[12] Nilesh Vaidya et al. “Spontaneous network formation among cooperative RNA replicators”. In: *Nature* 491.7422 (Oct. 2012), pp. 72–77.

[13] Marcel Hollenstein. “DNA Catalysis : The Chemical Repertoire of DNAzymes”. In: *Molecules* 20 (2015), pp. 20777–20804.

[14] Xi Chen, Na Li, and Andrew D Ellington. “Ribozyme Catalysis of Metabolism in the RNA World”. In: *Chem. Biodivers.* 4 (2007), pp. 633–655.

[15] Alexander I Taylor et al. “Catalysts from synthetic genetic polymers”. In: *Nature* 518.7539 (2015), pp. 427–430.

[17] A Blokhuis et al. “Selection Dynamics in Transient Compartmentalization”. In: *Phys. Rev. Lett.* 120.15 (Apr. 2018), p. 158101.

[18] M Eigen. “Self-organization of matter and the evolution of biological macromolecules”. In: *Naturwissenschaften* 58.10 (1971), pp. 465–523.

[19] Julien Derr et al. “Prebiotically plausible mechanisms increase compositional diversity of nucleic acid sequences”. In: *Nucleic Acids Res.* 40.10 (2012), pp. 4711–4722.

[20] Christof B Mast et al. “Escalation of polymerization in a thermal gradient”. In: *Proc. Natl. Acad. Sci. U.S.A.* 110.20 (2013), pp. 8030–5.

[21] Meng Wu and Paul G Higgs. “Origin of Self-Replicating Biopolymers: Autocatalytic Feedback Can Jump-Start the RNA World”. In: *J. Mol. Evol.* 69 (2009), pp. 541–554.

[22] Meng Wu and Paul G Higgs. “Comparison of the Roles of Nucleotide Synthesis, Polymerization, and Recombination in the Origin of Autocatalytic Sets of RNAs”. In: *Astrobiology* 11.9 (2011), pp. 895–906.

[23] Wim Hordijk, Jotun Hein, and Mike Steel. "Autocatalytic Sets and the Origin of Life". In: *Entropy* 12 (2010), pp. 1733–1742.

[24] Wim Hordijk, Stuart A Kauffman, and Mike Steel. "Required Levels of Catalysis for Emergence of Autocatalytic Sets in Models of Chemical Reaction Systems". In: *Int. J. Mol. Sci* 12 (2011), pp. 3085–3101.

[25] Gerald F Joyce. "The antiquity of RNA-based evolution". In: *Nature* 418 (2002), pp. 214–221.

[26] Alex Blokhuis and David Lacoste. "Length and sequence relaxation of copolymers under recombination reactions Length and sequence relaxation of copolymers under recombination reactions". In: *J. Chem. Phys.* 147 (2017), p. 094905.

[27] Ryo Mizuuchi et al. "Mineral surfaces select for longer RNA molecules". In: *Chem. Commun.* 55 (2019), pp. 2090–2093.

[28] Amir H Hoveyda and Adil R Zhugralin. "The remarkable metal-catalysed olefin metathesis reaction". In: *Nature* 450.7167 (2007), pp. 243–251.

[29] Damien Montarnal et al. "Silica-Like Malleable Materials from Permanent Organic Networks". In: *Science (80-.).* 334.6058 (2011), pp. 965–968.

[30] Mathieu Capelot et al. "Metal-Catalyzed Transesterification for Healing and Assembling of Thermosets". In: *J. Am. Chem. Soc.* 134.18 (2012), pp. 7664–7667.

[31] Yeon Lee and Donald C Rio. "Mechanism and Regulation of Alternative Pre-mRNA Splicing". In: *Annu. Rev. Biochem.* 84.1 (June 2015), pp. 291–323.

[32] Benedict A Smail et al. "Spontaneous advent of genetic diversity in RNA populations through multiple recombination mechanisms". In: *RNA* 25.4 (2019), pp. 453–464.

[33] Hannes Mutschler et al. "Random-sequence genetic oligomer pools display an innate potential for ligation and recombination". In: *Elife* 7 (2018), p. 43022.

[34] Will E Draper, Eric J Hayden, and Niles Lehman. "Mechanisms of covalent self-assembly of the Azoarcus ribozyme from four fragment oligonucleotides". In: *Nucleic Acids Res.* 36.2 (2008), pp. 520–531.

[35] Niles Lehman. "A Case for the Extreme Antiquity of Recombination". In: *J. Mol. Evol.* 56.6 (June 2003), pp. 770–777.

[36] S Lahiri et al. "Kinetics and Thermodynamics of reversible Polymerization in closed systems". In: *New J. Phys.* 17 (2015), p. 85008.

[38] D Andrieux and P Gaspard. "Nonequilibrium generation of information in copolymerization processes". In: *Proc. Natl. Acad. Sci. U.S.A.* 105.28 (2008), pp. 9516–9521.

[39] Yannick De Decker. "On the stochastic thermodynamics of reactive systems". In: *Physica A* 428 (2015), pp. 178–193.

[40] M Esposito and C Van den Broeck. "Ensemble and trajectory Thermodynamics: A brief introduction". In: *Physica A* 418 (2015), p. 6.

[41] Udo Seifert. "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports Prog. Phys.* 75.12 (2012), p. 126001.

[42] P Gaspard. "Fluctuation theorem for nonequilibrium reactions". In: *J. Chem. Phys.* 120.19 (2004), pp. 8898–8905.

[44] T Schmiedl, T Speck, and U Seifert. "Entropy Production for Mechanically or Chemically Driven Biomolecules". In: *J. Stat. Phys.* 128 (2007), p. 77.

[45] P J Flory. “Thermodynamics of Heterogeneous Polymers and Their Solutions”. In: *J. Chem. Phys.* 12 (1944), p. 425.

[46] P J Blatz and A V Tobolsky. “Note on the kinetics of systems manifesting simultaneous polymerization-depolymerization phenomena”. In: *J. Chem. Phys.* 49 (1945), p. 77.

[47] Matteo Polettini and Massimiliano Esposito. “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”. In: *J. Chem. Phys.* 141.2 (2014), pp. –.

[48] Riccardo Rao and Massimiliano Esposito. “Conservation laws and work fluctuation relations in chemical reaction networks”. In: *J. Chem. Phys.* 149.24 (2018).

[49] Alexei V Tkachenko and Sergei Maslov. “Spontaneous emergence of autocatalytic information-coding polymers”. In: *J. Chem. Phys.* 143.2015 (2015), p. 045102.

[50] Shinpei Tanaka, Harold Fellermann, and Steen Rasmussen. “Structure and selection in an autocatalytic binary polymer model”. In: *EPL (Europhysics Lett.* 107.2 (2014), p. 28004.

[51] Benedikt Obermayer et al. “Emergence of Information Transmission in a Prebiotic RNA Reactor”. In: *Phys. Rev. Lett.* 107 (2011), p. 018101.

[52] R. Rao, D. Lacoste, and M. Esposito. “Glucans monomer-exchange dynamics as an open chemical network”. In: *J. Chem. Phys.* 143 (2015), p. 244903.

[53] D T Gillespie. “Exact Stochastic Simulation of Coupled Chemical-reactions”. In: *J. Phys. Chem* 81 (1977), p. 2340.

[54] R. Rao and L. Peliti. “Thermodynamics of accuracy in kinetic proofreading: dissipation and efficiency trade-offs”. In: *J. Stat. Mech.* 2015.6 (2015), p. 06001.

[55] Paul J Flory. “Thermodynamics of High Polymer Solutions”. In: *J. Chem. Phys.* 660 (1941), pp. 8–10.

[57] Günter Wächtershäuser. “Before Enzymes and Templates: Theory of Surface Metabolism”. In: *Microbiol. Rev.* 52.4 (1988), pp. 452–484.

[58] Helen Greenwood Hansma. “Possible origin of life between mica sheets: Does life imitate mica?” In: *J. Biomol. Struct. Dyn.* 31.8 (2013), pp. 888–895.

[60] Elisa Biondi et al. “Montmorillonite protection of an UV-irradiated hairpin ribozyme : evolution of the RNA world in a mineral environment”. In: *BMC Evol. Biol.* 7 (2007), S2.

[61] Elisa Biondi et al. “Adsorption of RNA on mineral surfaces and mineral precipitates”. In: *Beilstein J. Org. Chem.* 13.i (2017), pp. 393–404.

[62] A Ricardo et al. “Borate Minerals Stabilize Ribose”. In: *Science (80-.).* 303.January (2004), p. 2004.

[63] Giovanna Costanzo et al. “Nucleoside Phosphorylation by Phosphate Minerals”. In: *J. Biol. Chem.* 282.23 (2007), pp. 16729–16735.

[64] James P. Ferris et al. “Synthesis of long prebiotic oligomers on mineral surfaces”. In: *Nature* 381 (1996), pp. 59–61.

[65] James P Ferris. “Montmorillonite-catalysed formation of RNA oligomers : the possible role of catalysis in the origins of life”. In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361 (2006), pp. 1777–1786.

[66] Giorgio Bernardi. “Chromatography of nucleic acids on hydroxyapatite”. In: *Nature* 206 (1965), pp. 779–783.

[67] D. Gibbs, R. Lohrmann, and L. E. Orgel. "Template-directed synthesis and selective adsorption of oligoadenylates on hydroxyapatite". In: *J. Mol. Evol.* 1 (1980), pp. 347–354.

[68] H James Cleaves et al. "The adsorption of short single-stranded DNA oligomers to mineral surfaces". In: *Chemosphere* 83.11 (2011), pp. 1560–1567.

[69] A.J. Ramirez-Pastor et al. "Statistical thermodynamics and transport of linear adsorbates". In: *Phys. Rev. B* 59.16 (1999), pp. 27–36.

[70] F. Romá, J.L. Riccardo, and A.J. Ramirez-Pastor. "Application of the Fractional Statistical Theory of Adsorption (FSTA) to Adsorption of Linear and Flexible k -mers on Two-Dimensional Surfaces". In: *Ind. Eng. Chem. Res.* 45.6 (2006), pp. 2046–2053.

[72] Matthew W Powner, Béatrice Gerland, and John D Sutherland. "Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions". In: *Nature* 459 (2009), pp. 239–242.

[73] Clémentine Gibard et al. "Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions". In: *Nat. Chem.* 10.November (2018), pp. 212–217.

[74] Albert C Fahrenbach et al. "Common and Potentially Prebiotic Origin for Precursors of Nucleotide Synthesis and Activation". In: *J. Am. Chem. Soc.* 139 (2017), pp. 8780–8783.

[75] Taifeng Wu and Leslie E Orgel. "Nonenzymatic Template-Directed Synthesis on Hairpin Oligonucleotides. 3. Incorporation of Adenosine and Uridine Residues". In: *J. Am. Chem. Soc.* 114.21 (1992), pp. 7963–7969.

[76] Riccardo Rao and Massimiliano Esposito. "Nonequilibrium Thermodynamics of Chemical Reaction Networks : Wisdom from Stochastic Thermodynamics". In: *Phys. Rev. X* 6 (2016), p. 041064.

[77] John D Sutherland. "Studies on the origin of life — the end of the beginning". In: *Nat. Rev. Chem.* 1 (2017), p. 0012.

[78] Varun Giri and Sanjay Jain. "The Origin of Large Molecules in Primordial Autocatalytic Reaction Networks". In: *PLoS One* 7.1 (2012), e29546.

[79] Peter Schuster et al. "From Sequences to Shapes and Back : A Case Study in RNA Secondary Structures". In: *Proc. R. Soc. B* 255 (1994), pp. 279–284.

[80] Kevin Leu et al. "Cascade of Reduced Speed and Accuracy after Errors in Enzyme-Free Copying of Nucleic Acid Sequences". In: *J. Am. Chem. Soc.* 135 (2013), pp. 354–366.

[81] Angelica Mariani and John D. Sutherland. "Non-Enzymatic RNA Backbone Proofreading through Energy-Dissipative Recycling". In: *Angew. Chemie - Int. Ed.* 56.23 (2017), pp. 6563–6566.

[82] Claudia Bonfio et al. "UV-light-driven prebiotic synthesis of iron–sulfur clusters". In: *Nat. Chem.* (2017).

[83] James Attwater et al. "Ribozyme-catalysed RNA synthesis using triplet building blocks". In: *Elife* 7 (2018), p. 35255.

[84] N Lehman. "A Recombination-Based Model for the Origin and Early Evolution of Genetic Information". In: *Chem. Biodivers.* 5.9 (2008), pp. 1707–1717.

[86] D A Usher and A H Mchale. "Hydrolytic stability of helical RNA : A selective advantage for the natural 3',5'-bond". In: *Proc. Natl. Acad. Sci. U.S.A.* 73.4 (1976), pp. 1149–1153.

[87] Garrett A Soukup and Ronald R Breaker. "Relationship between internucleotide linkage geometry and the stability of RNA". In: *RNA* 5 (1999), pp. 1308–1325.

[88] Nikola A Ivica et al. “The Paradox of Dual Roles in the RNA World : Resolving the Conflict Between Stable Folding and Templating Ability”. In: *J. Mol. Evol.* 77 (2013), pp. 55–63.

[89] E Szathmáry. “The origin of replicators and reproducors”. In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361 (2006), pp. 1761–1776.

[90] Vera Vasas et al. “Evolution before genes”. In: *Biol. Direct* 7 (2012), pp. 1–14.

[91] Jean-paul Douliez et al. “Artificial Cell Models Catanionic Coacervate Droplets as a Surfactant-Based Membrane-Free Protocell Model Angewandte”. In: *Angew. Chem. Int. Ed.* 56 (2017), pp. 13689–13693.

[92] J.G. Gall et al. “Assembly of the Nuclear Transcription and Processing Machinery: Cajal Bodies (Coiled Bodies) and Transcriptosomes”. In: *Mol. Biol. Cell* 10.12 (1999), pp. 4385–4402.

[93] C P Brangwynne et al. “Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation”. In: *Science (80-.).* 324.5935 (2009), pp. 1729–1732.

[94] S. Jain et al. “ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure”. In: *Cell* 164 (2016), pp. 487–498.

[95] Simon Alberti. “The wisdom of crowds : regulating cell function through condensed states of living matter”. In: *J. Cell Sci.* 130 (2017), pp. 2789–2796.

[96] E D Goddard. “Polymer-surfactant interaction part ii. polymer and surfactant of opposite charge”. In: *Colloids and Surfaces* 19 (1986), pp. 301–329.

[97] Björn Drobot et al. “Compartmentalised RNA catalysis in membrane-free coacervate protocells”. In: *Nat. Commun.* 9 (2018), p. 3643.

[98] Kevin Leu et al. “The prebiotic evolutionary advantage of transferring genetic information from RNA to DNA”. In: *Nucleic Acids Res.* 39 (2011), pp. 8135–8147.

[99] Lihi Adler-abramovich et al. “Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria”. In: *Nat. Chem. Biol.* 8.8 (2012), pp. 701–706.

Books

[3] A I Oparin. *Origin of Life*. Dover, 1952.

[16] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford: Oxford University Press, 1995.

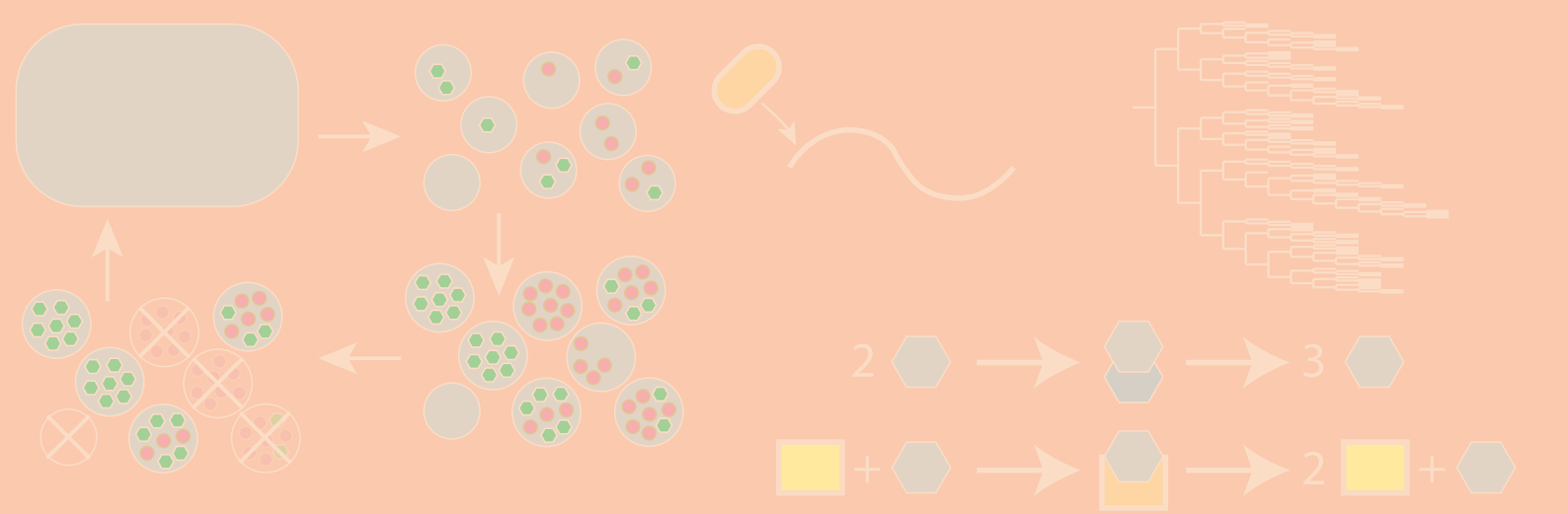
[37] Dilip Kondepudi and Ilya Prigogine. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. New York: Wiley, 1998.

[56] A.G. Cairns-Smith. *Seven Clues To The Origin of life: a scientific detective story*. Cambridge, UK: Cambridge University Press, 1985.

[71] Christian De Duve and Neil O Hardy. *A guided tour of the living cell*. New York: W. H. Freeman and Company, 1984.

[85] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press, 1993.

[100] W Zhang and Y Jin, eds. *Dynamic Covalent Chemistry: Principles, Reactions, and Applications*. 2018.



8. Transient compartments

The chemical composition of organisms is considerably different from their environment. Transport enzymes, carrier molecules [1, 2] (siderophores, ionophores, chalkophores), and vesicular secretion allow required nutrients to be picked up from the environment. Inside the cell, sophisticated molecular machinery is assembled from these nutrients, which are kept inside by virtue of cell membranes and other compartments.

Together, this allows chemical reactions to be precise and proceed at required rates, to remove waste from the system, to prevent side-reactions and to keep a metabolism thermodynamically spontaneous. Were the contents of a cell simply mixed with an environment, it would lack the precision, concentration and thermodynamic forces to remain viable. In fact, even mixing the contents within a single cell may be disastrous: a cell itself is a pastiche of coacervates and cellular compartments [3] with mutually incompatible chemistries, e.g. due to local pH, redox equilibria and digestive activity.

When considering the origins of life, it is not clear what chemical processes we start with. What is evident, however, is that such a process has constraints on concentrations, precision of reactions and associated thermodynamic forces. In terms of current scenarios, this means RNA/peptide/XNA/etc. worlds need abundant, chemically activated monomers, lipid world needs abundant amphiphiles, iron-sulfur world needs abundant iron and sulfur. These compounds then need to favorably engage in specific chemistry (and transport) and, in one way or another, perform some replication process on which 'chemical evolution' acts. For this and other tasks, compartmentalization is a benefit.

In this section, we will start by considering why compartments and heterogeneous environments are ubiquitous in biology and what kinds of capacities may be pertinent for abiogenesis. Subsequently, we will consider division in compartments and multilevel selection. For many, this is a key reason to consider compartmentalization. Then, the model based on 'transient compartmentalization' is discussed in detail. The sections on 'transient compartmentalization' largely recapitulate a work that is published in PRL[4] and another work in submission for which a preprint is available on BioRxiv[5].

Inspired by these papers, the specific case of transient compartmentalization of a replicase has

recently been studied by Laurent, Lacoste and Peliti, to appear in Life, and for which a preprint is available[6]. In the final section of this chapter, this parasite problem will be revisited using a stochastic description on the level of monomer incorporation and complex formation. This leads to the elucidation of a new type of parasite-induced catastrophe.

8.1 Physical aspects of Compartments

8.1.1 A definition of a compartment

In its colloquial sense, a compartment is a closed impermeable space. The Cambridge Dictionary describes a compartment as: *a separate part of a piece of furniture, equipment, or a container with a particular purpose*. For our purposes, we do not need to be very strict with the property of being closed, separate or impermeable, what we want to designate is: "a locally distinct physical-chemical environment". On the appropriate timescale of observation, this distinctness persists.

While separated in one way or another, a compartment may still be very similar to its surroundings. Consider a rock in contact with seawater, which has a little pore in it of length l. Let us inject some dye Y in this pore and consider a situation where it only exchanges with the sea by diffusion. In this situation, the characteristic timescale τ for concentration relaxation follows from the geometry and the diffusion constant D_Y

$$\tau = \frac{l^2}{D_Y}. \tag{8.1}$$

For a centimetric cylindrical pore filled with water and small dye molecules ($D \sim 10^{-5}\text{cm}^2/\text{s}$), this relaxation takes several days. If the contact area with the sea is small, the relaxation will take considerably longer. Within this time, the cylindrical pore can maintain a meaningfully different composition with respect to the surrounding seawater. For phenomena operating on shorter timescales, such pores then provide locally distinct chemical environments or 'compartments'.

In practice, we often think of compartments enclosed by membranes, or liquid droplets. Such compartments may readily exchange some compounds while other ones (e.g. large polymers, charged compounds) are barely exchanged at all.

8.1.2 Examples of compartments and their uses

In biochemistry, each biomolecule has its time and place, which can be enforced through the use of compartments. Here, let us take a quick look at a small number of biological uses of compartments, which come in very different sizes

- Storage and homeostasis: Many key compounds of interest to living systems are not externally supplied in the form or concentration desired, and these conditions may change considerably over time. Internally, however, concentrations can be controlled and regulated, by making compartments in which essential compounds can be stored. As an example, glucans are assembled in specialized compartments to form large polymeric reserves. This polymerization happens passively, and small sugar moieties can be exchanged selectively[7]. The dynamic equilibrium that ensues allows to retrieve sugar monomers when consumption outpaces production, and store excess sugar when production is high.
- Protection: an organism inherently has a different chemical composition compared to its environment, which can contain many toxic species or other inhospitable conditions. Controlled exchange via pores, transport proteins and other machinery helps to keep some of these risks at bay. Particular environments may be too inhospitable, a situation for which some organisms have devised cryptobiotic strategies. In cryptobiosis, an organism enters a state where metabolic processes are halted and a dormant state is entered. This happens notably during dessiccation, and cryptobiosis is often accompanied by a contraction of shape

and the introduction of an extra layer of insulation, such as akinetes (e.g. in cyanobacteria), or cysts made out of chitin (e.g. rotifers) [8].

- Separating incompatible chemistries: within a cell, different chemical processes need to be performed that can strongly interfere with each other, or have orthogonal requirements for their chemical environments. Bacteria performing the anammox (anaerobic ammonium oxidation) have specialized compartments where the overall conversion

$$NH_4^+ + NO_2^- \rightarrow N_2 + 2H_2O \tag{8.2}$$

powers the metabolism by the creation of a proton gradient. To maintain this gradient, and to keep toxic intermediates like hydrazine contained[9], the compartment is sealed off with a mix of highly specialized ladderane lipids, which let protons leak through around 10 times slower than regular lipids. Plant nodules are an example where compartments can be used to induce cooperation. In these nodules, nitrogen fixating bacteria are protected from the high oxygen levels of air, which destroy their fragile nitrogenase machinery used to fixate nitrogen, in an overall reaction that is often written as[10]:

$$N_2 + 16ATP + 8e^- + 8H^+ \longrightarrow NH_3 + H_2 + 16ADP + 16Pi \tag{8.3}$$

In fact, these bacteria do require minute amounts of oxygen, which is supplied and buffered at low concentration by the Leghemoglobin enzyme. The plant also supplies energy in the form of sugars, and notably receives fixated nitrogen species in exchange. Another macroscopic example can be found in the bombardier beetle[11], which has separate glands producing H_2O_2 and hydroquinone ($C_6H_4O_2$). The beetle derives its name from its capacity to inject these reactive chemicals in a compartment supplied with catalyst, to rapidly perform the highly exothermic reaction

$$C_6H_4O_2 + H_2O_2 \longrightarrow C_6H_6O_2 + H_2O, \tag{8.4}$$

the released heat evaporates the reaction mixture, generating a gas pressure used to propel a hot mix of noxious chemicals at predators.

Compartments in Origins of Life

The large number of organelles, vesicles and coacervates found in living systems are a testament to the successful exploitation of compartments on various scales. In the context of the origins of life, compartments have been considered as important ever since Oparin[12], notably as a means to locally concentrate the necessary prebiotic chemicals to high enough concentrations. Coacervates of peptides and RNA are now actively being explored [13, 14] in origins of life and are increasingly being proposed in prebiotic scenarios. Many other compartments have been advanced, such as rock pores[15], aerosols[16], lipid vesicles [17, 18, 19], mica sheets [20], FeS compartments [21] and mineral surfaces [22, 23] to name but a few.

Another key feature of compartmentalization, and the main object of this chapter, is the formation of a separate chemical collective. Such a higher order organization can be identified as a discrete entity on the compartment level. When compartments can show a difference in their rates of survival or reproductive success, a layer of selection is introduced that is absent in bulk chemistry. Such extra layer of selection provides an evolutionary platform that can stabilize coexistence and cooperation[24, 25] on lower levels, thereby maintaining states that would not survive in bulk.

When thinking of compartments and multilevel selection in origins of life, we tend to consider replicating entities akin to protocells. In the rest of this chapter, we will explore what happens if we relax this assumption, by having chemical collectives be encapsulated and released by the environment instead. In this picture, there is still multilevel selection, but discrete compartment-level states do not propagate themselves directly.

8.1.3 Mechanisms with compartments

In 1965, Spiegelman showed experimentally that RNA could be replicated by an enzyme called $Q\beta$ RNA replicase, in the presence of free nucleotides and salt[26]. Interestingly, he noticed that as the process is repeated, shorter and shorter RNA polymers appear, which he called parasites. Typically, these parasites are nonfunctional molecules that replicate faster than the RNA polymers introduced at the beginning of the experiment and that for this reason tend to dominate. Eventually, a polymer of only 218 bases remained out of the original chain of 4500 bases, which became known as Spiegelman's monster.

In 1971, Eigen conceptualized this observation by showing that for a given accuracy of replication and relative fitness of parasites, there is a maximal genome length that can be maintained without errors [27]. To improve the accuracy of replication, however, would require having a functional replicase and/or error-correcting machinery. These are, to the best of our knowledge, long (argued to be >200 nt for RNA, if they exist) species with specific sequences, which would imply that such species only becomes viable above a certain threshold length.

This result leads to an interesting paradox: functional, viable replicators need to be copied with high accuracy, which requires them to be long and with sufficient machinery to be accurate. However, an evolutionary trajectory to acquire these capacities would already need that large degree of accuracy.

This paradox and the associated error catastrophe due to parasites are now considered to be key aspects[25, 28] of Origins of life research*

In the eighties, an instructive theoretical solution to the parasite problem was proposed in the form of the Stochastic corrector model, [24, 29] (see Fig 8.1) inspired by ideas of group selection [30]. In the Stochastic corrector model, small groups of replicating molecules grow in a deterministic way in compartments, to a fixed final size called the carrying capacity. Then, the compartments are divided, and their content are stochastically partitioned between two daughter compartments (in principle, the number of daughters can be varied).

Thanks to the variability introduced by this stochastic division, and to selection acting on compartments, a coexistence is possible between replicators and parasites despite the difference in their growth rates. The Stochastic corrector has also been considered to explain other evolutionary processes, for example the emergence of large chromosomes at the expense of smaller genomes[25].

The efficiency of the stochastic corrector mechanism depends critically on the noise in the inoculation, which is controlled by the carrying capacity. If the carrying capacity is large, the final size before division is large. When division happens, fluctuations will be very small, and growth will be essentially deterministic, leaving no variation for group selection to act on. On the other end, if the final size is too small, division leads to giant fluctuations in the composition of offspring. Frequently, this will lead to random loss of replicators and compositions that hamper survival [29].

Ultimately, if too small a fraction of daughter cells is viable, the dividing compartments go extinct. If division yields, on average, n daughter cells, more than $1/n$ should survive, and subsequently grow in size by a factor n. In the particular instance that $n = 2$, at least half of the compartments need to be viable. This seemingly trivial fact is in fact an important consequence of permanent encapsulation, which binds the fate of a compartment to its contents, and vice versa. It paves the way for multilevel selection and the propagation of successful lineages, a prerequisite for much of the options that evolution has to offer.

It is instructive to note that we are here interested in the selection mechanism of the stochastic corrector, not its exact replication chemistry. Various instances of the GARD model [31, 32, 19]

*Note that the issue of Eigen's paradox is only a key question if such a situation occurred in the first place. It can equally well be thought of as a thought experiment that demonstrates the absurdity of this situation, and a motivation to find an alternative trajectory. Eigen's error catastrophe, on the other hand, can be considered as a more universal problem.

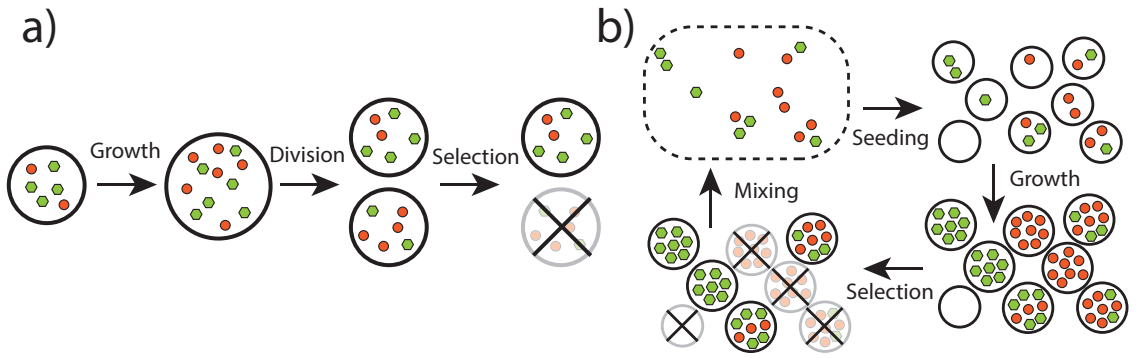


Figure 8.1: A sketch of a) the stochastic corrector model, b) transient compartmentalization. Both exhibit growth, selection and noisy inoculation of new compartments.

can have this division mechanism as well. In amphiphile-GARD, often taken as synonymous for GARD, the replication of its components relies on cross-catalytic incorporation, whereas in the stochastic corrector template replication is the mechanism of choice [33]. We hope that separate terminology will be proposed that generalize this phenomenon, at present we will simply refer to the growth-stochastic division-selection cycle as a stochastic corrector, irrespective of its chemistry.

A recently proposed mechanism, transient compartmentalization[34], is similar in spirit to the stochastic corrector, with the twist that compartments are destroyed after one round. Transient compartmentalization is schematically depicted in Fig. 8.1. It shares some important features with the Stochastic corrector[24] model: there is noise in the inoculation step in which molecules from a large pool are used to seed compartments, followed by growth. Then, selection is performed on a compartment level. The essential difference comes with the mixing step, where the selected compartments are removed and their contents mixed up, e.g. because of a process in the environment with a characteristic timescale τ_{cyc} that sets the cycle frequency. It can also be an experimentally controlled parameter, as was the case in the pioneering experiments[34] that inspired this chapter.

When replicating molecules are freed from their compartments, molecular replication is no longer constrained by a cell cycle, which means copy numbers can be much larger (or smaller) than double their initial number. Indeed, in Ref. [34], initial populations of RNA polymers were of size $n = O(1)$, which subsequently grew to $N = O(10^6)$ copies. As long as the environment supplies transient compartments, only a fraction larger than n/N needs to survive longer than τ_{cyc} to maintain the population of replicating molecules. If the success of a compartment depends on its contents (like for the stochastic corrector), group selection can act on transient compartments.

The composition of a compartment can enhance its reproductive success. For example some compounds may stabilize a compartment (e.g. amino acids can stabilize lipid vesicles [35]), chelate degradative catalysts (e.g. Mg^{2+} ions), buffer a desired chemical environment or improve the influx of metabolites[36] (see also Sec. 5.4). The reverse can also be true: compounds may destabilize a compartment, degrade metabolites, catalyze harmful side reactions, harness all replication machinery and so forth. While this may lead to rich and complex phenomena, the effect on survival can often be encoded by a low-dimensional composition-dependent fitness function $f(\bar{x})$. By deriving general results for large classes of fitness functions, we can learn something about a large variety of scenarios. Another example of a study with general implications on transient compartmentalization, quantified co-encapsulation effects in the context of directed evolution experiments [37].

Transient compartmentalization is a mode of multilevel selection that captures several mechanisms proposed in scenarios for the origins of life, based on various types of compartments (e.g. lipid vesicles [38], pores [39, 40], inorganic compartments [41], coacervates [12, 42] or aerosols

[16]) or various protocols of transient compartmentalization [43, 44]. Of particular interest here is a recent experiment, in which small droplets containing RNA in a microfluidic device [34] were used as compartments. In this experiment, a catalytic RNA was used as a proxy for a functional species/ functional replicator in competition with a nonactive parasite. We have used this system as a model to illustrate the theory of transient compartmentalization[4].

The related issue of cooperation between producers and non-producers has been discussed before [45]. Spatial clustering can lead to similar effects as compartmentalization in favoring the survival of cooperating replicators [46, 47]. These ideas were combined in a recent study of a population of individuals growing in a large number of compartmentalized habitats, called demes [48].

Simpler cells and plausibility

Overcoming error catastrophes has been an important motivation for studying compartment models, but the utility of these models goes well beyond this problem: they describe the foundations of multilevel selection. In formulating prebiotic scenarios and a synthesis of chemical evolution, multilevel selection can be expected to fulfill a major role. Although traditionally these approaches have focused on RNA replicase scenarios, the mathematical frameworks can also be extended to other systems, as is exemplified by GARD.

The true merit of the stochastic corrector and transient compartmentalization is that these models show interesting things about a ubiquitous unit in biology: the cell. Many of the characteristic properties of a cell derive from sophisticated machinery. Whether early compartmentalized replicators or protocells had all those properties is far from obvious.

Most eukaryotic cells partition their chromosomes among daughters. This property is facilitated by mitotic spindles. A machinery employed by some bacteria involves the placement of a septum in the middle of the cell, with a positioning informed by chemical gradients. This ensures that each daughter receives one of two circular DNAs. The stochastic corrector demonstrates that such a machinery, or indeed such a partitioning property, may not have been required from the start[25]. Their absence, however, introduces compositional noise that strongly affects the tolerated selection pressures and overall survival. This compositional noise is a key ingredient in the evolutionary tradeoffs that a stochastic corrector faces[29].

Ideally, a prebiotic scenario would allow to conceptualize most of chemical evolution as a sequence of gradual steps. A sudden transition from e.g. naked genes[33] on a surface [49] to encapsulated genes in dividing compartments is in this sense quite dramatic. Cell division requires machinery or mechanisms, which may not have been there from the start (but see Ref.[50] for an elegant proposal). In additions, the vesicles that are often proposed as protocells are impermeable to the biomonomers (nucleotides, amino acids)[36, 51] that accompany such a scenario.

As a selection mechanism, it is very primitive compared to the stochastic corrector. Its use is to enrich the pool in compounds that would replicate poorly in bulk. If those compounds are copolymers with particular subsets of sequences ('quasispecies'), such a mechanism can maintain sequence information. The mixing step, however, precludes the existence of lineages, and erases the information associated with the compartment composition. This is a key distinction with respect to e.g. GARD or the stochastic corrector, which transmit their composome to the next generation.

This absence of composome transmission is compensated by the simplicity of the mechanism. This simplicity requires less from the chemistry and the environment, making the mechanism a plausible and general stepping stone towards the development of more sophisticated selection mechanisms. In this sense, transient compartmentalization may have preceded cell-division[†].

[†]Or it may not, a mechanism is not a proof, especially when other interesting mechanisms[52, 53] can be imagined. See also Ch.9 for a complementary perspective: there can plausibly be a multitude of multilevel mechanisms operating simultaneously.

8.2 Transient Compartmentalization

Let us now introduce more formally the transient compartmentalization model for general replicating molecules in compartments exploiting a common resource. We start from a pool of molecules, which contains a large number of two types of replicating molecules, which we call for simplicity A and B. Let the fraction of A molecules in this pool be x. These molecules then seed a large number of compartments, which is considered to be infinite. A given compartment will contain n replicating molecules, out of which m will be of A type and the remaining ones of B type. Since this number is small in comparison with the number of molecules of the initial pool, n is a random variable drawn from a Poisson distribution of parameter λ, while the number m follows a binomial distribution $B_m(n,x)$. The resulting probability distribution for seeded compartments is then

$$P_\lambda(n,m,x) = \text{Poisson}(\lambda,n)B_m(n,x). \tag{8.5}$$

The replicating molecules A and B are involved in separate autocatalytic cycles, exploiting a common resource C, yielding the simplified overall reactions

$$\text{A}+\text{C} \underset{-g_1}{\overset{g_1}{\rightleftharpoons}} 2\text{A}+\text{D}, \tag{8.6}$$

$$\text{B}+\text{C} \underset{-g_2}{\overset{g_2}{\rightleftharpoons}} 2\text{B}+\text{F}, \tag{8.7}$$

where D and F are product molecules. In principle, we can add other resources and products and change the stoichiometry as long as we suppose C to be limiting and to reproduce autocatalytically. Any necessary non-replicating molecules and catalysts are assumed to be present in sufficiently large numbers in the compartments.

After seeding, the numbers of A molecules, m, and of B molecules, y, grow exponentially and independently so that

$$\bar{m} = me^{\alpha T}, \tag{8.8}$$

$$\bar{y} = (n-m)e^{\gamma T}, \tag{8.9}$$

with $T(m,n)$ the time which marks the end of the exponential growth phase, $\bar{m}$ the number of A molecules and $\bar{y}$ the number of B molecules at time T. The autocatalytic reactions of Eq. (8.6) eventually slow down, e.g. due to the exhaustion of a common resource C, or due to the saturation of nonreplicating catalytic sites.

For simplicity, let us assume that the growth phase ends when $N = \bar{m}+\bar{y}$, where N is the same constant for all compartments. Now, the final composition at this end time T is mainly controlled by the ratio $\Lambda = e^{(\gamma-\alpha)T}$. Here, we do not describe the saturation which could be done more precisely using the notion of carrying capacity [54]. In that case, the growth would be described by logistic equations and the carrying capacity would be equal to N. Note that N can be many times larger than n, due to the absence of a division step (which would impose $N \approx 2n$). This means that a smaller fraction (at least n/N) of compartments is enough to carry the functional molecules to the next generation. For a dividing cell, on average at least half of its daughter compartments must survive to avoid extinction of the population.

The fraction of A molecules at the end of growth phase can be well approximated as

$$\bar{x}(n,m) = \frac{\bar{m}}{N} = \frac{m}{n\Lambda-(\Lambda-1)m}. \tag{8.10}$$

If B grows faster, we have $\gamma > \alpha$, and thus $\Lambda > 1$, which is the regime considered in Ref. [4]. In Sec 8.3, we also consider regimes in which $\gamma < \alpha$.

We now implement selection at the compartment level. Selection can in general be described by a selection function $f(\bar{x}) \geq 0$. In our work, we have assumed that the selection function only

depends on the final composition $\bar{x}$ of the compartment. A natural choice for f is a monotonically increasing function of $\bar{x}$. As an example, we will use the sigmoidal function

$$f(\bar{x}) = 0.5\left(1+\tanh\left(\frac{\bar{x}-x_{th}}{x_w}\right)\right), \tag{8.11}$$

where x_{th} and x_w are dimensionless parameters, which describe respectively a threshold in the composition and the steepness of the function.

The compartments which have passed the selection step are then pooled together, forming a new pool of molecules from which future compartments can be seeded. The fraction of A molecules, x' of this new ensemble is the average of $\bar{x}$ among the selected compartments

$$x' = \frac{\langle \bar{x} f(\bar{x})\rangle}{\langle f(\bar{x})\rangle}, \tag{8.12}$$

which is equivalent[‡] to

$$x'(\lambda,x) = \frac{\sum_{n,m}\bar{x}(n,m)f(\bar{x}(n,m))P_\lambda(n,x,m)}{\sum_{n,m}f(\bar{x}(n,m))P_\lambda(n,x,m)}. \tag{8.13}$$

The transient compartmentalization cycle is then repeated, starting with the seeding of new compartments from that pool of composition x'. The formula Eq. (8.13) can be generalized to compartments with different final population sizes, by replacing the selection function with a modified selection function f'. Having N be the typical final size, we have for a compartment of final size N' the function

$$f'(\bar{x}) = \frac{N'}{N}f(\bar{x}). \tag{8.14}$$

Upon repetition of the protocol, the pool composition typically converges to a fixed point x^*, which is a solution of

$$x = x'(\lambda,x). \tag{8.15}$$

The stability of the fixed point x^* changes when

$$\left.\frac{dx'}{dx}\right|_{x=x^*} = 1. \tag{8.16}$$

It is implicitly assumed that $x'(x)$ is a sufficiently smooth function of x for this derivative to be defined.

8.2.1 Application to ribozyme-parasite dynamics

The above model has been introduced in Ref. [4] to describe replication of RNA ribozymes (resp. parasites) in compartments, which play the role of the A molecules (resp. B molecules). In this case, in addition to the replicating molecules, a large amount of $Q\beta$ replication enzymes $n_{Q\beta}$ and activated nucleotides n_u (serving as C molecules) is supplied in each compartment with the same concentration in each compartment. At the end of this growth phase, we have $n_{Q\beta} \approx N = \bar{m}+\bar{y}$, at which point further growth is limited by the number of replication enzymes. After time T, the

[‡]Of course, a fixed λ is not generally appropriate, and a more general model should consider a recursion for λ' too. This extension has been explored in Ref[6]. We can consider a fixed λ to arise from e.g. a carrying capacity in the mixed pool, a dilution, osmotic equilibration or some other restoring force, or simply suppose that it changes slowly. Alternatively, we may simply hope that we still learn something if we neglect this contribution.

growth will be linear instead of exponential. At that point, the composition x (defined by the relative fraction of ribozymes) no longer changes. The average number $\bar{m}$ of ribozymes and $\bar{y}$ of parasites grow according to

$$\begin{aligned} \bar{m} &= m\exp(\alpha T), \\ \bar{y} &= (n-m)\exp(\gamma T), \end{aligned} \tag{8.17}$$

where T denotes the time and α (resp. γ) denote the average growth rate of the ribozymes (resp. parasites) during this exponential growth phase[§]. The relevant quantity for this dynamics is the ratio of the number of daughters of one parasite molecule and that of the daughters of one ribozyme molecule: $\Lambda = \exp((\gamma-\alpha)T))$. Note that $\Lambda > 1$ since $\gamma > \alpha$.

The compartments are then selected according to a selection function $f(\bar{x}) \geq 0$. In Ref. [34], a measurement of the synthesis of a dye molecule by photodetection was used to accept or reject compartments. This selection served as a proxy for a more general fitness effect due to catalytically active RNA. A specific form which is compatible with [34] is the sigmoid function given by Eq. (8.11) with $x_{th} = 0.25$ and $x_w = 0.1$.

Note that this function takes a small but non-zero value for $\bar{x} = 0$, namely $0.5(1-\tanh(x_{th}/x_w)) = 0.0067$, which represents the fitness of a pure parasite compartment. As a comparison with this function and Ref.[55], we also studied a linear selection function with the same starting point

$$f_{lin}(\bar{x}) = 0.0067 + \bar{x}. \tag{8.18}$$

Dynamical and asymptotic behavior

Instead of finding the steady state value of x, it is easier to evaluate $\Delta x = x'(\lambda, x) - x$ as a function of λ, which shows how the composition evolves over one round. The steady-state value corresponds to the line $\Delta x = 0$ separating negative values above from positive values below as shown in Fig. 8.4.

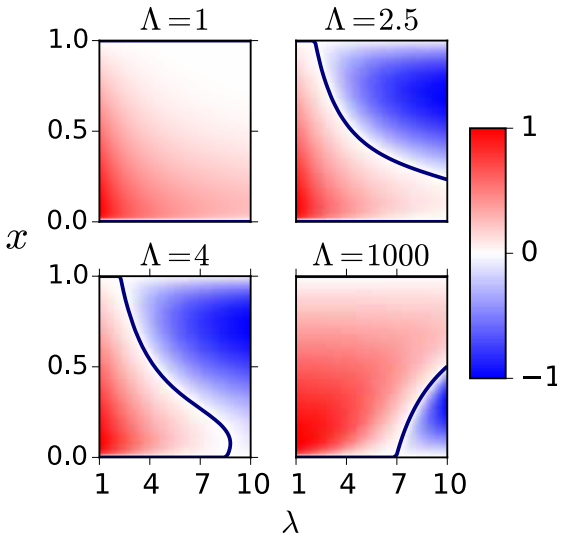


Figure 8.2: Contour plots of Δx for four values of $\Lambda = 1, 2.5, 4$ and 1000 in the plane (x, λ), with red (resp. blue) regions corresponding to $\Delta x > 0$ (resp. $\Delta x < 0$).

[§]The exact time T at which a compartment enters the linear regime depends on its initial composition m, n. In practice, we are often interested in $N \gg m, n$, where this dependence has a small effect on the results of the model, such that we can use the same value of $\mathbf{\Lambda}$ for all compartments, as demonstrated in detail in the Suppl. Mat. of Ref.[4].

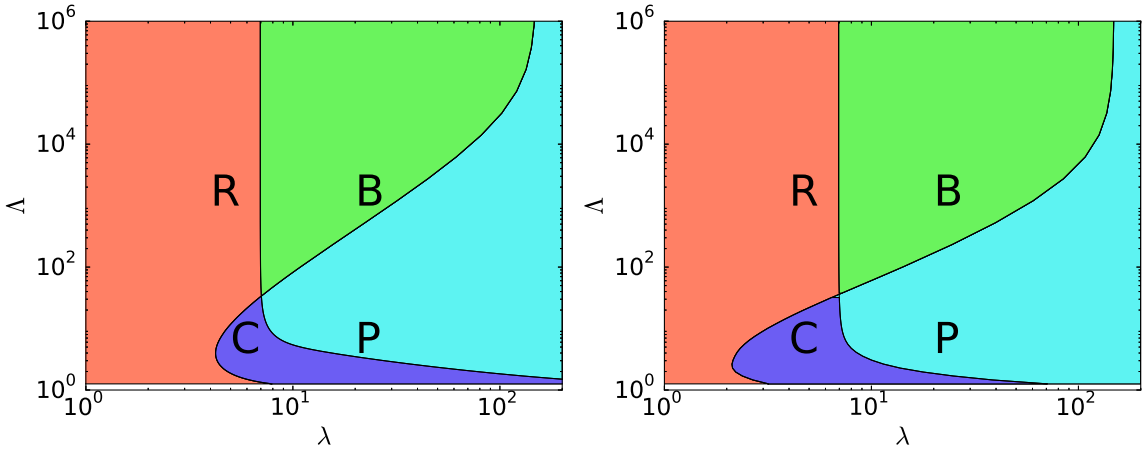


Figure 8.3: Left: Phase diagram of the transient compartmentalization dynamics with the linear selection function $f_{lin}(\bar{x}) = 0.0067 + \bar{x}$ in the (λ, Λ) plane. Right: idem with the sigmoidal selection function $f(\bar{x})$. Phases are R: pure Ribozyme, B: Bistable, C: Coexistence and P: pure Parasite.

We construct a phase diagram in the (λ, Λ) plane, by numerically evaluating the bounds of stability of the fixed point $x = 0$ from the condition:

$$\left.\frac{\partial x'}{\partial x}\right|_{x=0} = 1, \tag{8.19}$$

and similarly for the other fixed point $x = 1$. In Fig. 8.3, a phase diagram is plotted for both $f(\bar{x})$ and $f_{lin}(\bar{x})$ The phase diagram shows four distinct phases. In the orange (resp. light blue P region) region R, the only stable fixed point is $x = 1$ (resp. $x = 0$). In the green region, $x = 0$ and $x = 1$ are both stable fixed points. The system converges towards one fixed point or the other depending on the initial condition: for this reason, we call this region B for bistable. In the violet region, $x = 0$ and $x = 1$ are both unstable fixed points, but there exists a third stable fixed point x^* with $0 < x^* < 1$. We call this a coexistence region (C). All of these phases can be seen in Fig. 8.4.

The phase diagram is limited to four phases, because it is based on the evaluation of the stability of two fixed points. Where it is implicitly assumed that we will have at most three fixed points at the same time. This was found to be almost entirely valid, but small regions of the λ, Λ plane were found to have four fixed points, as shown in Fig. 8.4

Interestingly, the phase diagram for the linear function is very similar to the one for the sigmoidal function. This is a general feature which we will exploit on multiple occasions, and it follows from the asymptotes of the selection function. To this end, let us analyze some specific limits for which the asymptotes of the phase diagram can be computed exactly. Let us consider

- $\lambda \gg 1$: bulk behavior
- $\Lambda \gg 1$: *hard* parasites
- Λ close to 1: *soft* parasites

For large λ, we can neglect the fluctuations of n, i.e. the total number of replicating molecules (ribozymes plus parasites) in the seeded compartment. Indeed, n is Poisson distributed with parameter λ, therefore $\mathrm{Var}(n)/\lambda^2 = 1/\lambda \ll 1$. For large λ, Λ close to 1 and x close to 1 (resp. 0), the most abundant compartments verify $m = n$ or $m = n - 1$ (resp. $m = 0$ or $m = 1$). By considering only these compartments in the recursion relation, one finds that the condition of stability of the fixed point $x = 0$ leads to

$$\Lambda = 1 + \frac{f'(0)}{f(0)\lambda} + O\Big(\frac{1}{\lambda^2}\Big), \tag{8.20}$$

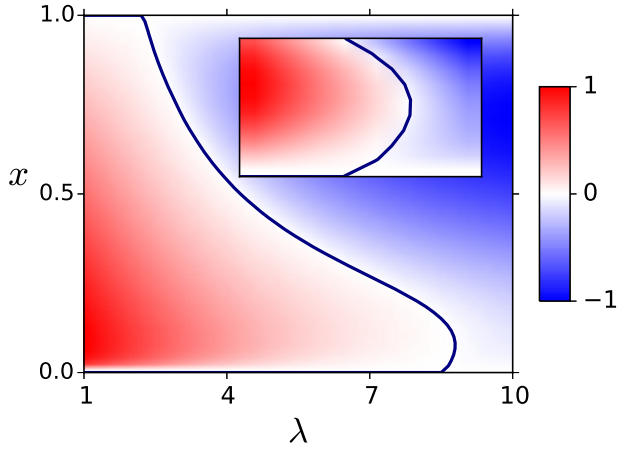


Figure 8.4: Contour plots of Δx vs. x for $\Lambda = 4$ in the plane (x, λ). Inset shows a blow-up of the region near $\lambda = 8$, which exhibits features of both the bistable and coexistence regions.

for an arbitrary selection function and $\Lambda \simeq 1 + 19.86/\lambda$ for the selection function of Eq. ((8.11)). This equation indeed characterizes the separation between the parasite and coexistence regime at large λ in Fig. 8.3. A similar equation is found for the fixed point at $x = 1$

$$\Lambda = 1 + \frac{f'(1)}{f(1)\lambda} + O\Big(\frac{1}{\lambda^2}\Big), \tag{8.21}$$

yielding $\Lambda \simeq 1 + 6.12\,10^{-6}/\lambda$ for this selection function for the separation between ribozyme and coexistence regions. For Λ close enough to 1, we have a ribozyme phase. The asymptotes given by (8.20) and (8.21) border the coexistence region in Fig. 8.3. This supports the observation that *soft* parasites can coexist with ribozymes.

Let us now study the *hard* parasite limit, namely $\Lambda \gg 1$, and finite λ. In this regime, we only need to consider three types of compartments: compartments made of pure ribozymes, such that $m = n \neq 0$, compartments containing parasites, and empty compartments, i.e. such that $n = 0$. One can introduce three inoculation probabilities for these cases p_{ribo}, p_{para}, and p_{zero}. Using Eq. (8.5), one finds

$$p_{ribo} = \sum_{n=1}^{\infty} \frac{x^n \lambda^n}{n!} e^{-\lambda} = (e^{\lambda x} - 1)e^{-\lambda}, \tag{8.22}$$

$$p_{zero} = e^{-\lambda}, \tag{8.23}$$

$$p_{para} = 1 - p_{ribo} - p_{zero} = 1 - e^{\lambda(x-1)}. \tag{8.24}$$

Let us assume that in compartments containing parasites, the parasites will overwhelm the ribozymes $x \to 0$. Inserting these values in (8.13), we find

$$x' = \frac{p_{ribo} f(1)}{p_{ribo} f(1) + p_{para} f(0)}. \tag{8.25}$$

Evaluating the fixed-point stability of $x = 1$ using (8.19), we find that the boundary value of λ satisfies

$$\lambda f(0) e^{\lambda} = (e^{\lambda} - 1) f(1), \tag{8.26}$$

for an arbitrary selection function. A similar calculation at the fixed point $x = 0$ leads to the other vertical separation line given by

$$\lambda f(1) = (e^{\lambda} - 1) f(0). \tag{8.27}$$

The solution of Eq. (8.26) (resp. Eq. (8.27)) is $\lambda \simeq 149.41$ (resp. $\lambda \simeq 6.95$) which compare well with the vertical separation lines in Fig. 8.3.

Let us now consider what happens in the limit $\lambda \to \infty$. For large λ, for Λ close to 1 and x close to 1 (resp. 0), the most abundant compartments verify $m = n$ or $m = n-1$ (resp. $m = 0$ or $m = 1$). As λ is large, we can neglect fluctuations in n and we can take $n = \lambda$. We therefore only look at the recursion for a typical compartment with $n = \lambda$, with a simplified notation $P_\lambda(n = \lambda, x, m) = P_\lambda(x, m)$, where

$$P_\lambda(x, m) = B_m(\lambda, x), \tag{8.28}$$

obtaining

$$x' = \frac{f(1) P_\lambda(x, \lambda) + \bar{x} P_\lambda(x, \lambda - 1) f(\bar{x})}{f(1) P_\lambda(x, \lambda) + P_\lambda(x, \lambda - 1) f(\bar{x})}, \tag{8.29}$$

where

$$\bar{x} = \bar{x}(\lambda, \lambda - 1) = \frac{\lambda - 1}{\lambda + \Lambda - 1} \simeq 1 - \frac{\Lambda}{\lambda}. \tag{8.30}$$

We have therefore

$$x' = \frac{x^\lambda f(1) + \lambda x^{\lambda-1}(1-x)\bar{x} f(\bar{x})}{x^\lambda f(1) + \lambda x^{\lambda-1} f(\bar{x})} = \frac{x f(1) + \lambda (1-x)\bar{x} f(\bar{x})}{x f(1) + \lambda (1-x) f(\bar{x})}. \tag{8.31}$$

Taking the derivative with respect to x we obtain

$$\frac{dx'}{dx} = \frac{\lambda (1-\bar{x}) f(1) f(\bar{x})}{(x f(1) + \lambda (1-x) f(\bar{x}))^2}, \tag{8.32}$$

which for $x = 1$ yields

$$\left.\frac{dx'}{dx}\right|_{x=1} = \frac{\lambda (1-\bar{x}) f(\bar{x})}{f(1)}. \tag{8.33}$$

Thus the boundary defined by the equation

$$\left.\frac{dx'}{dx}\right|_{x=1} = 1, \tag{8.34}$$

is given by

$$\Lambda \simeq 1 + \frac{f'(1)}{f(1)\lambda} = 1 + 6.12\, 10^{-6}/\lambda. \tag{8.35}$$

Evaluating the stability around the fixed point $x = 0$ we obtain likewise

$$x' = \frac{\lambda x (1-x)^{\lambda-1} \bar{x} f(\bar{x})}{(1-x)^\lambda f(0) + \lambda x (1-x)^{\lambda-1} f(\bar{x})} = \frac{\lambda x \bar{x} f(\bar{x})}{(1-x) f(0) + \lambda x f(\bar{x})}, \tag{8.36}$$

where now $\bar{x}$ is given by

$$\bar{x} = \bar{x}(\lambda, 1) = \frac{1}{(\lambda - 1)\Lambda + 1} \simeq \frac{1}{\Lambda\lambda}. \tag{8.37}$$

Evaluating the derivative of $x'(x)$ at $x = 0$ we obtain

$$\left.\frac{dx'}{dx}\right|_{x=0} = \frac{\lambda \bar{x} f(\bar{x})}{f(0)}. \tag{8.38}$$

This gives the boundary as

$$\Lambda = 1 + \frac{f'(0)}{f(0)\lambda} = 1 + 19.8661/\lambda. \tag{8.39}$$

In ref. [34] a comparison was made of the system behavior as a function of the number of selection rounds in three possible protocols: (i) No compartments (bulk behavior), (ii) compartments with no selection, (iii) compartments with selection. Such a comparison based on our theoretical model is shown in Fig. 8.5 for parameter values corresponding to the coexistence region of Fig. 8.3. As expected, the fraction of ribozymes decreases towards zero rapidly in case (i), and somewhat

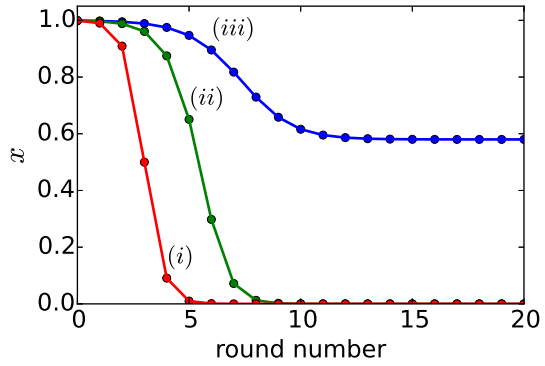


Figure 8.5: Evolution of the average ribozyme fraction x as function of the number of rounds for the three protocols, namely (i) No compartments (bulk behavior), (ii) compartments with no selection, (iii) compartments with selection. We choose $\lambda = 5$ and $\Lambda = 10$, corresponding to the coexistence region of Fig. 8.3.

less quickly in case (ii). Only in case (iii) is it possible to maintain a non-zero ribozyme fraction on long times. It is indeed observed that the ribozyme fraction eventually vanishes for protocols (i) and (ii) in the experiment of Ref. [34]. In case (iii), a stabilizing decrease of the ribozyme fraction towards coexistence is observed, which confirms the prediction in Ref. [34], where such a tendency was observed, but could be monitored for 10 rounds. In figure 8.6 we show the behavior of the distribution of the ribozyme fraction after the growth phase, i.e. $\bar{x}(n,m)$ (defined in Eq. (8.44)) as a function of round number. The parameters are $\Lambda = 5$ and $\lambda = 10$, corresponding to the parasite region, where the final state of the system is $x = 0$, and the initial condition is $x = 0.999$. Note that the distribution of $\bar{x}(n,m)$ is discrete, since many values are not accessible in the allowed range of n and m. At $t = 0$, it exhibits a sharp peak near $\bar{x} = 1$ coexisting with a broad peak at small values of $\bar{x}$. As time proceeds, the weight of the distribution shifts to the peak at small values of $\bar{x}$, since in this case selection is not sufficiently strong to favor the peak near $\bar{x} = 1$ and parasites eventually take over.

Comparison to experiments

In addition to predicting the phase diagram associated with the long-time compositions reached by this transient compartmentalization dynamics, our theoretical model makes also predictions regarding the evolution of the ribozyme fraction as function of the round number, *i.e.* the number

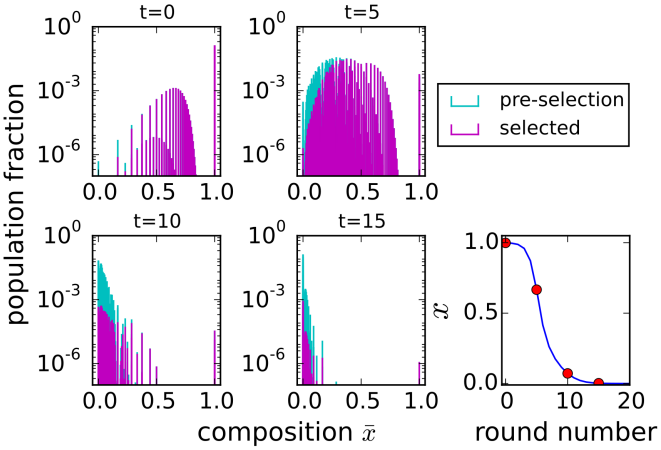


Figure 8.6: Evolution of the distributions of ribozyme fraction $\bar{x}(n,m)$ before and after selection at different times. The chosen times are shown as red circles in the lower right panel, which represents the evolution of the average fraction x as a function of the number of selection rounds.

Type	Length (nt) 2	$T_d(s)$	Relative r	Λ
Ribozyme	362	25.0	1.00	1
Parasite 1	245	20.7	1.21	13
Parasite 2	223	17.1	1.46	107
Parasite 3	129	14.6	1.71	473

Table 8.1: Lengths and doubling times for the parasites and ribozyme observed in Ref. [34], together with their relative aggressivity mesured by their relative growth rate r, and the corresponding values of Λ.

of completed cycles of compartmentalization. The model correctly reproduces that this fraction quickly goes to zero as function of the round number in bulk, less quickly with compartmentalization and no selection and even less quickly in the case of compartmentalization with selection. In the latter case, a finite fraction can be maintained for an infinite number of rounds provided λ is sufficiently small, corresponding to the coexistence region of the phase diagram.

In order to compare precisely the predictions of the model to the experiments of Ref. [34], it is important to know the value of key parameters such as Λ. Table 8.1 reports the experimental parameters measured in Ref. [34] for the ribozyme and three different parasites. The nucleotide length, its doubling time (T_d), its relative replication rate (r) from which we infer Λ in the final column. The doubling time T_d for the ribozyme is related to the growth rate α by $T_d = \ln(2)/\alpha$, and similarly the doubling times of the parasites is $T_d = \ln(2)/\gamma$.

In the experiment, a typical compartment contains λ RNA molecules that can be ribozymes or parasites, $2.6 \cdot 10^6$ molecules of Qβ replicase, and $1.0 \cdot 10^{10}$ molecules of each NTP. Replication takes place by complexation of RNA with Qβ replicase, which uses NTPs to make a complementary copy. This copy is then itself replicated to reproduce the original. There is a large amount of nucleotides, so that exponential growth of the target RNA proceeds until $N \approx n_{Q\beta}$. This large quantity of enzymes also means that in practice, the noise due to fluctuations in the number of enzymes should be very small. Starting from a single molecule, it takes $n_D = \log_2 n_{Q\beta} = 21.4$ doubling times to reach this regime. In a parasite-ribozyme mixture, we can estimate Λ using the

relative r:

$$\Lambda = \frac{2^{n_D}}{2^{n_D/r}} = 2^{n_D(1-\frac{1}{r})}. \tag{8.40}$$

8.3 A modified model with deterministic mutations

In the deterministic model, we assume that a fraction μ of replicated ribozyme strands mutate into parasites. Thus, the equations describing the evolution of m and y in the growth phase assumes the form

$$\begin{aligned} \dot{m} &= \alpha m - \mu m = (\alpha - \mu)m \\ \dot{y} &= \gamma y + \mu m, \end{aligned} \tag{8.41}$$

which yields for the first equation

$$\bar{m} = m e^{(\alpha-\mu)T}, \tag{8.42}$$

where $\bar{m}$ is again the number of ribozymes at the end of the growth phase and m the value at the initial time. Now substituting Eq. (8.42) into the equation for y, one finds

$$\bar{y} = \left(n - m + \mu m \frac{e^{(\alpha-\gamma-\mu)T} - 1}{\alpha - \mu - \gamma}\right) e^{\gamma T}. \tag{8.43}$$

The ratio between the number of daughters of one parasite molecule and the number of daughters of a ribozyme molecule is now renormalized by the rate μ: $\bar{\Lambda} = e^{(\gamma+\mu-\alpha)T} = e^{\mu T}\Lambda$, where Λ is the relative growth of parasites introduced previously in the mutation-free model.

The fraction of ribozymes at the end of the exponential phase is now given by

$$\bar{x}(n,m) = \frac{\bar{m}}{N} = \frac{m}{n\bar{\Lambda} - (\bar{\Lambda}-1)(1+\delta)m}, \tag{8.44}$$

where $\delta = \mu/(\alpha - \mu - \gamma)$. We call δ the mutation ratio, which is a dimensionless measure of mutation versus relative growth (competition). When $\delta \to 0$, we recover the mutation-free model, if $|\delta| \gg 0$ mutations become dominant.

Selected compartments are then pooled together, and the new average fraction of ribozymes becomes $x'(x, \lambda, \delta, \bar{\Lambda})$. Note that for nonzero mutation rate ($\mu > 0$), $x' = 1$ ceases to be a fixed point in this deterministic approach, since parasites will always appear at sufficiently long times. Therefore, the pure ribozyme (R) phase is no longer present in the phase diagram of fig. 8.7.

The fixed point $x' = 0$ however is still present. If this fixed point is stable, we have a pure parasite phase. If it is unstable, there is stable coexistence at a fixed composition. If more fixed points appear, multiple stable compositions are in principle be possible.

8.3.1 The prolific parasites regime ($\bar{\Lambda} \geq 1$)

Prolific parasites have a better bulk reproductive success than ribozymes, when $\bar{\Lambda} \geq 1$, which is equivalent to $\alpha \leq \mu + \gamma$ and $\delta < 0$. In a mutation-free model, this would imply necessarily a faster growth of parasites ($\alpha < \gamma$), but in the present case, we could also allow for slower parasites as compared to ribozymes (i.e. $\alpha > \gamma$), provided parasites are aided by a sufficiently high mutation rate μ.

The phase diagram is evaluated by testing the stability of the fixed point $x' = 0$. We find an asymptote behaving like $1/\lambda$ for large λ, and plateaus for small λ. The ends of these plateaus locate in the limit $\delta \to 0$ at the position of the vertical line separating the ribozyme and bistable phase in the original phase diagram.

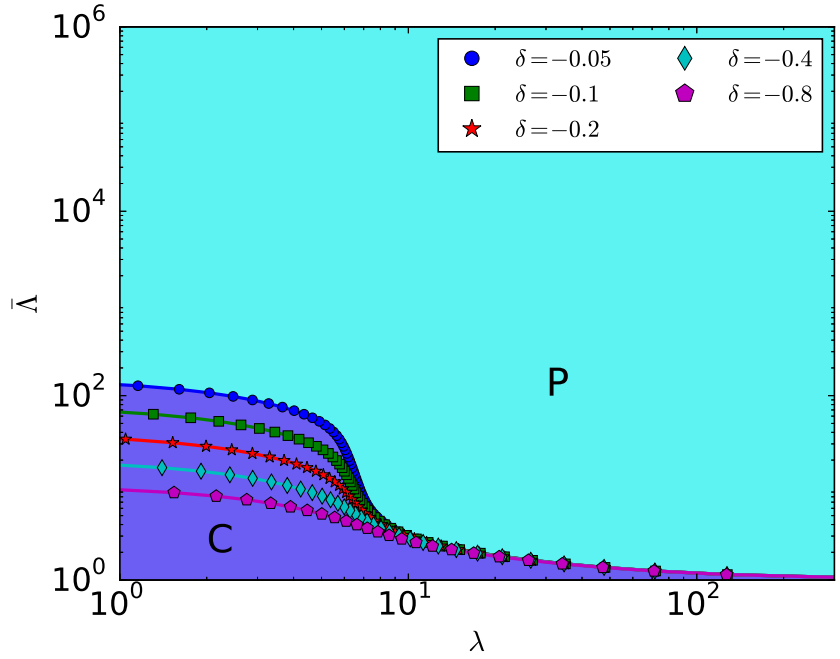


Figure 8.7: Phase diagram of the model with mutation in the case of prolific parasites. The selection function is given in Eq. (8.11). Phases are colored for $\delta = -0.05$, other separatrices are plotted for various mutation strengths δ. The possible phases are coexistence (C), pure parasite (P).

Let us first derive the right asymptote in the $\lambda \gg 1$ limit. In this limit, we evaluate x' by considering compartments of size λ

$$x' = \frac{\lambda x \bar{x} f(\bar{x})}{(1-x)f(0) + \lambda x f(\bar{x})}. \tag{8.45}$$

The fixed point stability condition $dx'/dx|_{x=0} = 1$ leads to

$$\left.\frac{dx'}{dx}\right|_{x=0} = \frac{\lambda \bar{x} f(\bar{x})}{f(0)}. \tag{8.46}$$

Upon substituting Eq. (8.44) evaluated at $m = 1, n = \lambda$ and approximating $f(\bar{x}) \approx f(0) + f'(0)\bar{x}$, (for $\lambda \gg 1, \bar{x} \ll 1$) we find a quadratic equation for $\bar{\Lambda}$, whose only physical solution ($\bar{\Lambda} \geq 1$) is

$$\bar{\Lambda} = \frac{\lambda - 2\delta - 2 + \sqrt{\lambda\left(4\frac{f'(0)}{f(0)} + \lambda\right)}}{2(\lambda - \delta - 1)}. \tag{8.47}$$

Since we consider monotonically increasing selection functions, $f'(0) > 0$. For $\lambda \gg -\delta$, we find

$$\bar{\Lambda} = 1 + \frac{f'(0)}{f(0)(\lambda - \delta - 1)} \approx 1 + \frac{f'(0)}{f(0)\lambda}, \tag{8.48}$$

which is the same expression as the one found in the mutation-free phase diagram [4]. This explains why there is a single asymptote as μ is varied in the $\lambda \gg 1$ limit.

The plateaus extend to very low values of λ. We can find their location by considering only compartments of size $n = 1$. In that case, the final compositions can be $\bar{x}(1,0) = 0$ or

$$\bar{x}(1,1) = \frac{1}{1 + \delta - \delta\bar{\Lambda}}. \tag{8.49}$$

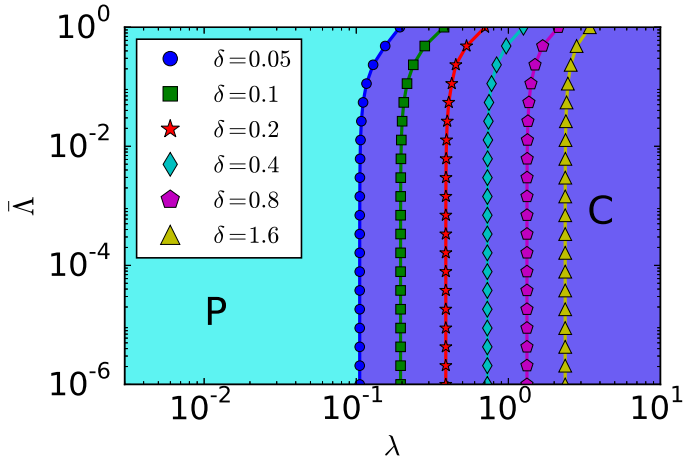


Figure 8.8: Phase diagram in absence of selection function for prolific ribozymes ($\bar{\Lambda} \leq 1$). Phases are colored for $\delta = 0.05$, separatrices are plotted for various mutation strengths δ. C: coexistence, P: pure parasite.

We then have for the composition recursion

$$x' = \frac{x\bar{x}f(\bar{x})}{(1-x)f(0)+xf(\bar{x})}. \tag{8.50}$$

Evaluating the derivative of $x'(x)$, we find

$$\frac{\bar{x}f(\bar{x})}{f(0)} = 1. \tag{8.51}$$

Substituting (8.49), we find that the location of plateaus obeys the implicit equation

$$\bar{\Lambda} = 1 + \frac{f(0)-f(\bar{x})}{f(0)\delta}. \tag{8.52}$$

8.3.2 The prolific ribozymes regime ($\bar{\Lambda} \leq 1$)

We now consider the opposite case where parasites are less prolific than ribozymes. This means $\alpha \geq \mu + \gamma$ and is equivalent to $\bar{\Lambda} \leq 1, \delta > 0$. This implies that $\alpha > \gamma$ (less aggressive parasites) and is reminiscent of a quasipecies scenario in which a fit ribozyme successfully outcompetes its parasites in bulk [27]. Since this can already happen in the absence of selection, we consider here the case where there is no selection, *i.e.* $f(\bar{x}) = 1$.

To analyze this regime we again assess the fixed point stability of $x' = 0$. We locate numerically the separatrix as shown in Fig 8.8. We obtain separatrices that for $\bar{\Lambda} \to 0$ tend to a fixed value of λ.

Let us start by observing that when $\bar{\Lambda} \to 0$, there are only two final compartment compositions for nonempty compartments: $\bar{x}(n,0) = 0$ or $\bar{x}(n,m) = 1/(1+\delta)$ for $m > 0$. We can now distinguish between three initial compartment compositions: (i) only parasites, (ii) no parasites, no ribozymes, and (iii) containing at least one ribozyme. Their associated seeding probabilities are:

$$\begin{aligned} p_{\text{para}} &= \sum_{n=1}^{\infty} \frac{(1-x)^n \lambda^n}{n!} e^{-\lambda} = (e^{\lambda(1-x)} - 1)e^{-\lambda} \\ p_{\text{zero}} &= e^{-\lambda} \\ p_{\text{ribo}} &= 1 - p_{\text{para}} - p_{\text{zero}} = 1 - e^{-\lambda x} \end{aligned} \tag{8.53}$$

In that case, we can write the composition recursion equation as

$$x' = \frac{1}{1+\delta}\frac{p_{\text{ribo}}}{p_{\text{para}}+p_{\text{ribo}}}, \tag{8.54}$$

The condition $dx'/dx|_{x=0} = 1$ yields the expression

$$\lambda = (1+\delta)(1-e^{-\lambda}), \tag{8.55}$$

for the asymptote. For $\lambda \ll 1$, we can expand the exponential in (8.55), to obtain

$$\lambda = \frac{2\delta}{1+\delta}, \tag{8.56}$$

which agrees very well with Fig 8.8.

Notice that here the coexistence phase is located to the right of the asymptotes, and the parasite phase to the left, whereas in Fig 8.7 it is the other way around. An intuitive way to understand this is to consider the limit $\lambda \to 0$. In this limit, nonempty compartments start with either a parasite or a ribozyme. The former will grow to a fully parasitic compartment, whereas the latter will contain ribozymes plus some parasites acquired by mutations. Therefore, at low λ, the ribozyme's capacity to outgrow parasites (competition) cannot be exploited, leading to ribozyme extinction.

It is only when ribozymes and parasites are seeded together that the differential growth rate becomes important, which becomes increasingly likely for higher λ. The phase boundaries in Fig. 8.10 mark the point where enough compartments engage in competition to allow for ribozyme survival. The mutation strength δ compares mutation rate to competition. When $\delta \to 0$, there is enough competition to ensure coexistence for all λ.

8.3.3 Error catastrophe

An error catastrophe corresponds to a situation where the accumulation of replication errors eventually causes the disappearance of ribozymes. Since there are only a parasite (P) and a coexistence phase (C) in the model with mutations, the error catastrophe means that the coexistence region shrinks at the benefit of the parasite phase as the mutation rate increases. One sees this effect in Fig. 8.10, which corresponds to the prolific parasites regime ($\bar{\Lambda} \geq 1$) discussed above. In this figure, we see a larger coexistence region in the small λ region, because there the compartmentalization is efficient to purge parasites. As the mutation rate increases however, this region shrinks because the compartmentalization fails to purge the more numerous parasites.

In Fig. 8.9, a particular example is provided where α and γ are fixed, such that $\bar{\Lambda}$ is fixed, and μ is varied. Since competition is fixed, we have $\mu \propto \delta$. The resulting steady-state value $x = x^*$ then decreases monotonically with μ, and reaches $x = 0$ when crossing the phase boundary in Fig 8.10. For small values of λ, this boundary corresponds to the plateau region, for larger values, this corresponds to the $1/\lambda$ asymptote. As can be seen in Fig 8.10, coexistence is stable for much higher values of the mutation rate μ when the compartment size λ is small. This means that compartmentalization with selection leads to a relaxed error threshold with respect to the bulk.

The error catastrophe was also studied in the absence of selection and was shown to be in the prolific ribozymes regime ($\bar{\Lambda} \leq 1$). In Fig. 8.11, an example of this case is shown, and there too, we see that the steady-state value of the ribozyme fraction x^* decreases as μ is increased, until it reaches the phase boundary in Fig 8.12. In contrast to Fig. 8.9, where the error threshold decreases as the size of compartments increases, the trend is just the opposite in Fig. 8.11, which is expected since the role of ribozymes and parasites are exchanged here as compared to the prolific parasites regime.

In the prolific parasites regime, $\bar{\Lambda} \leq 1$ with selection, it is interesting to recast the error threshold as a constraint on the length of a polymer to be copied accurately, as done in the original formulation

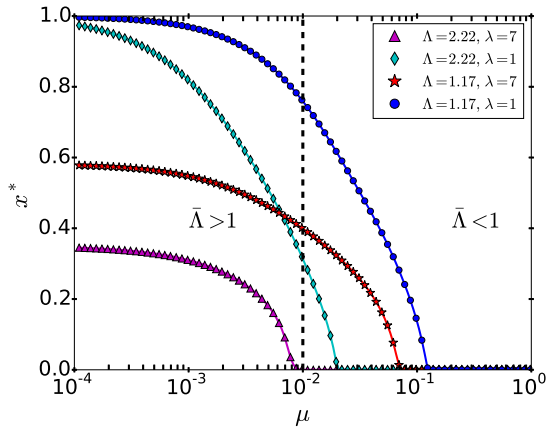


Figure 8.9: Steady state composition x^* as function of μ, $\alpha = 0.99, \gamma = 1.0$. Critical rates μ^* corresponds to separation between P and C phases in Fig. 8.10.

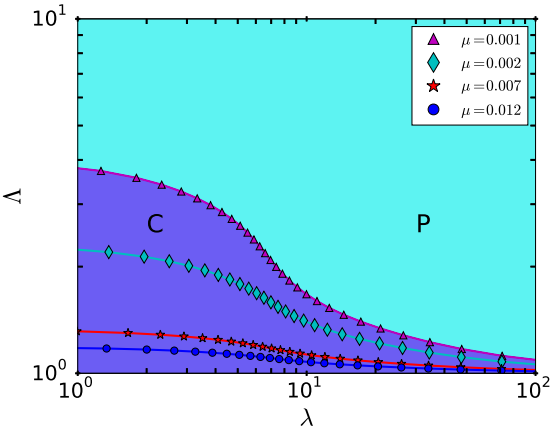


Figure 8.10: Phase diagram, drawn for f$\alpha = 0.99, \gamma = 1.0$. Separatrices are drawn for μ values close to μ^* in Fig. 8.9, corresponding to an error catastrophe.

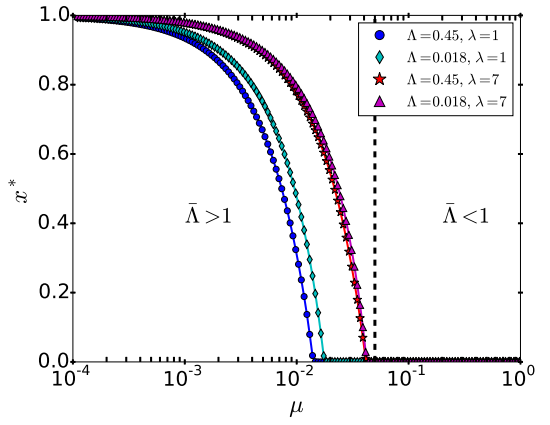


Figure 8.11: Steady state composition x^* as function of μ, $\alpha = 1.0, \gamma = 0.95$, in absence of selection ($f = 1.0$). Critical rates μ^* corresponds to separation between P and C phases in Fig. 8.12 .

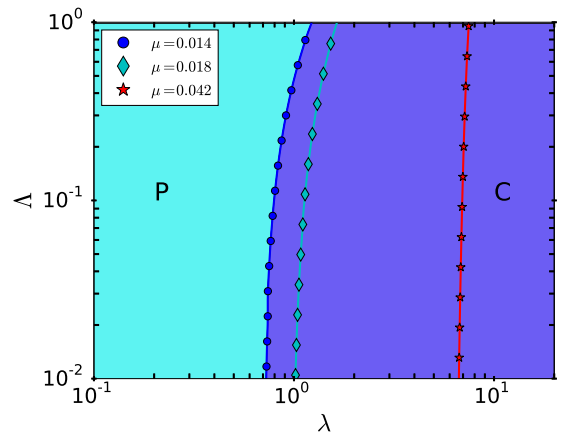


Figure 8.12: Phase diagram in absence of selection function ($f = 1.0$), drawn for $\alpha = 1.0, \gamma = 0.95$. Separatrices are drawn for μ values close to μ^* in Fig. 8.11.

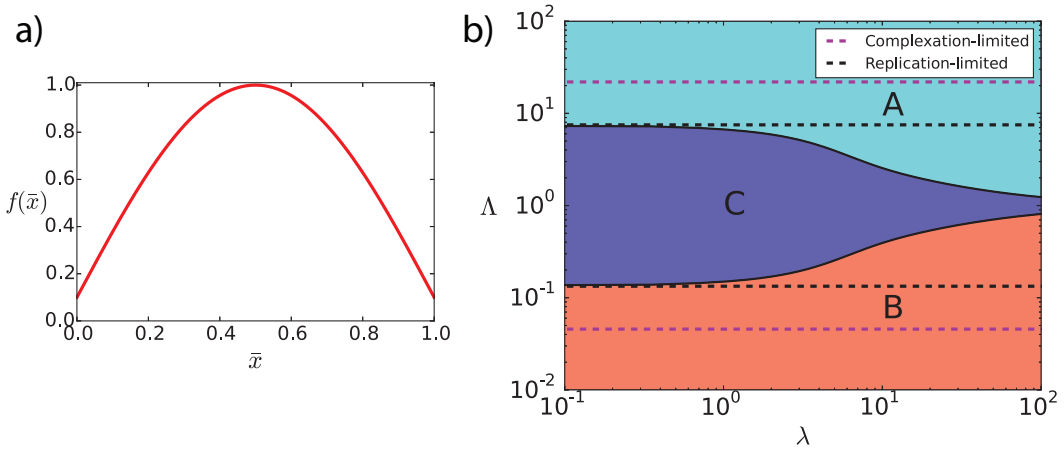


Figure 8.13: Phase diagram for selection function $f(\bar{x}) = 0.1 + 0.9\sin(\pi\bar{x})$, in absence of mutation. Dotted lines correspond to the phase boundary for $\lambda \to 0$, in the presence of strong noise and weak noise (see Sec. 8.4.8).

of the error threshold [27]. Let us introduce the error rate per nucleotide, ε. Then, for a sequence of length L, we have $\alpha - \mu = \alpha(1-\varepsilon)^L$. Since $\varepsilon \ll 1$, it follows from this that $\mu = \alpha\varepsilon L$. When $\alpha \simeq \gamma$, we have $\ln\bar{\Lambda} = \alpha\varepsilon LT$. Using Eq. (8.52), we find that the condition to copy the polymer accurately is

$$L \leq \frac{\ln(s)}{\varepsilon\alpha T}, \tag{8.57}$$

where $s = f(\bar{x})/f(0)$ and $\alpha T/\ln 2$ is the number of generations. This criterium has a form similar to the original error threshold [27], namely

$$L \leq \frac{\ln(s')}{\varepsilon}, \tag{8.58}$$

where $s' = \alpha/\gamma$ represents the selective superiority of the ribozyme. In our model, the equivalent of s' is s which characterizes the compartment selection.

8.3.4 Cooperation

We can also consider a case in which two species, A and B, can cooperate. In such an instance, the framework remains unchanged. To account for cooperation, (e.g. some metabolic task involving both species), we can have the selection function take larger values when both species have similar concentrations. Here, we consider the phase diagram obtained by the symmetric selection function $f(\bar{x}) = 0.1 + 0.9\sin(\pi\bar{x})$, for which we can evaluate the fixed point stability for $x = 0$ and $x = 1$, resulting in Fig. 8.13 By symmetry of the selection function, any stationary system composition $x = x^*$ observed for $\Lambda = \Lambda^*$ would have its mirror composition $x = 1 - x^*$ for $\Lambda = 1/\Lambda^*$. for $\lambda \gg 1$, we recover two asymptotes

$$\Lambda = 1 + \frac{f'(0)}{f(0)\lambda} \quad \Lambda > 1, \tag{8.59}$$

$$\Lambda = 1 + \frac{f'(1)}{f(1)\lambda} \quad \Lambda < 1. \tag{8.60}$$

where we note that $f'(0) = -f'(1)$ for our choice of selection function. For the parasite scenario, we had $f'(0), f'(1) > 0$. For cooperation, we also find a plateau for $\lambda \ll 1$. In this regime, most

compartments start out empty ($p_0 = e^{-\lambda}$), some with one strand ($p_1 = e^{-\lambda}\lambda$, and a small fraction with two ($p_2 = e^{-\lambda}\lambda^2/2$). Since $f(1) = f(0)$, the compartments with initial composition A, AA, B and BB will all propagate with the same success, so any fixed point instability must arise from an AB composition. Evaluating Eq. (8.13) in this regime, we find

$$x' = \frac{\frac{1}{\Lambda+1}\lambda^2 x(1-x) f\left(\frac{1}{\Lambda+1}\right) + (\lambda x + \frac{\lambda^2}{2}x^2) f(1)}{\lambda^2 x(1-x) f\left(\frac{1}{\Lambda+1}\right) + (\lambda + \frac{\lambda^2}{2}(1-2x)) f(1)} \tag{8.61}$$

where we have used $f(0) = f(1)$ and $\bar{x}(2,1) = 1/(\Lambda+1)$. If we now solve $\left.\frac{dx'}{dx}\right|_{x=0} = 1$, we find that

$$\Lambda = \frac{2f\left(\frac{1}{\Lambda+1}\right) - f(1)}{f(1)}. \tag{8.62}$$

Similarly, for $\left.\frac{dx'}{dx}\right|_{x=1} = 1$, we find

$$\frac{1}{\Lambda} = \frac{2f\left(\frac{\Lambda}{\Lambda+1}\right) - f(1)}{f(1)}, \tag{8.63}$$

which we would also find from the symmetry $x, \Lambda \leftrightarrow 1-x, 1/\Lambda$. In Fig.8.13, these plateaus are plotted. In Sec. 8.4.8, we derive how this plateau shifts to much higher values when noise is present in the growth step. The dotted lines in Fig.8.13 illustrate the amplitude of this effect for $\lambda \to 0$.

8.4 Noise in growth

For deterministic growth, given by Eqs. (8.8)-(8.9), fluctuations in the growth rates, denoted α for A molecules and γ for the B molecules, have been neglected. In order to estimate the magnitude and effect of fluctuations in the growth rates, we introduce in the next section a model for noisy replication. In particular, we consider a replication enzyme that stochastically binds to a strand, followed by the stochastic incorporation of L monomers. The model can either have i) a single rate-limiting step or ii) L rate-limiting steps. Case i) corresponds to simple autocatalytic reactions, or the rate-limiting binding of a replication enzyme. Case ii) corresponds to the rate-limiting polymerization of a polymer of length L, via a multistep replication process. For $L = 1$, all these descriptions become equivalent.

Importantly, this model assumes that the replicase, once bound, stays active until completion of the copy of the template. The possibility that the replicase falls off the template before completion of the copy is neglected. Similarly, any effects associated with the interaction of multiple replicases on the same template are neglected. In fact, when the replicase falls off of its template, the copying process is aborted and the shorter chain which has been produced in this way becomes a parasite. We can therefore describe such a process as a mutation using the framework of the previous section. To separate the effects due to mutations and noise clearly, we disregard from now on the possibility of mutations, and we focus in the following on the description of the noise associated with replication. Such a noise can stabilize the ribozyme phase at the expense of coexistence, and the coexistence phase at the expense of the parasite phase. The noise of replication becomes very small when the rate-limiting step is nucleotide incorporation, in which case one can use a deterministic approach. In case of a single rate limiting step, we obtain giant fluctuations.

8.4.1 A minimal model for the replication process

The replication of a polymer strand A by a replicase E can be considered to proceed through two stages. In the first stage, a strand A binds to a replicase E, to form a complex X_0

$$A + E \xrightarrow{\kappa_C} X_0, \tag{8.64}$$

with the rate κ_C.

Subsequently, activated nucleotides X are incorporated in a stepwise fashion to the complementary strand. A complex of E and A with a complementary strand of length n will be denoted by X_n, and the strand grows until the final length L is achieved, such that

$$\mathrm{X_n + X} \xrightarrow{\kappa} \mathrm{X_{n+1}},\ 0 \le n \le L-2 \tag{8.65}$$

$$\mathrm{X_{L-1} + X} \xrightarrow{\kappa} \mathrm{2A}, \tag{8.66}$$

where for simplicity we have assumed the same rate κ for both reactions. Let us denote by t the total time to yield 2A from A, which is the sum of the time associated with the step of complex formation, t_C and with the step of L nucleotide incorporations t_L. We thus have

$$t = t_C + t_L, \tag{8.67}$$

with $t_L = \sum_{i=0}^{L} t_i$ and t_i the time for adding one monomer, which we assumed is distributed according to

$$f(t_i) = \kappa e^{-\kappa t_i}. \tag{8.68}$$

For simplicity, we choose a single value κ for all monomer additions. The time for the formation of the complex, t_C is similarly distributed according to

$$f(t_C) = \kappa_C e^{-\kappa_C t_C}, \tag{8.69}$$

where $\kappa_C = 1/\langle t_C \rangle$.

Let us denote the moment generating function of t_C by $M_C(s)$ and similarly for t_L by $M_L(s)$ with :

$$\begin{aligned} M_C(s) &= \int_0^\infty dt_C \exp(-st_C) f(t_C) \\ &= \frac{\kappa_C}{s+\kappa_C}, \end{aligned} \tag{8.70}$$

$$\begin{aligned} M_L(s) &= \int_0^\infty dt_L \exp(-st_L) f(t_L) = \prod_{i=1}^{L} \Big[\int_0^\infty dt_i \exp(-st_i) f(t_i) \Big] \\ &= \left(\frac{\kappa}{s+\kappa} \right)^L. \end{aligned} \tag{8.71}$$

From M_L one obtains the distribution of replication time $f(t_L)$ by performing an inverse Laplace transform:

$$f(t_L) = \mathscr{L}^{-1}\left[M_L(s)\right] = \frac{\kappa^L t_L^{L-1} e^{-\kappa t_L}}{\Gamma(L)}, \tag{8.72}$$

where $\mathscr{L}^{-1}$ represents the inverse Laplace transform. This equation shows that the replication time distribution of one strand of length L follows a Gamma distribution [56]. For $L = 1$, Eq. (8.72) becomes a simple exponential distribution, which is a memoryless distribution. This distribution describes any process with a single rate-limiting step, such as simple autocatalysis or the binding of the replicase.

For $L > 1$, this distribution has memory and the growth in the number of RNA strands can no longer be described as a simple Markov process. Note that the Gamma distribution is peaked around the mean value of t_L, namely L/κ for $L \gg 1$. In this limit, the replication time has very small fluctuations. This feature has recently been exploited to construct a single-molecule clock, in which the dissociation of a molecular complex occurs after a well-controlled replication time[57].

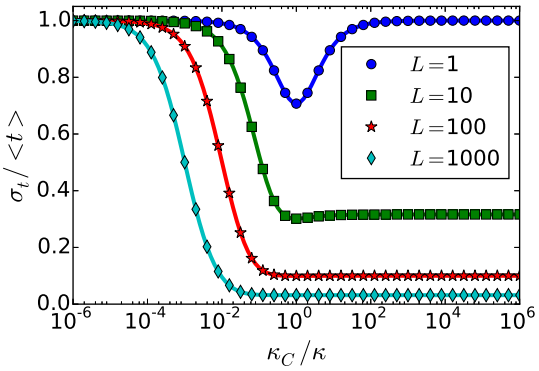


Figure 8.14: Waiting time variability $\sigma_t/\langle t\rangle$ for various polymer lengths L, as a function of the ratio of typical times for replication and complex formation

8.4.2 Coefficient of variation of the replication time

Let us now study the coefficient of variation of the full time t. For the simple replication model, this includes the diffusion of the replicase and the replication step. The generating function of t is clearly $M(s) = M_D(s)M_L(s)$. Thus, the cumulant-generating function defined as $K(s) = \ln M(s)$, yields the two moments of the distribution of t, namely the mean $\langle t\rangle$ and the variance σ_t^2. We have

$$\langle t\rangle = \langle t_C\rangle + \langle t_L\rangle = \frac{1}{\kappa_C} + \frac{L}{\kappa}, \tag{8.73}$$

$$\sigma_t^2 = \sigma_C^2 + \sigma_L^2 = \frac{1}{\kappa_C^2} + \frac{L}{\kappa^2}. \tag{8.74}$$

Thus the coefficient of variation of the replication time, namely $\sigma_t/\langle t\rangle$ is given by

$$\frac{\sigma_t}{\langle t\rangle} = \frac{\sqrt{\frac{1}{\kappa_C^2} + \frac{L}{\kappa^2}}}{\frac{1}{\kappa_C} + \frac{L}{\kappa}}. \tag{8.75}$$

Fig 8.14 shows this quantity as function of the length L and of the ratio of the rates (κ_C/κ).

There are two regimes: on one hand, when $L/\kappa \gg 1/\kappa_C$, the time taken by the replication step dominates over the time for the replicase to diffuse to its target. If in addition $\sigma_L^2 \gg \sigma_C^2$, the coefficient of variation of the time t scales as $1/\sqrt{L}$ and therefore becomes very small for long strands. This power-law regime is indeed visible as plateaus in Fig 8.14 and we will refer to this as the replication-limited regime.

On the other hand, when $1/\kappa_C \gg L/\kappa$, the time to form a complex between the replicase and its template dominates over the replication time. This regime has a large coefficient of variation since $\sigma_t \simeq \langle t\rangle$ as also seen in Fig 8.14. In this regime, the replication time is governed by the simple exponential distribution of Eq. (8.69). A simple autocatalytic reaction is governed by such a distribution, which is also equivalent to a replication-limited situation with $L = 1$. We will refer to this behavior as the diffusion-limited regime.

8.4.3 Phylogenetic noise due to asynchronous growth

In Fig. 8.15, phylogenetic trees are drawn for diffusion-limited and replication-limited growth. In both cases, growth starts from a single parent strand and descendants are depicted as function of their generation.

In this representation, the differences in the two growth regimes become very clear. In the replication-limited regime, generations are synchronized: lineages spread over the same numbers

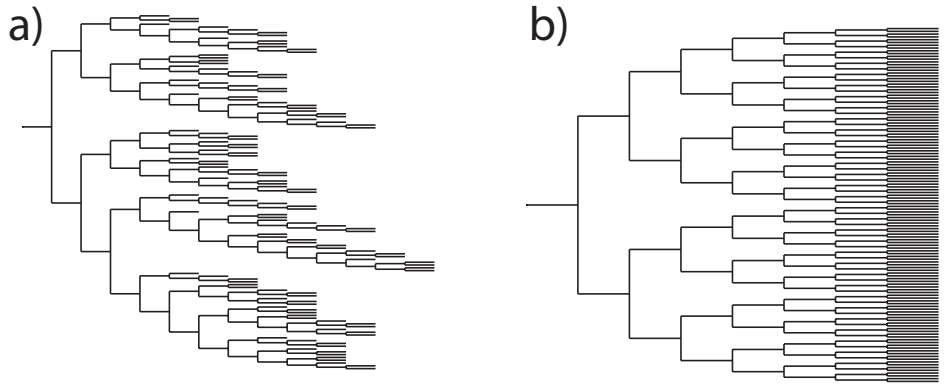


Figure 8.15: Phylogenetic trees, generations representation. a) diffusion-limited regime. b) replication limited regime. The simulation ends when the population size has reached 128. The horizontal axis corresponds to the generation number.

of generations. This happens because noise in replication time is small with respect to the typical replication time: $\sigma_t/\langle t\rangle \ll 1$ (see Eq. (8.75)). Replication slowly desynchronizes by the accumulation of noise over multiple generations. For two independent strains, generations become desynchronized after about $\sqrt{L}$ generations.

In contrast, in the diffusion-limited regime, fluctuations are of the order of the replication time: $\sigma_t/\langle t\rangle = 1$. In this memoryless case, each species is equally likely to perform the next replication event, yielding a desynchronized growth behavior with large gaps in the phylogenetic tree.

These figures have been obtained by simulating the growth of a replicating mixture starting from a single strand. The simulation follows k RNA-enzyme complexes, and for each the variable n_k measures the length of the growing complementary strand. For every nucleotide incorporation event, a strand i is chosen with probability $1/k$, after which its number of nucleotides is updated from n_i to n_i+1. When $n_i+1=L$, we set $n_i=0$, we update k to $k+1$, and then we introduce an extra strand variable n_{k+1} for the new strand. Both the replication-limited regime and the diffusion-limited regime can be modeled using this simulation. In the latter case, we choose $L=1$, which corresponds to exponentially distributed replication times as in Eq. (8.69). This also describes the case of simple autocatalysis.

8.4.4 Noise in population size due to growth

In sec 8.4.2, we have analyzed the noise associated with the replication of a single strand. Ultimately, we wish to quantify the compositional variation of the final population. In order to do so, we turn to the theory of branching processes with variable lifetimes taken randomly from a fixed distribution [58]. As explained in Appendix 10.5, this framework describes theoretically a population that grows exponentially starting from a single individual. In our molecular system, this single individual plays the role of the single molecule present in the initial condition before the replication starts; while the distribution of the lifetimes is the replication time distribution $f(t_L)$ obtained in Eq (8.72).

For $t_L \gg L/\kappa$, we find that the average population (starting from a single individual) $\mu^{(1)}$ scales as $\mu^{(1)}(t)=\mu^* e^{\alpha t}$, with a growth rate $\alpha \simeq \kappa \ln(2)/L$. The coefficient of variation of the population size $\sigma^{(1)}/\mu^{(1)}$ is

$$\frac{\sigma^{(1)}}{\mu^{(1)}} \approx \frac{\sqrt{2}\ln(2)}{\sqrt{L}}. \tag{8.76}$$

The renewal theory on which these results are based, can be generalized to the case that there are n individuals in the initial condition as shown in 10.5.1. The full solution is found by treating

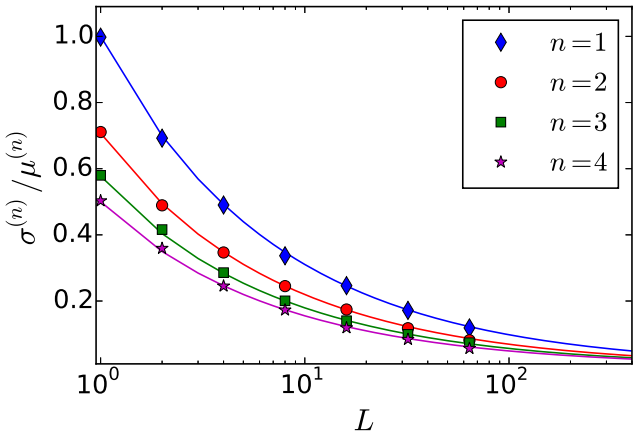


Figure 8.16: Coefficient of variation of the population size N as function of the initial population size n. The results have been averaged over 2000 runs. The solid lines represent the theoretical prediction: $1/\sqrt{nL}$.

the n initial molecules as n independent subpopulations, which all start at size 1 and follow the branching process described above and in 10.5. In that case, each subpopulation now has a mean $\mu^{(1)} = \mu^{(n)}/n$ and a standard deviation $\sigma^{(1)} \approx \mu^{(1)}/\sqrt{L}$. This then allows to write

$$\sigma^{(n)} \approx \sqrt{n}\sigma^{(1)} = \frac{\mu^{(1)}}{\sqrt{nL}}. \tag{8.77}$$

We show in Fig. 8.16 that the corresponding coefficient of variation, $\sigma^{(n)}/\mu^{(n)}$, agrees well with simulations of the branching process. The 2000 simulation runs were stopped after a time t^* such that $\langle N(t^*)\rangle \simeq 5000$.

8.4.5 Giant fluctuations in logistic growth of competing species

The problem of two species competing for the same resources has been studied in the literature and offers a complementary perspective on the role of noise in a growing population, which has been studied in the previous section. Let us consider two such species, which typically start with a few individuals and then grow according to logistic noise. As shown in Ref. [54], when the carrying capacity is reached, the number of each species is subject to giant fluctuations (the coefficient of variation is of the order of unity) when the two species have similar growth rates. In the terminology introduced in previous section, this model applies to the diffusion-limited regime ($L \to 1$, simple autocatalysis), where a Markov description of the population dynamics is applicable.

Keeping the notations of the first section, we denote by n the initial number of molecules, which splits into m ribozymes (or A autocatalysts) and y parasites (or B autocatalysts), and by N the final number of molecules in the compartment. In the neutral case ($\alpha = \gamma$), the moments of the number of ribozymes $\bar{m}$ are found to be [54] :

$$\langle \bar{m} \rangle = N\frac{m}{n}, \tag{8.78}$$

$$\sigma_{\bar{m}} = \sqrt{\frac{my}{n^2}\frac{N(N-n)}{n+1}}, \tag{8.79}$$

with again $y = n - m$. Since N remains fixed, $\sigma_{\bar{x}} = \sigma_{\bar{m}}/N$. This means that

$$\sigma_{\bar{x}} \approx \frac{1}{n}\sqrt{\frac{my}{n}} \tag{8.80}$$

for $N \gg n$, which means that the noise in the composition depends primarily on the number of individuals in the initial condition. Let us denote $s = \alpha/\gamma - 1 \ll 1$, with $s \ll 1$ and $\rho = \ln(N/n)$. In Ref. [54], it was shown that

$$\frac{\sigma_{\bar{m}}}{\langle \bar{m} \rangle} = \sqrt{\frac{y}{m(n+1)}} \left(1 - \frac{\rho s n (m+1)}{(n+1)(n+2)} \right) \tag{8.81}$$

In general, the dynamics of the composition has a large variability for: (i) small compartments ($n \sim O(1)$), (ii) mixed compartments ($m, y > 0$), and for $m \approx y$, (iii) comparable growth rates ($s \to 0$).

Such a coefficient of variation is asymptotically constant on long times and the constant only depends on the initial number of molecules. A similar scaling for the coefficient of variation holds in a number of other physical situations, such as for the fluctuations in the number of protein filaments formed in small volumes [59].

8.4.6 Noise for co-encapsulated growing populations

Let us now apply the results of the section 8.4.4 to analyze the effect of the growth noise on our transient compartmentalization dynamics. Let us assume that the length of the ribozymes is L_α and that of the parasites L_γ. For experimental values of these parameters we refer the reader to Table 8.1. In Sec. 8.2, we have defined m, y to be the initial number of ribozymes and parasites and $\bar{m}, \bar{y}$ to be the final mean number of ribozymes and parasites at the end of the growth phase in a given compartment. Using Eqs. (8.76)-(8.77), we obtain

$$\begin{aligned} \frac{\sigma_{\bar{m}}}{\langle \bar{m} \rangle} &\simeq \frac{1}{\sqrt{L_\alpha m}}, \\ \frac{\sigma_{\bar{y}}}{\langle \bar{y} \rangle} &\simeq \frac{1}{\sqrt{L_\gamma y}}. \end{aligned} \tag{8.82}$$

Since the ribozyme fraction $\bar{x}$ at the end of the exponential phase is given by $\bar{x}(n,m) = \bar{m}/N$ and $N \simeq n_{Q\beta}$, the noise on $\bar{x}(n,m)$ takes the following form :

$$\begin{aligned} \sigma_{\bar{x}} &= \sqrt{\left(\frac{\partial \bar{x}}{\partial \bar{m}}\right)^2 \sigma_{\bar{m}}^2 + \left(\frac{\partial \bar{x}}{\partial \bar{y}}\right)^2 \sigma_{\bar{y}}^2}, \\ &\simeq \sqrt{\left(\frac{\bar{y}}{N^2}\right)^2 \frac{\bar{m}^2}{m L_\alpha} + \left(\frac{-\bar{m}}{N^2}\right)^2 \frac{\bar{y}^2}{y L_\gamma}}, \\ &\simeq \bar{x}(1-\bar{x}) \sqrt{\left(\frac{1}{m L_\alpha} + \frac{1}{y L_\gamma}\right)}, \end{aligned} \tag{8.83}$$

where we have used Eq. (8.77) with $\mu_{\bar{m}} = \bar{m}, \mu_{\bar{y}} = \bar{y}$. The factor $\bar{x}(1-\bar{x})$ is largest for $\bar{x} = 1/2$ and vanishes for pure parasite and pure ribozyme compartments, which means that this noise can be neglected when $\Lambda \gg 1$ or $\Lambda \ll 1$. Note that if we choose $\alpha = \gamma$ (and thus $\bar{x} = m/n$), and $L_\alpha = L_\gamma = 1$, Eq. (8.83) becomes

$$\sigma_{\bar{x}} \simeq \frac{1}{n} \sqrt{\frac{my}{n}} \tag{8.84}$$

which is consistent with Eq. (8.80) which was found using a different formalism[54].

Eqs. (8.81) and (8.83) point to an interesting trade-off : the synchronization of growth rates comes at the cost of greater compositional noise. To have a stable coexistence, growth rates should not diverge too much. However, this also implies giant fluctuations in final composition. In the

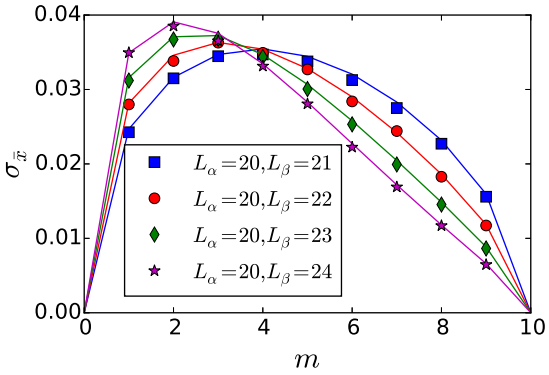


Figure 8.17: Standard deviation of the ribozyme fraction, $\sigma_{\bar{x}}$, as predicted from simulations (symbols), and compared with predictions from Eq. (8.83) (solid lines). For each initial composition (m,n), 10000 simulations were performed until a time t^* such that $\langle N(t^*)\rangle \simeq 5000$ and by choosing $\alpha/\gamma = L_\gamma/L_\alpha$.

presence of strong selection, noise will generate many unviable compositions, lowering the overall survival of compartments [29].

This reduction in survival is particularly detrimental if a compartment splits into only two daughter compartments [29]. To prevent extinction, at least half of the daughters should, on average, survive. This puts a strong constraint on more advanced selection mechanisms, such as the Stochastic corrector, for which growth noise can rapidly become more catastrophic than replication errors. It will be instructive to refer to this phenomenon as a noise catastrophe.

In transient compartmentalization, only a much smaller fraction of the order of λ/N compartments need to survive. The RNA experiments[34] are indeed performed in this regime since $N \approx O(10^6), \lambda = O(1)$). As such, a noise catastrophe needs to be considerably more severe before the molecules replicating in transient compartments go extinct.

By having multiple rate-limiting steps ($L > 1$), compositional noise is reduced. In this sense, polymerization on a template as considered here is inherently functional: the noise suppression it permits can increase the average compartment fitness. Noise suppression also increases evolvability, by giving the system access to more efficient mechanisms of heritability.

Using the parameters of Table 8.1 and (8.76), we can quantify the level of noise in the number of ribozymes or parasites in the RNA droplet experiment [34]. We find from this table that the ribozyme size was $L = 362$, and that the experiment should be in the replication-limited regime because the diffusion time scale should be approximately over $2 \cdot 10^4$ times smaller than replication times of the order of 10s. The noise in composition should be maximal when we start with one ribozyme and one parasite of equal length, and with $\alpha = \gamma$, which on average gives $\bar{x} = 1/2$. Consequently, the noise in composition is at most $\sigma_{\bar{x}} \approx 0.02$. In such a case, our deterministic approach used in [4] is applicable.

8.4.7 Phase diagram in the presence of weak noise

The growth equations given by Eqs. (8.8) and (8.9) are deterministic in nature, which means that a given initial condition (n,m) yields a unique final composition $\bar{x}(n,m)$. In contrast to that in a stochastic approach, a given n and m lead to many different trajectories, which means that $\bar{x}(n,m)$ is a random variable with a probability distribution $p(\bar{x}(n,m))$. Consequently, the ribozyme fraction

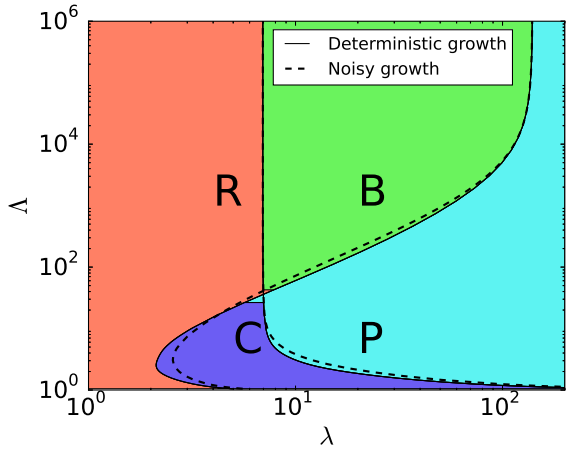


Figure 8.18: Phase diagram for ribozyme-parasite scenario in presence of noise given by Eq. (8.83), for $L_\alpha = L_\gamma = 3$.

after one round is

$$x' = \frac{\sum_{n,m} \int_0^1 d\bar{x}\bar{x}(n,m)p(\bar{x}(n,m))f(\bar{x})P_\lambda(n,x,m)}{\sum_{n,m} \int_0^1 d\bar{x}p(\bar{x}(n,m))f(\bar{x})P_\lambda(n,x,m)}. \tag{8.85}$$

This expression is computationally demanding to evaluate for $\lambda \gg 1$, but it can be simplified significantly in the weak noise limit.

In order to construct a phase diagram in this limit, we simplify Eq. (8.85), by considering $p(\bar{x}(n,m)) \approx \mathcal{N}(\bar{x}, \sigma_{\bar{x}})$, where $\mathcal{N}$ denotes a normal distribution with mean $\bar{x}$ and standard deviation defined by Eq. (8.83). From Eq. (8.83) we expect the effect of noise to be largest when λ, L and Λ are close to 1 (if $\Lambda \gg 1, \bar{x} \to 0$). In Fig. 8.18, the original phase diagram from Ref. [4] is shown together with the modified phase boundaries (dotted lines) due to the presence of Gaussian noise using Eq. (8.85) for the case that $L_\alpha = L_\gamma = 3$.

Given that the amplitude of this type of noise should rapidly diminish for larger L, and that $L \sim O(100)$ in the experiment, we expect our ribozyme-parasite scenario to be well-described by a deterministic dynamics. We also see that the noise stabilizes the pure ribozyme phase (R) with respect to the coexistence phase (C) because in the presence of noise, the R region has grown at the expense of the C region. Similarly, the noise stabilizes the coexistence region (C) against the parasite region (P).

8.4.8 Example: noise-induced cooperation

For the cooperation scenario, Eq. (8.85) can also be evaluated exactly in the limit $\lambda \ll 1$. As an example of the effect of growth noise, we consider the complexation-limited regime for a cooperation scenario, with a symmetric selection function $f = 0.1 + 0.9\sin(\pi\bar{x})$. In particular, we are interested in the plateau regions for $\lambda \ll 1$. We start by considering Eq. (8.85), which obeys

$$x' = \frac{\int_0^1 d\bar{x}\bar{x}\lambda x(1-x)f(\bar{x}) + (x + \frac{\lambda}{2}x^2)f(1)}{\int_0^1 d\bar{x}\lambda x(1-x)f(\bar{x}) + (1 + \frac{\lambda}{2}(1-2x))f(1)}. \tag{8.86}$$

Note that the AB compartment is the only compartment with a distribution of compositions final compositions $p(\bar{x})$, as the A, AA, B and BB compartments have a pure composition from the start.

Using $\left.\frac{dx'}{dx}\right|_{x=0} = 1$, and the procedure outlined in Sec. 8.3.1, we find that the new upper plateau must obey

$$\int_0^1 d\bar{x}\, \bar{x} f(\bar{x}) p(\bar{x}) = \frac{f(1)}{2}. \tag{8.87}$$

To solve this equation, we use the method outlined in [54] to numerically find $p(barx)$ for a given relative rate of growth $r = \alpha/\beta$ and fixed final population N. From this, we find an r for which the equation holds, and we subsequently find Λ from N and r. By symmetry, we can immediately find the lower plateau at $1/\Lambda$.

Both plateaus are plotted in the original phase diagram in Fig. 8.13.

8.5 Parasites and time allocation in replication

So far, we considered a parasite to be detrimental to the collective by lowering the overall fitness. A type of parasite often explicitly considered in origins of life acts by hijacking a replication enzyme or ribozyme to make more copies of itself, thus directing all resources to its own production. This then leads to an overall lower overall production of the replication machinery and ultimately extinction. It is this type of parasite that comes to prominence above Eigen's original error threshold.

Such problems have in the past [27, 25, 28] been analyzed using ordinary differential equations. Its multilevel selection in transient compartments was recently treated in Ref[6]. In these ODE approaches, replication is treated as instantaneous.

When replication is rate-limiting, however, we need to consider that it is a composition of sub-steps in which the strand-replicase complex incorporates monomers. These steps may individually be instantaneous, but not collectively. In this replication-limited regime, the process has a waiting time that becomes increasingly peaked as the strand length increases and during this time, the replicase is occupied. This subtlety is lost in instantaneous ODE approximations.

In this section, we will explore how this seemingly small specification drastically affects the dynamics. The capacity to occupy the replication machinery inhibits the production of competitors. Here, we will show that this feature makes parasites far more detrimental to replicase survival than one would infer from the typical ODEs used to study the problem.

8.5.1 Model Setup

We consider a replicase R which, upon meeting another replicase R, makes a new copy of the replicase, such that for the event $m \to m+1$ (where m is a discrete number of replicases) there is a transition rate

$$W_{m\to m+1} = \alpha m(m-1), \tag{8.88}$$

where α is a rate constant (we will absorb any volume dependence in rate constants). Let us also consider a parasite P, which, upon meeting a replicase R, is copied, giving 2P. We then have

$$W_{y\to y+1} = \gamma m y, \tag{8.89}$$

where γ is a again a rate constant and y denotes the discrete number of parasites. In Fig. 8.19 these processes are drawn schematically. Note that here, unlike the population dynamics in Refs.[4, 5], parasites cannot grow when $m = 0$, they require the presence of a replicase.

Deterministic approach

For $m, y \gg 1$, we can write the set of ODEs, corresponding to a mass-action approach.

$$\dot{m} = \alpha m(m-1), \tag{8.90}$$

$$\dot{y} = \gamma m y. \tag{8.91}$$

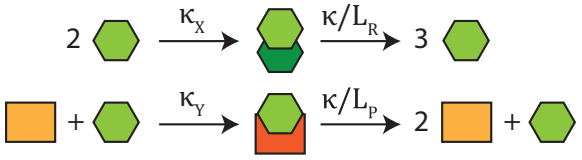


Figure 8.19: Schematic picture for replication of replicases (green hexagons) and parasites (yellow squares), which first form complexes X, Y, and subsequently incorporate L_R (resp. L_P) monomers, with a per monomer rate of κ. After a time κ/L, a copy has been formed. For a detailed description on the monomer-incorporation level, see Sec. 8.4.4.

supposing $m(0) > 1$. Integrating (8.90) from 0 to t yields the solution

$$\ln\left(\frac{(1-m(t))m(0)}{m(t)(1-m(0))}\right) = \alpha t \tag{8.92}$$

which can be rewritten to yield

$$m(t) = \frac{1}{1-\left(1-\frac{1}{m(0)}\right)e^{\alpha t}} \tag{8.93}$$

which blows up at finite time

$$\tau = \frac{1}{\alpha}\ln\left(1-\frac{1}{m(0)}\right). \tag{8.94}$$

Substituting this solution for (8.91) and supposing $y(0) > 0$ we find

$$\ln\left(\frac{y(t)}{y(0)}\right) = \gamma\left[t - \frac{1}{\alpha}\ln\left(m(0)-(m(0)-1)e^{\alpha t}\right)\right], \tag{8.95}$$

which simplifies to

$$y(t) = y(0)e^{\gamma t}m(0)^{\alpha/\gamma}\left(\frac{1}{1-\left(1-\frac{1}{m(0)}\right)e^{\alpha t}}\right)^{\gamma/\alpha}. \tag{8.96}$$

As can be expected, both parasite and replicase blow up in finite time, the ratio γ/α quantifies the relative degree of divergence.

8.5.2 Stochastic approach: variation of blow-up time

Let us again consider the process of m replicases, that copy upon collision (that is, collision is rate-limiting, nucleotide incorporation is rapid). Suppose they meet in memoryless fashion, with an exponentially distributed waiting time t_m for the next collision, described by

$$p(t_m) = \alpha m(m-1)\exp(\alpha m(m-1)t) \tag{8.97}$$

On average, to move from m to $m+1$ replicases, we then have

$$\langle t_m\rangle = \frac{1}{\alpha m(m-1)}. \tag{8.98}$$

let $\langle t \rangle$ be the sum of these average waiting times, then

$$\langle t \rangle = \sum_{m=m(0)}^{\infty} \frac{1}{\alpha m(m-1)} = \frac{1}{\alpha(m(0)-1)}. \tag{8.99}$$

where the final result follows from $\frac{1}{m(m-1)} = \frac{1}{m-1} - \frac{1}{m}$ whose sum yields a telescoping series. For $m(0) = 2$, we have $\langle t \rangle = 1/\alpha$, which means that on average, a blowup occurs after $2\langle t_2 \rangle$ has elapsed. The variance the exponential waiting time obeys

$$\sigma_{t_m}^2 = \frac{1}{\alpha^2 m^2 (m-1)^2}. \tag{8.100}$$

The variance of the total waiting time is then

$$\sigma_t^2 = \sum_{m=m(0)}^{\infty} \frac{1}{\alpha^2 m^2 (m-1)^2}. \tag{8.101}$$

For $m(0) = 2$, this yields $\sigma_t^2/\langle t \rangle^2 = (\pi^2 - 9)/3 \approx 0.29$. leading to a coeficient of variation $\sigma_t/\langle t \rangle \approx 0.54$. This large variation is mainly due to the first step, as $\sigma_{t_2}^2/\langle t_2 \rangle^2 = 0.25$.

A dominant effect of the first replication steps is also observed for exponential growth, which was discussed in the last section and in Ref.[54]. We thus expect that a compositional behavior that may deviate strongly from the deterministic Eq. (8.96). While the deterministic solution may be finite up till a time τ (Eq. (8.94)), a significant fraction of stochastic trajecteries will already blow up well before this time.

Clearly, we cannot expect such an accelerating encounter process to remain the rate-limiting step. In the next section we will consider resource limitations as a solution to curb this effect.

8.5.3 Stochastic approach: population composition

When considering the composition of a replicase-parasite population, we generally try to avoid cases with a blowup. In principle this can be done by considering a short enough time frame such that Eqs. (8.90) (8.91) remain good approximations, but as shown in the previous section, this approximation breaks down very fast and the actual growth dynamics can show considerable variance in replication times.

A more common way to avoid blowups is to consider limitations, e.g. limiting resources. One way this can be accounted for is by introducing a carrying capacity N in the growth rates, e.g.

$$W_{m \to m+1} = \alpha m(m-1)\frac{N-m-y}{N}, \tag{8.102}$$

$$W_{y \to y+1} = \gamma m y \frac{N-m-y}{N}, \tag{8.103}$$

which yields our former equations (8.88) and (8.89) for $N \gg m, y$.

Let us now consider a master-equation formalism, in which we use a population size $n = m + y$ in favor of time, to describe a population which grows to its carrying capacity where $n = N$. We can then write

$$P_{m,n+1} = a_n^{m-1} P_{m-1,n} + b_n^{n-m-1} P_{m,n}. \tag{8.104}$$

Where a_n^m, denotes the probability to go from a state (n,m) to $(n+1, m+1)$. Since the only alternative replication course would be to go to $(n+1, m)$, whose probability is denoted by b_n^m, we have by total probability

$$a_n^m + b_n^m = 1. \tag{8.105}$$

From the transition rates defined by eqs. (8.102) and (8.103), we have

$$a_n^m = \frac{W(n,m \to n+1,m)}{W(n,m \to n+1,m) + W(n,m \to n,m+1)} \tag{8.106}$$

From eqs. (8.102) and (8.103), we arrive at

$$a_n^m = \frac{r(m-1)}{n-1+(r-1)(m-1)}, \tag{8.107}$$

$$b_n^m = \frac{n-m-1}{n-1+(r-1)(m-1)}. \tag{8.108}$$

where $r = \alpha/\gamma$. If we now define $m' = m-1$ and $n' = y+m'$, our master equation exactly coincides with the master equation proposed for memoryless growth of two species by B. Houchmandzadeh [54]. Such a master equation was shown to exhibit giant fluctuations in the composition of the population (see Sec. 8.4.5 for a complementary perspective). Consequently, the regime in which an ODE approach can be considered (diffusion-limited growth) is inherently noisy. This aspect of the composition dynamics is not captured the ODE approach.

While a carrying capacity curbs the blowup in finite time, its use in an ODE framework may be insufficient to capture all important details. One important detail is that a typical replicase forms a complex with the species it replicates. During this replication process, it is unavailable for other species. This is not reflected in differential equations, which treat such a process as an enzymatic process with a single rate-limiting step. Such an approximation is no longer appropriate in the replication-limited growth regime. We will see that in this regime parasites become considerably more detrimental.

8.5.4 Replication-limited growth

In the diffusion-limited regime, the hyperbolic growth accelerates so rapidly, that it takes a pair of particles $\langle t \rangle = \frac{2}{\kappa_C}$ to reach a blow-up on average, which is twice the typical time of the first encounter. Evidently, such a blow-up in finite time is unphysical: we cannot accelerate all processes indefinitely nor supply the material for such growth. The problem stems from our approximations: a variety of timescales can be fast when $m(0) = 2$ and may initially be neglected. However, every subsequent encounter event occurs $(m+2)/m$ times faster than the former, such that the encounter rate rapidly outgrows the rate of all other relevant processes, after which it is no longer the rate-limiting step.

In the former section, we discussed that resource depletion through a carrying capacity could be introduced as such a limitation. Another way to account for this, is by considering that the replication process consists of several steps, of which only the complexation step should accelerate. For the minimal model for polymer replication discussed in Sec. 8.4, this corresponds to the replication-limited regime.

We can extend our stochastic model for independently elongating replicators, to a model for growing parasites and replicators. In the first step, a replicase R or a parasite P complexes with another replicase R:

$$\mathrm{R} + \mathrm{R} \xrightarrow{\kappa_X} \mathrm{X}_0, \qquad \mathrm{P} + \mathrm{R} \xrightarrow{\kappa_Y} \mathrm{Y}_0 \tag{8.109}$$

with the rates κ_X, κ_Y .

Subsequently, activated nucleotides X are incorporated in a stepwise fashion to the complemen-

tary strand and the strand grows until the final length L_R (resp. L_P) is attained, such that

$$X_n + X \xrightarrow{\kappa} X_{n+1}, \quad 0 \leq n \leq L_R - 2 \tag{8.110}$$

$$X_{L-1} + X \xrightarrow{\kappa} 3R, \tag{8.111}$$

$$Y_n + X \xrightarrow{\kappa} Y_{n+1}, \quad 0 \leq n \leq L_P - 2 \tag{8.112}$$

$$Y_{L'-1} + X \xrightarrow{\kappa} 2P + R, \tag{8.113}$$

An important difference with the deterministic model, is that replicases are occupied while they perform their replication, which prevents them from copying other species during that time. In the deterministic model and the the diffusion-limited regime, this memory effect is absent. The model is reminiscent of the branching process discussed in Appendix 10.5. However, there are now two growing and interacting populations which affect each other's growth, which means we can no longer apply the renewal theorem. This means that exact approaches from queueing theory cannot be applied: the queues are not independent. A discussion on population statistics will, at this point, be confined to numerical simulations of the process. However, the phenomenology can be well understood in terms of simple analytical arguments, which will be discussed in detail. We hope this problem may motivate new approaches to study interacting queues.

8.5.5 Complexation routes

To understand the fate of a replicator-parasite population, we must consider what kind of compositions unbound species can occur, and how they respond to replication events. Let us denote N_P, N_R the number of free parasites and replicases. If $\kappa_X, \kappa_Y \gg \kappa/L_R, \kappa/L_P$, replicases will directly bind to any available target upon their release. Let us now consider a number of population compositions, $\{N_R, N_P\}$, and their response to the finishing of replication for X, Y complexes.

- $\{N_R, N_P\} = \{0,0\}$

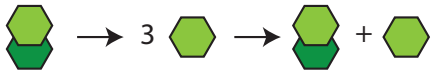


Figure 8.20: Starting from X, we can only form X + R.

$X \to 3R$: $\{N_R, N_P\} = \{3,0\}$, which immediately leads to a new replicase complex X, $N_X + 1$, and $\{N_R, N_P\} = \{3,0\}$

Figure 8.21: Starting from Y, we can only form Y + P.

$Y \to 2P + R$: $\{N_R, N_P\} = \{1,2\}$, which can only be immediately followed by forming parasite complex Y, $N_Y + 1$ and $\{N_R, N_P\} = \{0,1\}$.

- $\{N_R, N_P\} = \{1,0\}$

$X \to 3R$: $\{N_R, N_P\} = \{4,0\}$, which immediately forms two new replicase complexes X, $N_X + 2$, and $\{N_R, N_P\} = \{0,0\}$

$Y \to 2P + R$: $\{N_R, N_P\} = \{2,2\}$. New complexations can then take two routes:
i) Formation of a single replicase complex X, $N_X + 1$, $\{N_R, N_P\} = \{0,2\}$

Figure 8.22: Starting from X + R, we can only form 2X.

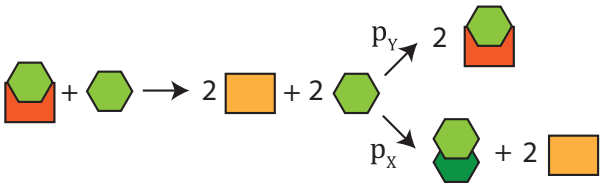


Figure 8.23: Starting from Y + R, we either form 2Y or X + 2P.

ii) Formation of two parasite complexes Y, $N_Y + 2$ $\{N_R, N_P\} = \{0,0\}$ The relative probabilities of these processes are given by the statistics of the formation a the first complex

$$p_X = \frac{\kappa_X}{\kappa_X + 2\kappa_Y}, \tag{8.114}$$

$$p_Y = \frac{2\kappa_Y}{\kappa_X + 2\kappa_Y}. \tag{8.115}$$

•$\{N_R, N_P\} = \{0,1\}$

Figure 8.24: Starting from X + P, we can only form X + Y.

X → 3R: $\{N_R, N_P\} = \{3,1\}$ leads to an X and Y complex, $N_X + 1, N_Y + 1$, $\{N'_R, N'_P\} = \{0,0\}$

Y → 2P + R: $\{N_R, N_P\} = \{1,3\}$ leads to a Y complex $N_Y + 1$, $\{N'_R, N'_P\} = \{0,2\}$

•$\{N_R, N_P\} = \{0,2\}$,

X → 3R: $\{N_R, N_P\} = \{3,2\}$ leads either to:
i) an X and Y complex, $N_X + 1, N_Y + 1$, $\{N'_R, N'_P\} = \{0,1\}$, or
ii) two Y complexes, $N_Y + 2$ $\{N'_R, N'_P\} = \{1,0\}$, with a relative probability

$$p_{X,Y} = \frac{(\kappa_Y)^2 + 2\kappa_X\kappa_Y}{(\kappa_X + \kappa_Y)^2} \tag{8.116}$$

$$p_{2Y} = \frac{\kappa_Y^2}{(\kappa_X + \kappa_Y)^2}. \tag{8.117}$$

Y → 2P + R: $\{N_R, N_P\} = \{1,4\}$ forms a new Y complex $N_Y + 1$, $\{N'_R, N'_P\} = \{0,3\}$.

•$\{N_R, N_P\} = \{0,k\}, \quad k \geq 3$

Figure 8.25: Starting from Y + P, we can only form Y + 2P.

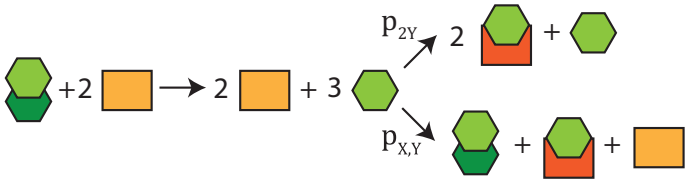


Figure 8.26: Starting from X + 2P, we can either form 2Y + R or Y + X + P

X → 3R: $\{N_R, N_P\} = \{3, k\}$ either leads to:
i) an X and Y complex, $N_X + 1$, $N_Y + 1$, $\{N'_R, N'_P\} = \{0, k-1\}$, or
ii) three Y complexes $N_Y + 3$, $\{N'_R, N'_P\} = \{0, k-3\}$. With relative probabilities

$$\begin{aligned} p_{X,Y} &= \frac{2\kappa_X}{2\kappa_X + k\kappa_Y} + \frac{k\kappa_Y}{2\kappa_X + k\kappa_Y}\frac{\kappa_X}{\kappa_X + (k-1)\kappa_Y} = \frac{2(\kappa'_C)^2 + (3k-2)\kappa'_C\kappa_C}{2(\kappa'_C)^2 + (3k-2)\kappa'_C\kappa_C + k(k-1)\kappa_C^2}, \\ p_{3Y} &= \frac{k\kappa_C}{2\kappa'_C + k\kappa_C}\frac{(k-1)\kappa_C}{\kappa'_C + (k-1)\kappa_C} = \frac{k(k-1)\kappa_C^2}{2(\kappa'_C)^2 + (3k-2)\kappa'_C\kappa_C + k(k-1)\kappa_C^2}. \end{aligned} \tag{8.118}$$

These complexation routes highlight how parasites are inherently competent at frustrating the replication of R. In the case where complexation is strongly biased towards X, $\kappa_X \gg \kappa_Y$ (but still $\kappa_X, \kappa_Y \gg \kappa/L_R, \kappa/L_P$), this analysis simplifies greatly: after any event, replicase complexes are formed until $N_R \leq 1$, after which parasite complexes are formed.

Growth regimes

As a first illustration, consider the case where we start with $N_P = 1$, and a large population of replication complexes $N_X \gg N_P$ that are largely desynchronized. If X complexes finish in rapid succession and assemble new X complexes, we can write a net process

$$2X \xrightarrow{\kappa/L_R} 3X. \tag{8.119}$$

with κ/L the average waiting time. If a small population of parasites starts to grow, it will equally well encounter its necessary replicases in rapid succession ($N_X \gg N_P$ and replicases are produced in odd amounts), and perform a net process

$$Y \xrightarrow{\kappa/L_P} 2Y. \tag{8.120}$$

Neglecting the frustration of growth of R by the parasite on short timescales, we thus find doubling times

$$\tau_X = \frac{L_R \ln 2}{\kappa \ln 3/2} \approx 1.71\frac{L_R}{\kappa}, \tag{8.121}$$

$$\tau_Y = \frac{L_P}{\kappa}, \tag{8.122}$$

which show that parasites are inherently advantaged, even if they are equally long $L_R = L_P$. This is because a replicase occupies another replicase to replicate itself, meaning only half of replicases can be replicating at a given time. Parasites do not hinder each other in this way. Here, replicase growth can only match that of parasites that are considerably longer: $L_P > (\ln(2)/\ln(3/2))L$.

Figure 8.27: Starting from Y + 2P, we can only form Y + 3P.

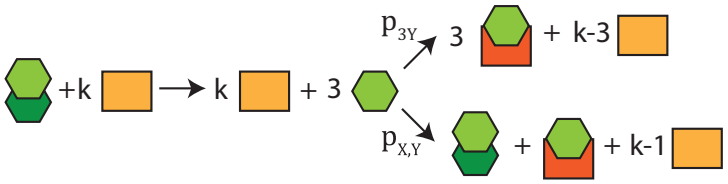


Figure 8.28: Starting from X+k P, we either form 3Y + (k−3) P or X + Y + (k−1) P.

When parasites become abundant enough, replication events of R will no longer manage to convert all their products to new X complexes, because the third R molecule will rapidly complex with a parasite. If $\kappa_X \gg N_P \kappa_Y$, the first two R molecules will still proceed to form a complex X, but the X population will no longer grow. In this regime, R molecules are produced at a constant rate, since the net reaction becomes

$$X \xrightarrow{\kappa/L_R} X + R. \tag{8.123}$$

As parasites do not produce R themselves, supply and production of R becomes the limiting factor. Their growth is linear in supplied R and this supply itself also grows linearly due to Eq. (8.123). Thus, the growth becomes quadratic instead of exponential.

If instead, $\kappa_X = \kappa_Y$, abundant parasites will occupy all replicase molecules R to form themselves via Y. Any remaining X will have little chance of yielding more X complexes after replication, as the corresponding probability $p_{X,Y}$ (Eq. (8.118)) of such an event is now inversely proportional to the free parasite population

$$p_{X,Y} = \frac{3}{(N_P + 2)}. \tag{8.124}$$

With no further production of R, the parasite population grows linearly.

8.5.6 Stochastic simulations

When we start with small numbers of parasites and replicases, the arguments from Sec. 8.5.5 do not capture all the phenomenology. To illustrate this, the system of equations (8.111)-(8.113) was simulated using Gillespie's algorithm[60] on the level of monomer incorporation, in the limit $\kappa_X, \kappa_Y \gg \kappa/L_R, \kappa/L_P$. Complexations are performed near instantly at the end of every replication, with complexation probabilities as detailed in Sec. 8.5.5.

For an initial condition $\{N_P^0, N_R^0\} = \{1,2\}$, $L_P = L_R = 10$, 100000 simulations were performed, with a fixed end time of $\tau_{sim} = \ln(10^6/(N_P^0 + N_R^0))/\kappa \ln(2)$. The resulting distributions show a considerable amount of fine structure, even in regions where sampling is high.

In all populations, parasites far outnumber replicases, which becomes more evident when we plot the replicase fraction $\bar{x}$, defined as

$$\bar{x} = \frac{N_R + N_Y + 2N_X}{N_R + N_P + 2N_Y + 2N_X}. \tag{8.125}$$

This fraction counts all the fully assembled replicases and parasites, in free and complexed form. As seen in Fig. 8.32, parasites are over 10 times more abundant than replicators.

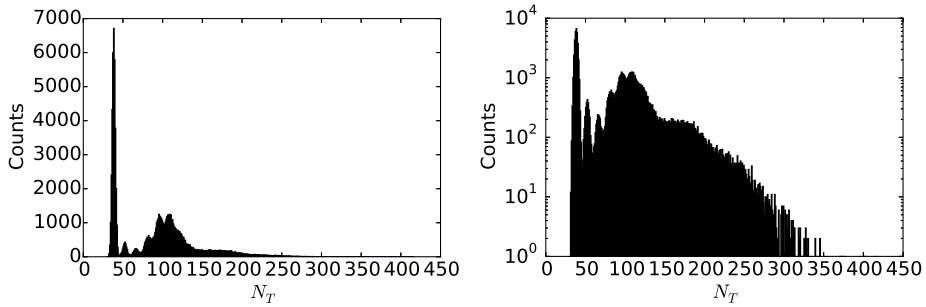


Figure 8.29: Histogram for total population size left) linear right) logarithmic, after $\tau_{sim} = \ln(10^6/(N_{\mathrm{P}}^0 + N_{\mathrm{R}}^0))/\kappa\ln(2)$ has elapsed, for $\{N_{\mathrm{P}}^0, N_{\mathrm{R}}^0\} = \{1,2\}$, $L_{\mathrm{R}} = L_P = 10$, $\kappa_{\mathrm{X}}, \kappa_{\mathrm{Y}} \gg \kappa/L_{\mathrm{R}}, \kappa/L_{\mathrm{P}}$.

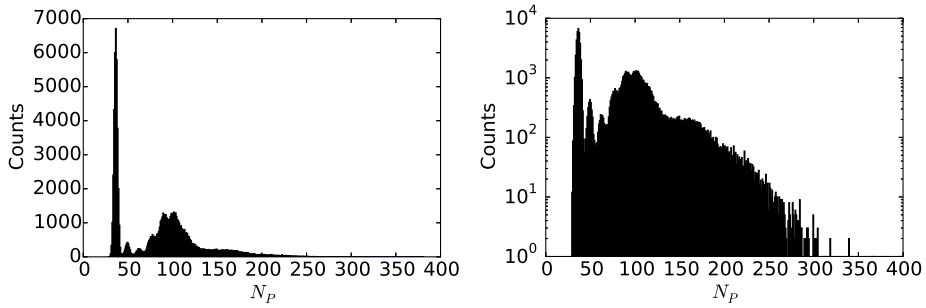


Figure 8.30: Histogram for number of parasites in the population size left) linear right) logarithmic, after $\tau_{sim} = \ln(10^6/(N_{\mathrm{P}}^0 + N_{\mathrm{R}}^0))/\kappa\ln(2)$ has elapsed, for $\{N_{\mathrm{P}}^0, N_{\mathrm{R}}^0\} = \{1,2\}$, $L_{\mathrm{R}} = L_{\mathrm{P}} = 10$, $\kappa_{\mathrm{X}}, \kappa_{\mathrm{Y}} \gg \kappa/L_{\mathrm{R}}, \kappa/L_{\mathrm{P}}$.

The Complexation Catastrophe and the Error Catastrophe

In Sec. 8.4.6, it was argued that complexation and polymerization could be beneficial, by approaching deterministic waiting times and thereby stabilizing synchronization. Here, we see that such contributions can also be very beneficial to certain types of parasites often considered in the literature[27]. It will be instructive to see how robust the conclusions from ODE models in the literature are to this modification.

The complexation catastrophe is similar in spirit to an error catastrophe: parasites take over. There is a fundamental difference however: a complexation catastrophe, once set in motion, cannot be overcome by accurate replication. Such catastrophes warn us that maintaining replicating polymers is not trivial and that our prebiotic scenarios are constrained, similar to how tradeoffs constrain evolution. One strategy to address such catastrophes is to start off with a catastrophe-prone self-replicating species, and subsequently consider selection mechanisms that help it overcome catastrophes, such as transient compartments, a stochastic corrector [25] or spatial diffusion[28].

Another strategy would be to see if we can start off with self-replicating species that are less prone to catastrophes from the start. To attenuate the complexation catastrophe, we can consider a replicase R and an inactive complementary species R$'$, such that we can perform

$$\mathrm{R} + \mathrm{R}' \to 2\mathrm{R} + \mathrm{R}', \qquad 2\mathrm{R} \to 2\mathrm{R} + \mathrm{R}', \tag{8.126}$$

then templates R$'$ can hang around, and newly released replicases R have a chance to bind to more templates instead of being sequestered directly by parasites. Given the template nature of DNA and RNA, this model would be a more correct starting point. By opposing two templates to a single-species model, a division of labor between the different templates realized, an effect also

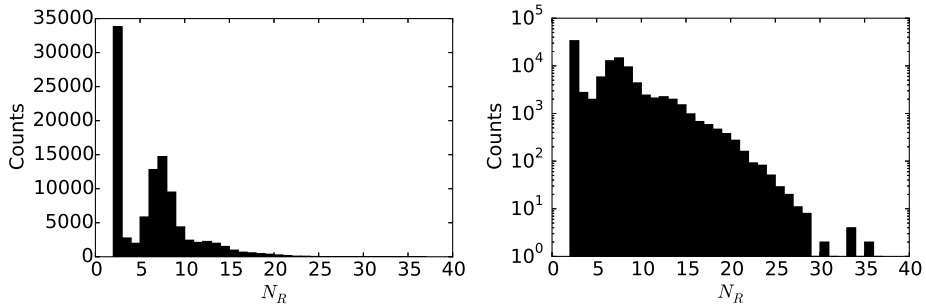


Figure 8.31: Histogram for number of replicases in the population size left) linear right) logarithmic, after $\tau_{sim} = \ln(10^6/(N_{\mathrm{P}}^0 + N_{\mathrm{R}}^0))/\kappa\ln(2)$ has elapsed, for $\{N_{\mathrm{P}}^0, N_{\mathrm{R}}^0\} = \{1,2\}$, $L_{\mathrm{R}} = L_{\mathrm{P}} = 10$, $\kappa_{\mathrm{X}}, \kappa_{\mathrm{Y}} \gg \kappa/L_{\mathrm{R}}, \kappa/L_{\mathrm{P}}$.

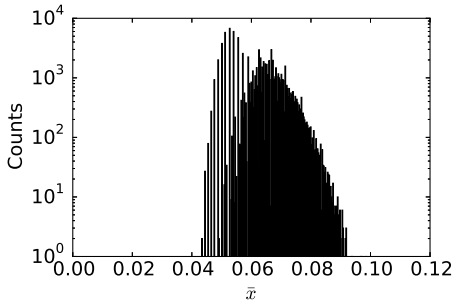


Figure 8.32: Histogram of composition $\bar{x}$, after $\tau_{sim} = \ln(10^6/(N_{\mathrm{P}}^0 + N_{\mathrm{R}}^0))/\kappa\ln(2)$ has elapsed, for $\{N_{\mathrm{P}}^0, N_{\mathrm{R}}^0\} = \{1,2\}$, $L_{\mathrm{R}} = L_P = 10$, $\kappa_{\mathrm{X}}, \kappa_{\mathrm{Y}} \gg \kappa/L_{\mathrm{R}}, \kappa/L_{\mathrm{P}}$.

considered by[61, 28]. Takeuchi and Hogeweg considered that going from a pure RNA system to a DNA-RNA system would stabilize a system against parasites. Many of these works are formulated in a hypothetical prebiotic context where considerable sophistication is present (e.g. an RNA world with genetic machinery and enzymes). However, catastrophes are expected to be relevant before such sophistication came into play.

It will be interesting to see how systems chemistry provides new pathways to attack these questions. Self-assembly, dynamic chemistry and self-sorting are but some of the features that may allow us to come up with more gradual scenarios. Perhaps these chemical collectives went through intermediate stages that were not prone to the catastrophes that have been invoked for an early replicase. In this sense, catastrophes can be interpreted as constraints on evolutionary trajectories derived from thought experiments, similar to how thermodynamics was further constrained by thought experiments, e.g. on mixing [62] and information[63].

Bibliography

Articles

[2] Stefan Schuster et al. “Cooperation and cheating in microbial exoenzyme production – Theoretical analysis for biotechnological applications”. In: *Biotechnol. J.* 5 (2010), pp. 751–758.

[3] Simon Alberti. “The wisdom of crowds : regulating cell function through condensed states of living matter”. In: *J. Cell Sci.* 130 (2017), pp. 2789–2796.

[4] A Blokhuis et al. “Selection Dynamics in Transient Compartmentalization”. In: *Phys. Rev. Lett.* 120.15 (Apr. 2018), p. 158101.

[5] Alex Blokhuis et al. “The generality of transient compartmentalization and its associated error thresholds”. In: *BioRxiv* (2018). arXiv: 521211 [10.1101].

[6] Gabin Laurent, Luca Peliti, and David Lacoste. “Survival of self-replicating molecules under transient compartmentalization with natural selection”. In: *Life* (2019), accepted. BiorXiv755355.

[7] Ö Kartal et al. “Carbohydrate-active enzymes exemplify entropic principles in metabolism”. In: *Mol. Syst. Biol.* 1 (7), p. 542.

[9] Frank R Moss et al. “Ladderane phospholipids form a densely packed membrane with normal hydrazine and anomalously low proton / hydroxide permeability”. In: *Proc. Natl. Acad. Sci. U.S.A.* 115.37 (2018), pp. 9098–9103.

[10] Stephen Hunt and David B Layzell. “Gas exchange of legume nodules and the regulation of nitrogenase activity”. In: *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44 (1993), pp. 483–511.

[11] Daniel J. Aneshansley and Thomas Eisner. “Biochemistry at 100 C : Explosive Secretory Discharge of Bombardier Beetles (Brachinus)”. In: *Science (80-.).* 165 (1969), pp. 61–63.

[13] Jean-paul Douliez et al. “Artificial Cell Models Catanionic Coacervate Droplets as a Surfactant-Based Membrane-Free Protocell Model Angewandte”. In: *Angew. Chem. Int. Ed.* 56 (2017), pp. 13689–13693.

[14] Björn Drobot et al. “Compartmentalised RNA catalysis in membrane-free coacervate protocells”. In: *Nat. Commun.* 9 (2018), p. 3643.

[15] Christof B Mast et al. “Escalation of polymerization in a thermal gradient”. In: *Proc. Natl. Acad. Sci. U.S.A.* 110.20 (2013), pp. 8030–5.

[16] Christopher M Dobson et al. “Atmospheric aerosols as prebiotic chemical reactors”. In: *Proc. Natl. Acad. Sci. U.S.A.* 22.97 (2000), pp. 11864–11868.

[18] Marc Tessera. “Is pre-Darwinian evolution plausible ?” In: *Biol. Direct* 13 (2018), pp. 1–18.

[19] D Lancet, R Zidovetzki, and O Markovitch. “Systems protobiology: origin of life in lipid catalytic networks”. In: *J.R.Soc.Interface* 15 (2018), p. 20180159.

[20] Helen Greenwood Hansma. “Possible origin of life between mica sheets: Does life imitate mica?” In: *J. Biomol. Struct. Dyn.* 31.8 (2013), pp. 888–895.

[21] Eugene V Koonin and William Martin. “On the origin of genomes and cells within inorganic compartments”. In: *Trends Genet.* 21.12 (2005).

[22] Günter Wächtershäuser. “Before Enzymes and Templates: Theory of Surface Metabolism”. In: *Microbiol. Rev.* 52.4 (1988), pp. 452–484.

[24] E Szathmáry and L Demeter. “Group selection of early replicators and the origin of life”. In: *J. Theor. Biol.* 128.4 (1987), pp. 463–486.

[26] S Spiegelman et al. “The synthesis of a self-propagating infectious nucleic acid with a purified enzyme”. In: *Proc Natl Acad Sci U S A* 54 (1965), pp. 919–927.

[27] M Eigen. “Self-organization of matter and the evolution of biological macromolecules”. In: *Naturwissenschaften* 58.10 (1971), pp. 465–523.

[28] Nobuto Takeuchi and Paulien Hogeweg. “Evolutionary dynamics of RNA-like replicator systems: A bioinformatic approach to the origin of life”. In: *Phys. Life Rev.* 9.3 (2012), pp. 219–263.

[29] D Grey, V Hutson, and E Szathmáry. “A re-examination of the stochastic corrector model”. In: *Proc. R. Soc. Lond. B* 262 (1995), pp. 29–35.

[30] David Sloan Wilson. “A Theory of Group Selection”. In: *Proc. Natl. Acad. Sci. USA* 72.1 (1975), pp. 143–146.

[31] Daniel Segré et al. “Graded autocatalysis replication domain (GARD): kinetic analysis of self-replication in mutually catalytic sets”. In: *Orig. Life Evol. Biosph.* 28 (1998), pp. 501–514.

[32] Daniel Segré, Dafna Ben-eli, and Doron Lancet. “Compositional genomes : Prebiotic information transfer in mutually catalytic noncovalent assemblies”. In: *Proc. Natl. Acad. Sci. U.S.A.* 97.8 (2000), pp. 4112–4117.

[33] E Szathmáry. “The origin of replicators and reproducors”. In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361 (2006), pp. 1761–1776.

[34] Shigeyoshi Matsumura et al. “Transient compartmentalization of RNA replicators prevents extinction due to parasites”. In: *Science (80-.).* 354.6317 (2016), pp. 1293–1296.

[35] Caitlin E Cornell et al. “Prebiotic amino acids bind to and stabilize prebiotic fatty acid membranes”. In: *Proc. Natl. Acad. Sci. U.S.A.* 116.35 (2019), pp. 17239–17244.

[36] William Stillwell. “Facilitated Diffusion of Amino Acids across Bimolecular Lipid Membranes as a Model for Selective Accumulation of Amino Acids in a Primordial Protocell”. In: *BioSystems* 8 (1976), pp. 111–117.

[37] A.S. Zadorin and Y. Rondelez. “Natural selection in compartmentalized environment with reshuffling”. In: *J. Math. Biol.* (2019), pp. 1–54.

[38] P L Luisi, P Walde, and T Oberholzer. “Lipid vesicles as possible intermediates in the origin of life”. In: *Curr. Op. Coll. Int. Sci.* 4.1 (1999), pp. 33–39.

[39] Moritz Kreysing et al. “Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length”. In: *Nat. Chem.* 7 (2015), pp. 203–208.

[40] Philipp Baaske et al. “Extreme accumulation of nucleotides in simulated hydrothermal pore systems”. In: *Proc. Natl. Acad. Sci. U.S.A.* 104.22 (2007), pp. 9346–9351.

[41] E V Koonin and W Martin. “On the origin of genomes and cells within inorganic compartments”. In: *Trends Genet.* 21.12 (2005), pp. 647–654.

[42] D Zwicker et al. “Growth and division of active droplets provides a model for protocells”. In: *Nat. Phys.* 13.4 (2017), pp. 408–413.

[43] Bruce Damer and David Deamer. “Coupled Phases and Combinatorial Selection in Fluctuating Hydrothermal Pools: A Scenario to Guide Experimental Approaches to the Origin of Cellular Life”. In: *Life* 5.1 (2015), pp. 872–887.

[44] Taro Furubayashi and Norikazu Ichihashi. “Sustainability of a Compartmentalized Host-Parasite Replicator System under Periodic Washout-Mixing Cycles”. In: *Life* 8.3 (2018), p. 10.

[45] John S Chuang, Olivier Rivoire, and Stanislas Leibler. “Simpson’s Paradox in a Synthetic Microbial System”. In: *Science (80-.).* 323.5911 (2009), pp. 272–275.

[46] Andrew S Tupper and Paul G Higgs. “Error tresholds for RNA replication in the presence of both point mutations and premature termination errors”. In: *J. Theor. Biol.* 428 (2017), pp. 34–42.

[47] Ye Eun Kim and Paul G Higgs. “Co-operation between polymerases and Nucleotide Synthethases in the RNA world”. In: *PLoS Comput. Biol.* 12.11 (2016), e1005161.

[48] Lukas Geyrhofer and Naama Brenner. “Coexistence and cooperation in structured habitats”. In: *bioRxiv* (2018).

[50] Roger Wick, Peter Walde, and Pier Luigi Luisi. “Light Microscopic Investigations of the Autocatalytic Self-Reproduction of Giant Vesicles”. In: *J. Am. Chem. Soc.* 117 (1995), pp. 1435–1436.

[51] William Stillwell. “Facilitated diffusion as a method for selective accumulation of materials from the primordial oceans by a lipid-vesicle protocell”. In: *Orig. Life* 10 (1980), pp. 277–292.

[52] Tamás Czárán, Balázs Könny, and Eörs Szathmáry. “Metabolically coupled replicator systems : Overview of an RNA-world model concept of prebiotic evolution on mineral surfaces”. In: *J. Theor. Biol.* 381 (2015), pp. 39–54.

[53] György Károlyi et al. “Chaotic flow : The physics of species coexistence”. In: *Proc. Natl. Acad. Sci. U.S.A.* 97.25 (2000), pp. 13661–13665.

[54] Bahram Houchmandzadeh. “Giant fluctuations in logistic growth of two species competing for limited resources”. In: *Phys. Rev. E* 98.4 (Oct. 2018), p. 42118.

[55] A Dramé-Maigné et al. “Quantifying the performance of high-throughput directed evolution protocols”. In: *ArXiv* (2018). arXiv: 1811.05288.

[56] Daniel L Floyd, Stephen C Harrison, and Antoine M van Oijen. “Analysis of Kinetic Intermediates in Single-Particle Dwell-Time Distributions”. In: *Biophys. J.* 99.2 (2010), pp. 360–366.

[57] Alexander Johnson-Buck and William M Shih. “Single-Molecule Clocks Controlled by Serial Chemical Reactions”. In: *Nano Lett.* 17.12 (2017), pp. 7940–7944.

[59] Thomas C T Michaels, Alexander J Dear, and Tuomas P J Knowles. “Stochastic calculus of protein filament formation under spatial confinement”. In: *New J. Phys.* 20.5 (2018), p. 55007.

[60] D T Gillespie. “Exact Stochastic Simulation of Coupled Chemical-reactions”. In: *J. Phys. Chem* 81 (1977), p. 2340.

[61] Nobuto Takeuchi, Paulien Hogeweg, and Kunihiko Kaneko. “The origin of a primordial genome through spontaneous symmetry breaking”. In: *Nat. Commun.* 8 (2017), p. 250.

[63] Leo Szilard. “Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen”. In: *Zeitschrift für Phys.* 53 (1929), pp. 11–12.

Books

[1] Pierre Cornelis and Simon C. Andrews, eds. *Iron Uptake and Homeostatis in Microogranisms*. Poole, UK: Caister Academic Press, 2010, p. 292.

[8] Thomas Nogrady, Robert L. Wallace, and Terry W. Snell. *Rotifera, Vol. 1: Biology, Ecology and Systematics*. SPB Academic Publishing, 1993.

[12] A I Oparin. *Origin of Life*. Dover, 1952.

[17] Harold J. Morowitz. *Beginnings of cellular life: Metabolism recapitulates biogenesis*. London: Yale University Press, 1992.

[25] John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford: Oxford University Press, 1995.

[49] Pier Luigi Luisi. *The Emergence of Life: From Chemical Origins to Synthetic Biology*. 2nd ed. Cambridge, UK: Cambridge University Press, 2016, p. 478.

[58] S Karlin and H M Taylor. *A first course in stochastic processes*. New York: Academic Press, 1975.

[62] J. Willard Gibbs. *On the equilibrium of heterogeneous substances*. New Haven: The Academy, 1874.

9. A new mechanism-based scenario

In this final chapter, a new prebiotic scenario is proposed that recapitulates concepts introduced in the previous chapters. The word 'scenario' is employed in the sense 'a postulated sequence or development of events'. In recent years, the word scenario is increasingly being used in the sense 'a specific geological environment' (e.g. ponds, hydrothermal vents, deep rock compartments), with a particular finding showing support for a particular environment. The present discussion makes no explicit argument for any environment in particular, it makes an argument for a novel set of mechanisms and transitions.

The outlined scenario is meant to be thought-provoking, we do not pretend that our present work can give an accurate picture of abiogenesis. Our scenario attempts to explain a transition from prebiotic chemistry to Darwinian evolution in terms of physical-chemical mechanisms. In particular, it describes i) prebiotic chemistry undergoing multienvironment chemical evolution, ii) multilevel selection through cross-catalytic formation of synthetic environments iii) a gradual transition towards Darwinian evolution due to coencapsulation and cohabitation of synthetic environments and their subsequent dynamics. The present scenario uses a collection of mechanisms introduced and explored in our previous chapters. Whether these transitions actually happened and whether they employed our particular set of mechanisms remains a speculation, like all prebiotic scenarios.

The starting point of these speculations, however, is different from other scenarios, which are inspired by extant biochemistry. In this scenario, only physical-chemical processes are described, some of which have not been proposed before. It is only in stage iii) that we postulate the gradual introduction of cell division. Genetic machinery is never invoked and can be a very late invention. Heredity and metabolism are an inherent consequence of generalized autocatalysis coupled with reservoirs.

Gradualism

If chemical evolution is anything like modern evolution, we must consider its gradual nature. This lesson is well understood when considering the postulated evolution of organs: a primitive eye may come without cones and rods, thus limiting any detailed vision. However, such an organ may already follow a circadian rhythm [1]. While its building instructions may be less convoluted, such a primitive organ may still render itself indispensable.

A similar idea comes from Cairns-Smith[2], who notes that an arcade of bricks is inherently unstable when it is built up brick by brick, only the finished arcade is stable. In this final structure, we do not see the pile of sand or some equivalent scaffold that permitted the construction in the first place, as they are now removed.

In this vein we may also think of key molecules and structures in life. If genetic polymers have such diverse and important tasks today, this can be thought to be the product of extensive chemical evolution. This then begs the question: what did these molecules do before their promotion to such important tasks? Many similar problems arise on a metabolic level, e.g. at what point can it be functional to have phosphorylation?

As has been argued throughout this manuscript, molecules can do much simpler things that render them indispensable: they may serve as donors or sinks of electrons and protons, be a fuel or activating agent, perform allocatalysis or autocatalysis, mediate intercompartment exchange, sequester and store compounds and synthesize new environments. To understand a molecule's incorporation in nature's repertoire, we may not only consider what it does today, but also what it did 'before'. Wächtershäusers [3] scenario provides an illustrative example: phosphates are the bread and butter of surface attachment and hydrophilicity.

We cannot and do not claim to understand where the extant biomolecules came from, but it is hoped we have to provided instructive new ways to think about the problem.

9.1 A physical chemist's dream

In analogy with a molecular biochemist's dream[4], we formulate here a physical chemist's dream. It is an idealized, dreamt up trajectory of chemical evolution. Since we assume that most of prebiotic chemistry is not preserved in extant biochemistry, our molecular actors are ipso facto anonymous.

9.1.1 I: Towards an organized soup

Our story starts with a reactive chemical soup, externally resupplied with chemicals and thus maintained far from equilibrium. The reactive soup was in contact with a heterogeneous environment with transport barriers, leading automatically to dynamic self-sorting. The reactive soup transformed this heterogeneous environment to a collection of chemically distinct subsoups in microenvironments. As a consequence of transport barriers and local chemistries, subsoups selectively started to exchange compounds locally.

It is at this stage that autocatalytic chemical evolution make its first move: through the nucleation of autocatalytic cycles, permanent modifications in local chemistries and transport are introduced, which also modify the local self-sorting.

The seeds of these new chemistries spread, as certain autocatalytic species or precursors reach new microenvironments further away. Most of this happens quite locally, due to diffusion. But advection and convection make sure that molecules venture to more distant lands. Small species also evaporate and spread over long distances, performing 'prebiotic pollination' of distant collectives of soups.

New autocatalytic cycles lead to diversification in the globally available chemistries, while local chemistries remain sufficiently specific* to remain functional. Some of the local† chemical networks become more than the sum of their parts. Reactions start to proceed with degrees of improvement beyond the single-reaction limit (e.g. by DKR, proofreading), and dynamically

*There is a clear tradeoff between diversity and efficiency, due to side reactions.

†here, local pertains either to a single compartment or compartment and their not too distant neighbors, coupled by exchange processes

interconverting compounds start to form a variety of species, of which some are used for further autocatalysis, error-correction and the assembly of higher-order structures.

All this leads to more diverse microenvironments, which in turn start the formation of new 'synthetic' environments e.g. new minerals by precipitation, new vesicles by surfactant production, new microcompartments by local dissolution of a solid substrate, new coacervates by polyelectrolyte synthesis, fibrils, sheets and other superstructures by the self-assembly of small molecules, complex coacervates formed by polyelectrolytes and micelles, or by polyelectrolyte block copolymers. The coupling of recombination and intercompartment exchange promotes the formation of long oligomers and polymers, which form key building blocks for further synthetic compartments.

The new environments provide the occasion for a further diversification and new autocatalytic cycles. By redirecting resources and promoting new pathways, other autocatalytic cycles lose their viability. An evolutionary process of multicompartment autocatalysis weaves the reactive soup in an intricate web of chemical and spatial currents.

9.1.2 II: Towards organized units of multilevel selection

With the introduction of multicompartment systems higher order selection mechanisms can make their entry. The viability of different networks critically hinges on the continued occupation of compartments. For synthetic ones, their continued production is of key importance as well. Transient compartmentalization and protocell division (Ch.9) are some examples of the new forms of multilevel selection that come into play. These selection forces steer compartmentalized chemistries in new directions.

As our organized soup starts to create synthetic environments, a new level of selection is introduced. We must now both consider i) the viability of the chemical network, and ii) the viability of the synthetic environment.

Both viabilities are determined by the internal chemistry and the surrounding microenvironments, and various synthetic environments compete for a common feedstock to produce their building blocks. These building blocks are formed from an amalgamation of self-sorted chemistries in different environments, that would quickly become incompatible in a single pot.

A large diversity of synthetic environments can harbor various chemistries, furnishing building blocks for different compartments. In this sense, autocatalysis can be envisaged on the compartment level, by a succession of environments making, assembling and recruiting each other's building blocks. For example, the hydrophobic core of a micelle provides an excellent environment to perform (and catalyze, in the presence of acid) condensation chemistry[5]. Thereby, a polymer is synthesized that upon entering the water phase is quickly protonated, forming a polyelectrolyte, which in turn can dramatically enhance the local concentration of monomers [6] with respect to the bulk phase. The charged polymer attracts amphiphile counterions, whose local concentration templates their future incorporation in a micelle.

Cross-catalytic environments predisposed to be in each other's vicinity had a great advantage, as they could rapidly exchange material with minimal losses. Rudimentary forms of coencapsulation, cohabitation, attachment, anchoring etc. started to make their marks, forming sticky multicompartment assemblies. These assemblies would spread locally by their continuing coassembly. On longer distances, mechanical fracturing coupled with flow would displace part of the assembly. These assemblies would show considerable spatial heterogeneity and undergo further autocatalytic evolution, e.g. when encountering new environments. The multicompartment assembly would in turn also be fertile ground for parasitic chemistries, compartments and assemblies, which do not aid in spreading the initial multicompartment assembly, but consume its resources and products for their own propagation. Their presence formed an important source for new pathways, innovations and adaptations. It is in this stage of expansion, multilevel competition and innovation that evolutionary arms races[7] can start to take off. Multicompartment assemblies that propagate more easily gain a

distinct advantage. Rudimentary forms of 'adaptative chemistry' make their entrance: dynamic combinatorial chemistries allow to explore a large set of options and flexibly redirect resources to those that are consumed in a given local context (e.g. by the local autocatalytic networks or the assembly of local structures).

This arms race, on a large scale, favored the capacity to adapt. Assemblies that could explore more solutions would find ways to outpace and consume others. Systems displaying increased capacities of evolution had an inherent advantage in this context, leading to an indirect selection for evolvability itself. Gradually, dynamic chemistry in multicompartment autocatalysis started to acquire favorable genetic characteristics. Transitioning to dividing multicompartment assemblies (and ultimately cells) provided a powerful means of dispersal and parallelized the search for new adaptations. This ultimately led to the modern dogma of biology.

9.1.3 Problems addressed, problems not addressed

Let us first reiterate, that the aforementioned is a scenario. It is inevitably short-sighted and limited by the imagination of the author. It addresses perceived shortcomings of other scenarios, thereby hoping to provide a marginal improvement on which future scenarios can be built.

Let us here consider some of the points we have tried to address

- Chemical Evolution: our framework for autocatalytic chemical evolution encompasses all forms of autocatalysis (to our knowledge) that are described in the OOL literature. The most important critiques on works on chemical evolution are: a lack of chemical realism and a lack of open-ended evolution (is there enough autocatalysis to do this?). We share this sentiment: successive autocatalysis of one particular type of chemistry on the scale of a single compartment does not seem sufficiently open-ended.
  Our framework (Ch.5, Ch.6) considers every chemical reaction and transport process as a potential building block for a allocatalytic and autocatalytic cycles. In our scenario, more layers of selection are introduced through multilevel selection (e.g. compartment division, transient compartmentalization, Ch.8), and higher order feedback, like the synthesis of new environments. We consider network motifs (App. 10.3) beyond the single reaction that may yield improvements, such as proofreading and dynamic kinetic resolution, which in turn may feed into selection on an ecological level.
  The problem of 'open-ended chemical evolution' has thereby been considered from many angles and with a focus on chemical realism. We are far from finished with this endeavor, new functional networks and selection mechanisms are in the making, along with experiments to back them up. We have not proven that our approach provides enough open-endedness, but the combined force of all these perspectives may progressively bring that goal closer.
- Gradualism: Our tentative scenario proceeds by small steps, we do not invoke extremely rare events. While their absence is not a logical necessity, we should perceive it as problematic to include them if we can avoid it.
- Generality: Not a single reactant was made explicit, the scenario is mechanistic. Instead of considering how we obtain a modern biomolecule, we look at the general structural motifs that chemical reaction networks can exhibit. This is in line with the ideas expressed by Krishnamurthy[8]: .. *the experimental and conceptual approaches for understanding or demonstrating the emergence of life's processes on Earth could benefit from deliberately ignoring the goal to exactly replicate extant life's process in a prebiotic setting*. We probably do not know which molecules constituted prebiotic chemistry and pre-life. By looking at the structure of chemical networks and chemical evolution, our approach aims to provide an insight in spite of this ignorance.

Let us now make some aspects explicit, that better scenarios can improve on

- Chemical Candidates and environments: There is factual information about the composition

and chemistry of the early earth. Geological and geochemical insights can provide important inputs to constrain the scenario, its chemistry and its synthetic environments.

- Transitions: We have mentioned transitions, such as one from dynamic combinatorial chemistry towards rudimentary genetics. This is a pure speculation, for which we have not provided a more detailed trajectory, despite our suggestion for a gradual transition. The 'emergence of a code' remains one of the mysteries in the field, that is still lacking a convincing mechanism that is in accord with chemistry, physics and evolution.
- Selection mechanisms: The compartment division and transient compartmentalization are only two multilevel selection mechanisms. There may be an abundance of other multilevel selection mechanisms. These mechanisms need to be found and thoroughly explored.
- Functional networks: Chemical network evolution has classically been considered in terms of autocatalysis and the rate of autocatalytic cycles. Here, we have drawn attention to some other things chemical networks can do, such as proofreading, dynamic kinetic resolution and use dynamic combinatorial chemistry to find new solutions. Still, this only scratches the surface of how chemical networks may affect chemical evolution in an ecological context, where interactions become key. A systematic understanding of functional chemical networks must emerge, to provide an appreciation of what molecules can and cannot do on the systems level.

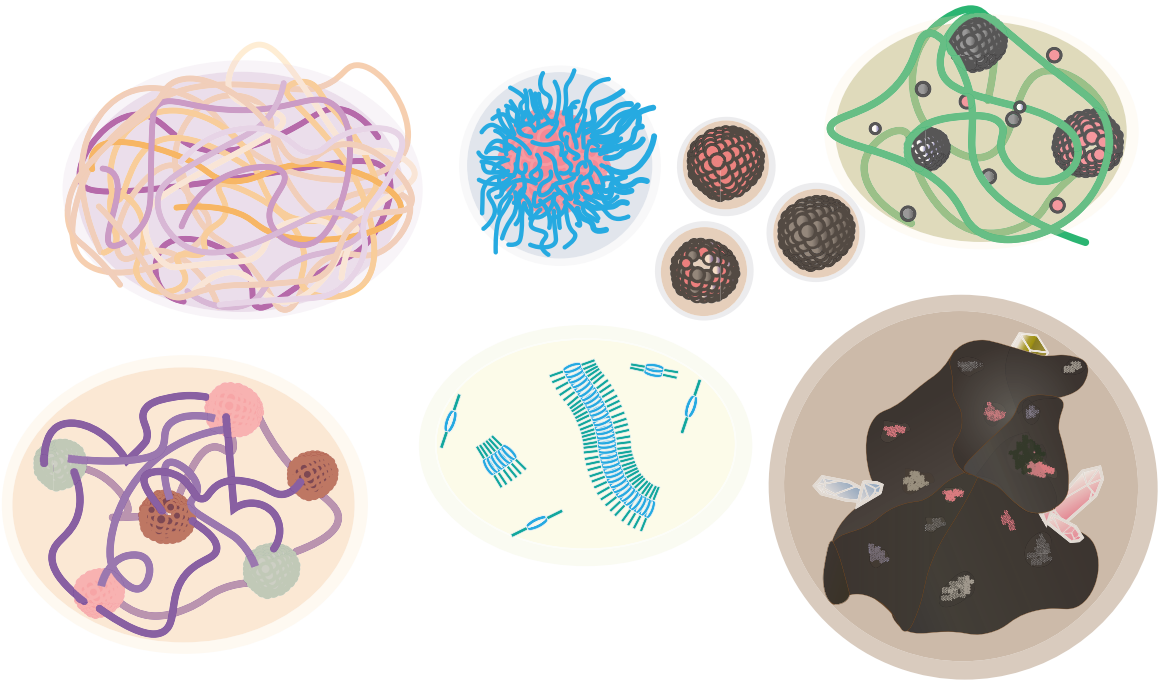


Figure 9.1: Some examples of synthetic environments: a coacervate formed by oppositely charged polyelectrolytes, a core-shell particle made from diblock copolymers, micelles, a micelle-polyelectrolyte complex coacervate, a complex coacervate with local phase separation (due to triblock-copolymers), supramolecular fibrils and patches of modified mineral surface (surface metabolisms)

Closing remarks: towards a systems-level view

The mindset of systems chemistry is to study emergent phenomena due to interacting molecular components at different levels. It is in this spirit that we must consider this scenario, and in this sense that we must assess: i) what a molecule can do, ii) what collectives of molecules can do and

iii) what collectives of compartments and their contents can do[‡].

Insights that address these fundamental questions will be a valuable advance in its own right. How much the systems-level approach will teach us about the Origins of Life is unknown. Perhaps our molecular ancestors are too far gone, obscured by molecular evolution[8]. This possibility should not discourage us from looking for them: even if we may not find them, we may learn some of the fundamental principles that brought us where we are today. Let us end with Oparin's words, which echo this optimism:

Сейчас, когда подробно изучена внутренняя организация живых существ, есть все основания считать, что мы сможем, рано или поздно, искусственно воспроизвести эту организацию и тем непосредственно показать, что жизнь есть не что иное, как особая форма существования материи.

— А.И. Опарин

Now that the internal organization of living things has been studied in detail, there is every reason to assume that, sooner or later, we can artificially reproduce this organization and thereby directly show that life is nothing more than a special form of existence of matter.

— A.I. Oparin

Bibliography

Articles

[1] Michael F Land. "The evolution of eyes". In: 1990 (1992), pp. 1–29.

[3] Günter Wächtershäuser. "Before Enzymes and Templates: Theory of Surface Metabolism". In: *Microbiol. Rev.* 52.4 (1988), pp. 452–484.

[4] Leslie E. Orgel. "Prebiotic chemistry and the origin of the RNA world". In: *Crit. Rev. Biochem. Mol. Biol.* 39.2 (2004), pp. 99–123.

[6] Jean-paul Douliez et al. "Artificial Cell Models Catanionic Coacervate Droplets as a Surfactant-Based Membrane-Free Protocell Model Angewandte". In: *Angew. Chem. Int. Ed.* 56 (2017), pp. 13689–13693.

[7] R. Dawkins and J.R. Krebs. "Arms races between and within species". In: *Proc. R. Soc. B* 205.1161 (1997), pp. 489–511.

[8] Ramanarayanan Krishnamurthy. "Life's Biological Chemistry: A Destiny or Destination Starting from Prebiotic Chemistry?" In: *Chem. Eur. J.* 24 (2018), p. 16708.

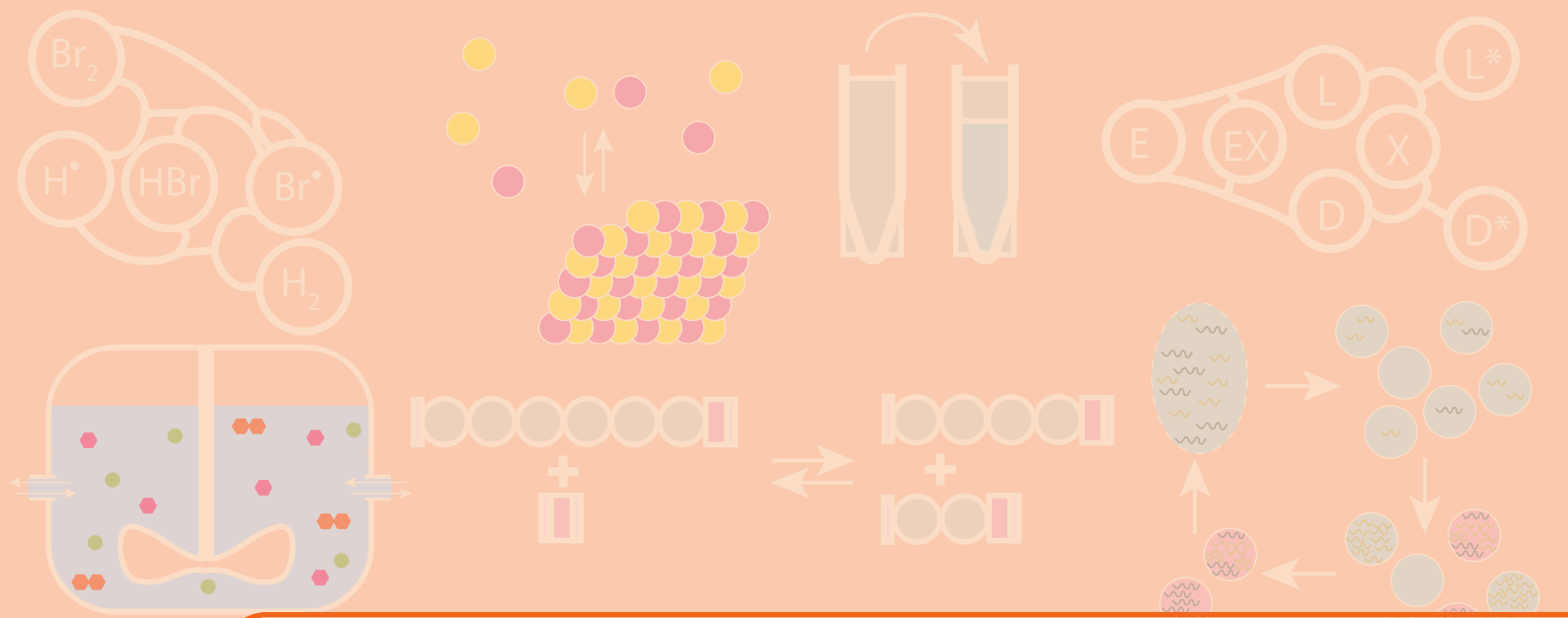
[9] Florian Hinzpeter, Ulrich Gerland, and Filipe Tostevin. "Optimal Compartmentalization Strategies for Metabolic Microcompartments". In: *Biophys. J.* 112.4 (2017), pp. 767–779.

Books

[2] A.G. Cairns-Smith. *Seven Clues To The Origin of life: a scientific detective story*. Cambridge, UK: Cambridge University Press, 1985.

[5] Marcus U. Lindström, ed. *Organic Reactions in Water*. Blackwell Publishing, 2007.

[‡]An interesting consideration of such a question is put forward in Ref.[9], where different compartmentalization strategies for enzymes in metabolic microcompartments are discussed, to optimize their yield.



10. Appendix

10.1 Appendix: Stochastic thermodynamics, the 0th law and currents

The thermodynamic chemostat is a key ingredient found in many models. As exhibited in Sec 3.1, such chemostats are far from hypothetical. While it is often practical to fix a concentration without further specification, it can be instructive to have a more detailed picture of what chemostatting is and what it can and cannot do.

To further flesh out the concept of chemostats, this section will look at them from the viewpoint of stochastic thermodynamics. To have chemostats that are in accord with the second law, we will see that, in general, we cannot always equalize concentrations or chemical potentials. Instead, a general chemostat will annul the net current, characterized by an equality of statistical moments of numbers of molecules. Only a limited class of current propensities are consistent with this more detailed description, and this constrains the transport of conserved quantities in general.

10.1.1 Two-particle exchange

Let us set up a simple, passive, two-particle transport process between a reservoir (R) and a system (I), for the species A, described by a reaction vector $\boldsymbol{g}$:

$$2\mathrm{A}^{\mathrm{I}} \underset{-\boldsymbol{g}}{\overset{\boldsymbol{g}}{\rightleftarrows}} 2\mathrm{A}^{\mathrm{R}}. \tag{10.1}$$

Such a transport process can e.g. be accomplished by some particular enzyme, in which case it is a catalytic cycle for which a more elaborate graph can be drawn (see Fig. 10.1). Another situation in which this may happens is when a species A may traverse the barrier separating reservoir and system only in its dimeric form A_2

$$2\mathrm{A}^{\mathrm{I}} \rightleftharpoons \mathrm{A}_2{}^{\mathrm{I}} \rightleftharpoons \mathrm{A}_2{}^{\mathrm{R}} \rightleftharpoons 2\mathrm{A}^{\mathrm{R}} \tag{10.2}$$

The next scheme presents a minimal example of a six-state enzyme transporting two particles at a time, between a small compartment and a large reservoir:

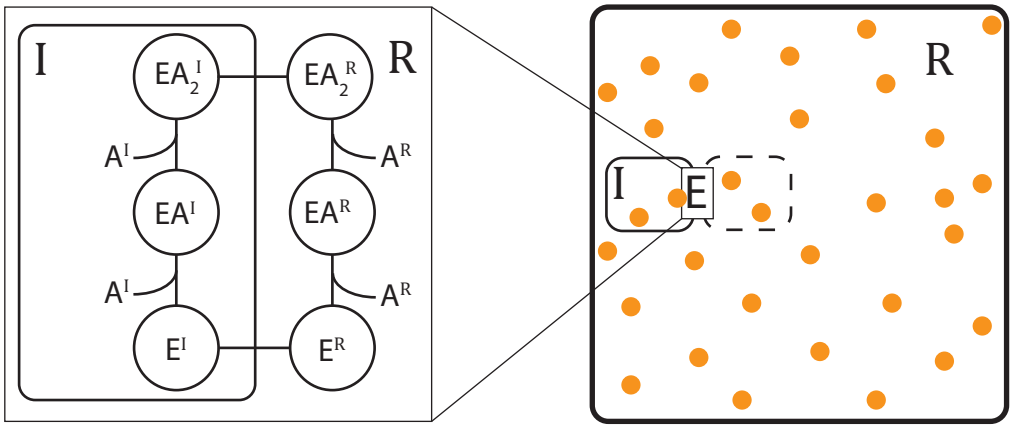


Figure 10.1: A minimalist scheme for a passive nonlinear transport enzyme transporting 2A at a time between compartment I and large reservoir R. The dashed box provides a reference volume v equal to the compartment size, such that the reservoir concentration of A is λ/v.

For both the system and the reservoir, we have two binding steps:

$$\mathrm{A}+\mathrm{E} \rightleftharpoons \mathrm{EA} \tag{10.3}$$
$$\mathrm{A}+\mathrm{EA} \rightleftharpoons \mathrm{EA}_2. \tag{10.4}$$

And two steps where the enzyme switches its orientation between reservoir and system.

$$\mathrm{E}^\mathrm{I} \rightleftharpoons \mathrm{E}^\mathrm{R} \tag{10.5}$$
$$\mathrm{EA}^\mathrm{I}_2 \rightleftharpoons \mathrm{EA}^\mathrm{R}_2. \tag{10.6}$$

It is the latter reaction that accounts for the net transfer of two particles. For simplicity *, we will suppose that the energy landscape is perfectly symmetric with respect to facing the compartment or reservoir

$$\mu^\circ_{\mathrm{E}^\mathrm{I}} = \mu^\circ_{\mathrm{E}^\mathrm{R}} = \mu^\circ_\mathrm{E}, \tag{10.7}$$
$$\mu^\circ_{\mathrm{EA}^\mathrm{I}} = \mu^\circ_{\mathrm{EA}^\mathrm{R}} = \mu^\circ_\mathrm{EA}, \tag{10.8}$$
$$\mu^\circ_{\mathrm{EA}_2{}^\mathrm{I}} = \mu^\circ_{\mathrm{EA}_2{}^\mathrm{R}} = \mu^\circ_{\mathrm{EA}_2} \tag{10.9}$$

Let us denote the total number of internal A molecules (roaming freely plus bound in EA^I, EA_2^I) as N_A. We can then define equilibrium constants for pairs of enzyme states

$$K_1 = \frac{P_{EA^I}}{N_A P_{E^I}} = \frac{P_{EA^R}}{\lambda P_{E^R}} = \exp(-\beta(\mu^\circ_\mathrm{EA} - \mu^\circ_\mathrm{E})), \tag{10.10}$$
$$K_2 = \frac{P_{EA^I_2}}{(N_A-1)P_{EA^I}} = \frac{P_{EA^R_2}}{\lambda P_{EA^R}} = \exp(-\beta(\mu^\circ_\mathrm{EA} - \mu^\circ_\mathrm{E})), \tag{10.11}$$
$$K_3 = \frac{P_{E^I}}{P_{E^R}} = \exp(-\beta(\mu^\circ_\mathrm{E} - \mu^\circ_\mathrm{E})) = 1. \tag{10.12}$$

Where we used for K_2 that a first A is already bound in EA^I, so only the remaining $N_A - 1$ that roam freely can bind to form EA^II. Letting λ denote the abundance of A molecules per volume unit v (corresponding to the region within the dotted lines Fig. 10.1) in the large reservoir ($m \gg 1$). On the reservoir side, binding an A from the bath reduces the mean concentration to $\lambda - \frac{1}{m}$. Since we have $m \gg 1$, binding does not noticeable alter the reservoir concentration: $\lambda - \frac{1}{m} \to \lambda$.

*This assumption can be relaxed for equivalent results

A master equation model

Let us now suppose that the rate of the exchange reaction (10.6) is much slower than any other transition, such that between subsequent transfers rapid equilibrium between the enzyme states is established. Between such transfers, the number of A molecules in I (free or bound to the enzyme facing I) N_A is fixed. For the macrostate N_A, the equilibrium occupation of microstates $\mathrm{EA}_2^{\mathrm{I}}$, $\mathrm{EA}_2^{\mathrm{R}}$ becomes

$$\frac{P_{EA_2^I}}{P_{E^I}+P_{EA^I}+P_{EA_2^I}+P_{E^R}+P_{EA^R}+P_{EA_2^R}} = \frac{K_1K_2N_A(N_A-1)}{2+K_1(N_A+\lambda)+K_1K_2(N_A(N_A-1)+\lambda^2)}, \tag{10.13}$$

$$\frac{P_{EA_2^R}}{P_{E^I}+P_{EA^I}+P_{EA_2^I}+P_{E^R}+P_{EA^R}+P_{EA_2^R}} = \frac{K_1K_2\lambda^2}{2+K_1(N_A+\lambda)+K_1K_2(N_A(N_A-1)+\lambda^2)}. \tag{10.14}$$

Let us in addition suppose $K_1\lambda, K_2\lambda^2 \ll 1$, such that binding sites are vacant most of the time, such that the equilibrium occupations approach

$$P_{EA_2^I} = \frac{K_1K_2N_A(N_A-1)}{2}, \tag{10.15}$$

$$P_{EA_2^R} = \frac{K_1K_2\lambda^2}{2}. \tag{10.16}$$

Where we have used the total probability $P_{E^I}+P_{EA^I}+P_{EA_2^I}+P_{E^R}+P_{EA^R}+P_{EA_2^R}=1$. Then, for a given macrostate N_A, we can define rates J for two-particle transport from I to R and from R to I as

$$J_{EA_2^I\to EA_2^R} \propto N_A(N_A-1), \tag{10.17}$$

$$J_{EA_2^R\to EA_2^I} \propto \lambda^2. \tag{10.18}$$

Where the first reaction corresponds to a transation from N_A to N_A-2 and the second to a transition from N_A to N_A+2. We can then write a master equation to track the fluctuations of the population of A in the internal compartment

$$\begin{aligned}\frac{dP(N_\mathrm{A})}{dt} &= k_h\lambda^2\big[P(N_\mathrm{A}-2)-P(N_\mathrm{A})\big] \\ &- k_h\big[N_\mathrm{A}(N_\mathrm{A}-1)P(N_\mathrm{A})-(N_\mathrm{A}+2)(N_\mathrm{A}+1)P(N_A+2)\big] \quad N_A\geq 2\end{aligned} \tag{10.19}$$

$$\frac{dP(N_\mathrm{A})}{dt} = -k_h\lambda^2P(N_\mathrm{A})+k_h(N_\mathrm{A}+2)(N_\mathrm{A}+1)P(N_A+2) \qquad 0\leq N_A\leq 2 \tag{10.20}$$

Where k_h is a rate constant. The stationary equilbrium distribution of this equation is a generalization of the Poisson distribution typically encountered for such equations

$$P(N_\mathrm{A}) = \frac{\lambda^{N_\mathrm{A}}C_0}{N_\mathrm{A}!} \tag{10.21}$$

Where C_0 is a normalization constant.

Normalization and moments of the distribution

Since we transport 2 particles at a time, $N_\mathrm{A} \mod 2$ is constant. To characterize C_0, we need to distinguish between odd and even values for N_A. For even N_A, we have:

$$\frac{1}{C_0} = \sum_{n=0}^{\infty}\frac{\lambda^{2n}}{2n!} = \cosh(\lambda) \tag{10.22}$$

Which yields, for the moments of N_A:

$$\begin{aligned}\langle N_\mathrm{A}^m\rangle &= \frac{1}{\cosh(\lambda)}\sum_n \frac{(2n)^m\lambda^{2n}}{2n!} \\ &= \frac{(\lambda\frac{d}{d\lambda})^m}{\cosh(\lambda)}\sum_n\frac{\lambda^{2n}}{2n!} = \frac{(\lambda\frac{d}{d\lambda})^m\cosh(\lambda)}{\cosh(\lambda)} \\ &= \sum_{n=1}^{m}\lambda^n(\delta^1_{n\ mod2}\tanh(\lambda)+\delta^0_{n\ mod2})S(m,n)\end{aligned} \tag{10.23}$$

where δ^b_a is the Kronecker delta symbol, and $S(m,n)$ denotes a Stirling number of the second kind, which quantifies the number of ways m objects can be partitioned in n nonempty sets. The first two moments are then

$$\langle N_\mathrm{A}\rangle = \lambda\tanh(\lambda),\ \ \langle N_\mathrm{A}^2\rangle = \lambda^2+\lambda\tanh(\lambda) \tag{10.24}$$

For odd N_A, we have:

$$\frac{1}{C_0} = \sum_{n=0}^{\infty}\frac{\lambda^{2n+1}}{(2n+1)!} = \sinh(\lambda). \tag{10.25}$$

Which yields different values for the moments of N_A:

$$\begin{aligned}\langle N_\mathrm{A}^m\rangle &= \frac{1}{\sinh(\lambda)}\sum_n \frac{(2n)^m\lambda^{2n}}{2n!} \\ &= \frac{(\lambda\frac{d}{d\lambda})^m}{\sinh(\lambda)}\sum_n\frac{\lambda^{2n}}{2n!} = \frac{(\lambda\frac{d}{d\lambda})^m\sinh(\lambda)}{\sinh(\lambda)} \\ &= \sum_{n=1}^{m}\lambda^n(\delta^1_{n\ mod2}\coth(\lambda)+\delta^0_{n\ mod2})S(m,n).\end{aligned} \tag{10.26}$$

With the first two moments being

$$\langle N_\mathrm{A}\rangle = \lambda\coth(\lambda),\ \ \langle N_\mathrm{A}^2\rangle = \lambda^2+\lambda\coth(\lambda) \tag{10.27}$$

Interestingly, the odd and even systems have dissimilar moments of their respective distributions for N_A, despite being in contact with the same reservoir. For $\lambda \gg 1$, $\coth(\lambda)\to 1, \tanh(\lambda)\to 1$ and thus these effects vanish. While the system was in contact with a chemostat of concentration λ/V, we clearly did not obtain the average concentration λ/V as concentration for our system. This is due to the nonlinear coupling.

The reservoir influx $J^+_{R\to I} = k_h\lambda^2$ needs to balance the system outflux $J^+_{I\to R} = k_h\langle N_\mathrm{A}^2\rangle - \langle N_\mathrm{A}\rangle$, in order to establish equilibrium:

$$k_h\lambda^2 = k_h\sum_{N_\mathrm{A}} N_\mathrm{A}(N_\mathrm{A}-1)P(N_\mathrm{A}) = k_h(\langle N_\mathrm{A}^2\rangle - \langle N_\mathrm{A}\rangle). \tag{10.28}$$

Upon inspection, we see that this condition is verified using the moments for the 'even' and the 'odd' compartment. Cancelling a net current is thus more fundamental than equalizing a concentration.

10.1.2 Coupled Compartments

Let us now take an odd compartment I and couple it to an even compartment II (see Fig. 10.2), with both coupled to a reservoir. We suppose that reservoir exchange is rapid, and intercompartment exchange is slow (e.g. due to a less efficient enzyme, or different enzyme abundances at the

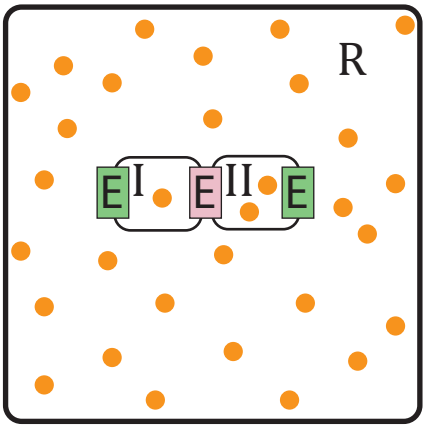


Figure 10.2: Setup for transport between compartments I and II, both coupled with the same reservoir R. Enzymes (E) given in green perform rapid exchange compared to the pink one in the middle ($\kappa \gg k_h$), which slowly exchanges 2A. On average, no net currents should establish since there is no driving force.

contacts), such that molecular abundances follow Eq. (10.21). We furthermore suppose the same current propensities for this coupling

$$J^+_{I\to II} = \kappa N^I_A (N^I_A - 1), \tag{10.29}$$
$$J^+_{II\to I} = \kappa N^{II}_A (N^{II}_A - 1). \tag{10.30}$$

where $\kappa \ll k_h$. Since we assume that transport between compartments is slow ($\kappa \ll k_h$), we can use the steady-state solutions for reservoir exchange to find the mean currents

$$\langle J^+_{I\to II}\rangle = \kappa \langle N^I_A (N^I_A - 1)\rangle = \kappa \lambda^2, \tag{10.31}$$
$$\langle J^+_{II\to I}\rangle = \kappa \langle N^{II}_A (N^{II}_A - 1)\rangle = \kappa \lambda^2. \tag{10.32}$$

In this example where N mod 2 is preserved, exchange with a reservoir reaches equilibrium for an even compartment and an odd compartment. At equilibrium, however, these compartments do not have the same average concentration (see Eqs. (10.27) and (10.24)). In fact, if they did have the same equilibrium concentration, the current $\langle J^+_{I\to II}\rangle - \langle J^+_{II\to I}\rangle$ would not vanish. In the absence of a driving force (both are connected to the same reservoir), this is in contradictation with the second law, which precludes persistent currents at equilibrium.

One can define a $\Delta\mu$ for the process, whose exponential vanishes on average. Let us define a chemical potential for reservoir species of the form

$$\bar{\mu} = \mu^\circ + k_b T \ln \lambda, \tag{10.33}$$

By equality of opposing currents between compartment and reservoir (Eq. (10.28)) and pairs of compartments, it follows that

$$\langle N^I_A (N^I_A - 1)\rangle \exp(2\beta\mu^\circ) = \langle N^{II}_A (N^{II}_A - 1)\rangle \exp(2\beta\mu^\circ) = \exp(2\beta\bar{\mu}) \tag{10.34}$$

For the compartment, we can then introduce a chemical potential of the form

$$\mu(N_A) = \mu^\circ + k_b T \ln N_A, \tag{10.35}$$

the rationale behind such a form is that it very naturally recovers the Gibbs free energy G in the statistical sense for a state N_A

$$G(N_A) = \sum_{k=1}^{N_A} (\mu^\circ + k_b T \ln k) = N_A \mu^\circ + k_b T \ln N_A!. \tag{10.36}$$

This expression then modifies Eq. (10.34) to

$$\langle \exp(\beta\mu(N_A^I)+\beta\mu(N_A^I-1))\rangle = \langle \exp(\beta\mu(N_A^{II})+\beta\mu(N_A^{II}-1))\rangle = \exp(2\beta\bar{\mu}). \tag{10.37}$$

Eq. (10.37) has a functional form reminiscent of a Jarzinsky equality[1], although its interpretation is quite different.

The Jarzynski equality relates the statistics of the work W done in taking the system from state α to β, to the Helmholtz free energy F of those states

$$\langle \exp(W)\rangle = \exp(-\beta(F(\alpha)-F(\beta))), \tag{10.38}$$

In Eq. (10.37), no work is extracted or injected and the relation is an equilibrium relation. Its utility here is that it characterizes equilibrium, whereas more straightforward quantities do not. The process does not equalize chemical potentials in the sense of a uniporter (Eq. (3.1)), since we know from Eqs. (10.27), (10.24)) that

$$\langle N_A^I\rangle \neq \langle N_A^{II}\rangle \neq \lambda, \tag{10.39}$$

and by taking into account Eq. (10.35) we see that an average chemical potential does not equalize either

$$\langle \mu(N_A^I)\rangle \neq \langle \mu(N_A^{II})\rangle \neq \bar{\mu}. \tag{10.40}$$

nor do the exponentials that one would obtain from (10.39)

$$\langle \exp(-\beta\mu(N_A^I))\rangle \neq \langle \exp(-\beta\mu(N_A^{II}))\rangle \neq \lambda, \tag{10.41}$$

Introducing

$$\begin{aligned}
\Delta\mu^I &= \mu(N_A^I)+\mu(N_A^I-1)-2\bar{\mu}, && (10.42)\\
\Delta\mu^{II} &= \mu(N_A^{II})+\mu(N_A^{II}-1)-2\bar{\mu}, && (10.43)\\
\Delta\mu^{III} &= \mu(N_A^I)+\mu(N_A^I-1)-\mu(N_A^{II})-\mu(N_A^{II}-1), && (10.44)
\end{aligned}$$

we can write Eq. (10.41) in more compact form by pairing a compartment with another compartment or a reservoir

$$\langle \exp(-\beta\Delta\mu^I)\rangle = \langle \exp(-\beta\Delta\mu^{II})\rangle = \langle \exp(-\beta\Delta\mu^{III})\rangle = 1. \tag{10.45}$$

While Eq. (10.38) pertains to the ΔF of one pair of states, the chemical relation Eq. (10.37) gives a weighted average over all accessible pairs of N_A, N_A-2 states to find the quantity $\mu(N_A)+\mu(N_A-1)$, which is subsequently rewritten as a $\Delta\mu$ on the level of the whole system.

Out of equilibrium

Let us now slightly modify the situation, by linking compartment I and II to different reservoirs with chemical potentials $\bar{\mu}^I, \bar{\mu}^{II}$ (Fig. 10.3), with bath concentrations λ^I, λ^{II}, respectively. Maintaining that transport between reservoirs is much faster than transport between compartments ($k_h \gg \kappa$), the occupations are fixed by

$$\begin{aligned}
k_h\langle N_A^I(N_A^I-1)\rangle &= k_h(\lambda^I)^2, && (10.46)\\
k_h\langle N_A^{II}(N_A^{II}-1)\rangle &= k_h(\lambda^{II})^2. && (10.47)
\end{aligned}$$

Maintaining the same current propensities for intercompartment transport, (Eqs. (10.29)-(10.30)), we find a net transport rate

$$\langle J^+_{I\to II}\rangle - \langle J^+_{II\to I}\rangle = \kappa(\langle N_A^I(N_A^I-1)\rangle - \langle N_A^{II}(N_A^{II}-1)\rangle) = \kappa((\lambda^I)^2-(\lambda^{II})^2) \tag{10.48}$$

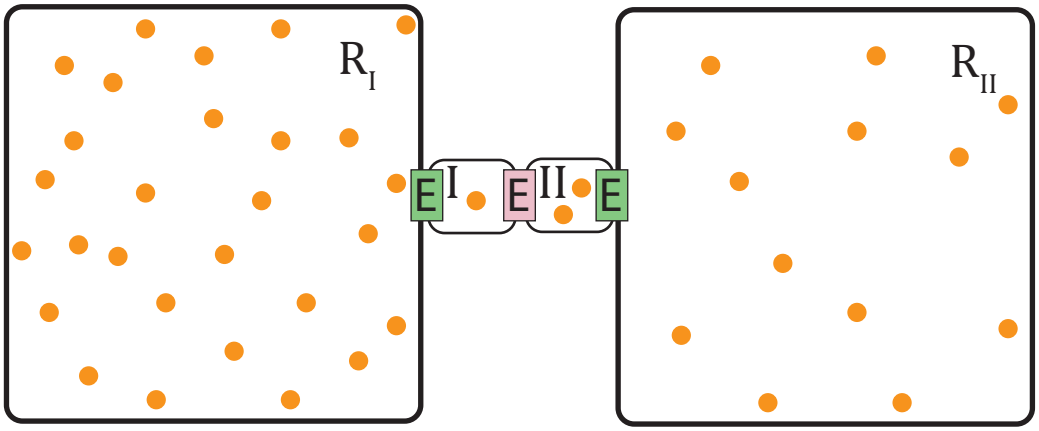


Figure 10.3: Setup for transport between reservoir R_I and R_{II}, through two coupled compartments. Enzymes (E) given in green perform rapid exchange compared to the pink one in the middle ($\kappa \gg k_h$), which slowly exchanges 2A.

Taking the ratio of currents, we find a de Donder [2] form,

$$\frac{\langle J^+_{I\to II}\rangle}{\langle J^+_{II\to I}\rangle} = \frac{\langle N^I_A(N^I_A-1)\rangle}{\langle N^{II}_A(N^{II}_A-1\rangle} = \exp(-2\beta(\Delta\bar{\mu})), \tag{10.49}$$

with an affinity of $2\Delta\bar{\mu}$, where $\Delta\bar{\mu} = \bar{\mu}^I - \bar{\mu}^{II}$. Written in terms of chemical potentials, we find

$$\langle \exp(-\beta[\mu(N^I_A)+\mu(N^I_A-1)-\mu(N^{II}_A)-\mu(N^{II}_A-1)])\rangle = \exp(-2\beta\Delta\bar{\mu}). \tag{10.50}$$

Due to rapid equilibration between reservoir and compartment, local equilibrium can be supposed on intermediate timescales $1/k_h \ll t \ll 1/\kappa$,

$$\langle \exp(-\beta\Delta\mu^I)\rangle = \langle \exp(-\beta\Delta\mu^{II})\rangle = 1. \tag{10.51}$$

where

$$\Delta\mu^I = \mu(N^I_A)+\mu(N^I_A-1)-2\bar{\mu}^I, \tag{10.52}$$
$$\Delta\mu^{II} = \mu(N^{II}_A)+\mu(N^{II}_A-1)-2\bar{\mu}^{II}. \tag{10.53}$$

10.1.3 A link with the zeroth law of thermodynamics

The zeroth law of thermodynamics states that thermodynamic equilibrium between systems is a transitive property. Its formulation typically focuses on the equilibrium in the exchange of energy, thereby fixing the temperature, but this specification is typically unnecessary. Often, the zeroth law is given as:

If a system A is in equilibrium with a system B and A is in equilibrium with C, then B is in equilibrium with C. [3]

Where B is in equilibrium with C is often specified with 'if B is brought in contact with C, no further change occurs'.

A Macroscopic application of the zeroth law

Pippard[3] gives a useful argument for how the zeroth law implies the existence of an equation of state, which we will give here in a more the generalized sense. The original argument holds that contact between phases to introduces one constraint, as only transport of heat is considered, but in principle it can be more.

To establish this, Pippard considers three systems (A,B,C). These systems are in local equilibrium within themselves, and the state is characterized by a set of n variables $(x_1^A, x_2^A, \dots)$, (e.g. Volume, pressure, etc.). Now let us first consider placing two systems in contact via a diathermal wall. These systems are not in equilibrium for any arbitrary choice of $\{x_i^A\}$, $\{x_i^C\}$, but fixing $2n-1$ of them will lead the unspecified variable to adopt its equilibrium value upon relaxation by heat flow. The condition for which the variables reach equilibrium are captured by an equation

$$F_1(x_1^A,\dots,x_n^A,x_1^C,\dots,x_n^C) = 0. \tag{10.54}$$

Which we will rewrite for x_1^C as

$$x_1^C = f_1(x_1^A,\dots,x_n^A,x_2^C,\dots,x_n^C), \tag{10.55}$$

By the exact same argument, equilibrium between B and C is captured by

$$F_2(x_1^B,\dots,x_n^B,x_1^C,\dots,x_n^C) = 0. \tag{10.56}$$

From which we can again find x_1^C

$$x_1^C = f_2(x_1^B,\dots,x_n^B,x_2^C,\dots,x_n^C). \tag{10.57}$$

Which means that

$$f_1(x_1^A,\dots,x_n^A,x_2^C,\dots,x_n^C) = f_2(x_1^B,\dots,x_n^B,x_2^C,\dots,x_n^C). \tag{10.58}$$

According to the zeroth law of thermodynamics stated before, we have now established equilibrium between A and C as well, as if they were put in direct contact. Therefore, Eq. (10.58) must be equivalent to

$$F_3(x_1^A,\dots,x_n^A,x_1^B,\dots,x_n^B) = 0 \tag{10.59}$$

Which means the functional dependence in f_1 and f_2 on variables in C must be such that they cancel each other out, e.g.

$$\phi_1(x_1^A,\dots,x_n^A)\zeta(x_2^C,\dots,x_n^C) + \eta(x_2^C,\dots,x_n^C) \tag{10.60}$$

Then, we find that upon cancellation, we can write the equilibrium as a relation between two functions of state ϕ_1,ϕ_2, as ϕ_1,ϕ_2 only depends on variables of their corresponding system,

$$\phi_1(x_1^A,\dots,x_n^A) = \phi_2(x_1^B,\dots,x_n^B). \tag{10.61}$$

Extending this argument, for system C, we find

$$\phi_1(x_1^A,\dots,x_n^A) = \phi_2(x_1^B,\dots,x_n^B) = \phi_3(x_1^C,\dots,x_n^C). \tag{10.62}$$

Which has the elegant property that it takes the same value for all systems at equilibrium, and for a diathermal wall, ϕ can be identified with the quantity that we call temperature T. Strictly speaking, equation (10.62) is valid for any function $G(T)$ that only depends on temperature. Since thermal equilibrium implies that all temperatures are the same, the zeroth law has been rephrased as 'all diathermal walls are equivalent' [4] or the more demeaning version 'you can build a thermometer'.

While we repeated Pippard's argument for the case of temperature, the argument could also concern a different quantity. Let us now take systems A,B,C in contact with a heat bath at temperature T. If we then establish a connection between A and C, e.g. by a movable diathermal wall or a semipermeable membrane, we again have one variable that we cannot fix and whose value

is set by equilibrium. Like for Eq. (10.54), we can write the condition for equilibrium as function of the variables as

$$H_1(x_1^A,\ldots,x_n^A,x_1^C,\ldots,x_n^C)=0. \tag{10.63}$$

By pursuing Pippard's line of reasoning outlined before, we find a property χ associated to the quantity that is equalized

$$\chi_1(x_1^A,\ldots,x_n^A)=\chi_2(x_1^B,\ldots,x_n^B)=\chi_3(x_1^C,\ldots,x_n^C). \tag{10.64}$$

Now, the quantity χ is for example the chemical potential (selective membrane), or the pressure (movable wall). However, we must be slightly more careful with a statement like 'all diachemical' walls are equivalent.

Let us first note that Pippard's argument is rooted in classical thermodynamics and the line of reasoning alludes to macroscopic systems for which we can neglect fluctuations. For small, nanoscopic systems, some of our variables (e.g. local temperature, number of particles) become fluctuating quantities, and expressions like Eq. (10.62) or Eq. (10.64) become true on average. For example, in the ideal chemostat performing single-particle exchange, the average, fluctuating concentration (Eq. (3.15)) is equalized with that of the reservoir.

The zeroth law has recently been reinspected in a number of regimes beyond its typical perceived validity. Pradhan, Amann and Seifert[5] showed that nonequilibrium-steady states of lattice gases could be put in contact, to yield an approximate zeroth law. Bera et al [6] generalized laws of thermodynamics from information theory, to address quantum-mechanical situations where strong coupling with baths can exist. In such cases, the classical zeroth law does not suffice. In the upcoming section, we will show that we do not need to be in a NESS or the quantum regime to expect violations of certain formulations of the zeroth law. We present a case in this happens for three coupled classical chemical systems at equilibrium.

Nonequivalent walls in classical thermodynamics

For integer quantities like the number of particles, there is another important difference: the number of species transported dictates the equilibrium that is established. Let us consider three systems (α,β,γ), and let α be connected to β through an enzyme performing N mod 2 transport (see Fig. 10.4). Furthermore, let β be connected to γ through an enzyme performing N mod 3 transport. In the present case, a net equilibrium will be established between α, β and γ, where the transport fixes

$$\langle N^\alpha(N^\alpha-1)\rangle = \langle N^\beta(N^\beta-1)\rangle \tag{10.65}$$

$$\langle N^\beta(N^\beta-1)(N^\beta-2)\rangle = \langle N^\gamma(N^\gamma-1)(N^\gamma-2)\rangle. \tag{10.66}$$

Transport will preserve α's mod 2 state and γ's mod 3 state, while β will have a mod 1 state. Although these systems establish an equilibrium by exchanging their common component, they have distinct equilibrium states. A quick way to see this is by having one compartment e.g. (β) be considerably larger (or coupled with a reservoir fixing the concentration at λ). Setting the size of α and γ equal, we can write

$$P(N^\alpha)=\frac{\lambda^{N^\alpha}C_0^\alpha}{N^\alpha!}, \qquad P(N^\beta)=\frac{\lambda^{N_\beta}C_0^\beta}{N_\beta!}. \tag{10.67}$$

where C_0^α is a mod 2 normalization constant and C_0^α a mod 3 normalization constant. Let $m=N^\gamma$modn, the normalization constant is then constructed by taking every nth term in the expression of an exponential.

$$\frac{1}{C_0}=\sum_{k=0}^{\infty}\frac{\lambda^{nk+m}}{(nk+m)!} \tag{10.68}$$

For $n = 1,2$ this respectively yields the well-known exponential function and the hyperbolic functions.

By the coupling of higher moments, the local equilibrium states are distinct and the average abundances do not equalize

$$\langle N^\alpha \rangle \neq \langle N^\beta \rangle \neq \langle N^\gamma \rangle. \tag{10.69}$$

Now, if α and γ would subsequently be connected (e.g. by a pore) to perform mod 1 transport, we would suddenly fix

$$\langle N^\alpha \rangle = \langle N^\gamma \rangle. \tag{10.70}$$

By this introduction, the populations of α (resp. γ) are no longer constrained by Nmod2 (resp. Nmod3), instead putting them all at Nmod1. Let us now consider what modulo constraints are preserved for different connections, also summarized in Fig. 10.4): connecting α to γ and γ to β fixes

$$N^\alpha \text{mod1}, \quad \text{N}^\beta \text{mod3}, \quad \text{N}^\gamma \text{mod1} \tag{10.71}$$

Whereas connecting α to β and β to γ fixes

$$N^\alpha \text{mod2}, \quad \text{N}^\beta \text{mod1}, \quad \text{N}^\gamma \text{mod3} \tag{10.72}$$

And connecting A to β and α to γ fixes

$$N^\alpha \text{mod1}, \quad \text{N}^\beta \text{mod2}, \quad \text{N}^\gamma \text{mod1} \tag{10.73}$$

It follows that not all diachemical walls are equivalent, and by extension the order in which we connect systems matters. Any of the proposed configurations reaches equilibrium, they are simply not the same equilibria. One may now wonder what it means for compartments that are not in direct contact to be in equilibrium 'with each other'. Macroscopically, it means that if we connect α to β an β to γ, then connecting α to γ leads to no further change, as they are already in equilibrium.

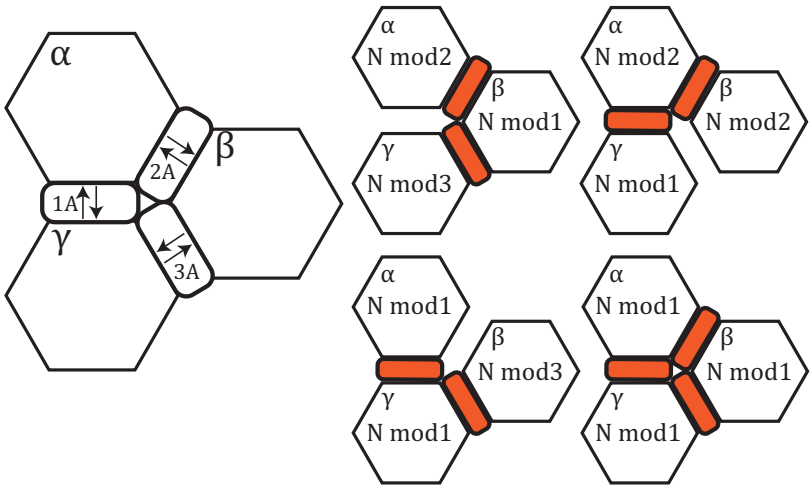


Figure 10.4: Different equilibrium states achieved by connecting (indicated in orange) different diachemical walls, transporting 1, 2, or 3 α simultaneously.

In our microscopic example, this ceases to be true if we have different types of diachemical walls: to preserve the modulo constraints, the number of particles exchanged between α and γ must

be a common divisor of 2 and 3. The smallest number of particles that we can transport between α and γ is then 6, thereby fixing

$$\langle (N^{\alpha})_6 \rangle = \langle (N^{\gamma})_6 \rangle. \tag{10.74}$$

Where we have used the falling factorial expression as defined in Eq. (3.22). With this choice connecting α to β and β to γ is followed by no further change when α is connected to γ.

In this case however, connecting α to γ and γ to β fixes (see Fig. 10.5)

$$N^{\alpha}\text{mod6}, \quad \text{N}^{\beta}\text{mod3}, \quad \text{N}^{\gamma}\text{mod3} \tag{10.75}$$

Whereas connecting α to β and β to γ fixes

$$N^{\alpha}\text{mod2}, \quad \text{N}^{\beta}\text{mod1}, \quad \text{N}^{\gamma}\text{mod3} \tag{10.76}$$

And connecting α to β and α to γ fixes

$$N^{\alpha}\text{mod2}, \quad \text{N}^{\beta}\text{mod2}, \quad \text{N}^{\gamma}\text{mod6} \tag{10.77}$$

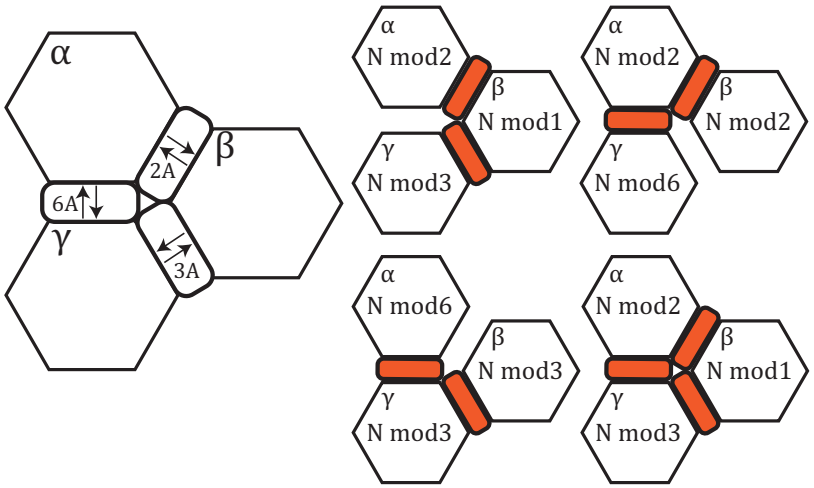


Figure 10.5: Different equilibrium states achieved by connecting (indicated in orange) different diachemical walls, transporting 2, 3 or 6 A simultaneously.

For macroscopic systems, the differences that follow from this are minute. As was shown for mod 2 systems, we have for even compartments $\langle N \rangle = \lambda \tanh(\lambda)$. Noting that $\tanh(\lambda)$ can be written as

$$\tanh(\lambda) = 1 - \frac{2}{\exp(2\lambda)+1}, \tag{10.78}$$

we find that $\tanh(\lambda)$ converges to 1 exponentially fast with growing λ, meaning corrections will rapidly become so small that we will not be able to measure a deviation from a mod 1 state. This argument is applicable to any Nmodn state (provided $n \ll N$).

For all practical purposes, a system can thus be large enough such that all diachemical walls lead to indistinguishable changes and no further caution is needed when considering the zeroth law. We can then comfortably return to using Pippard's expression $\phi_1(x_1^A,\ldots,x_n^A) = \phi_2(x_1^B,\ldots,x_n^B) = \phi_3(x_1^C,\ldots,x_n^C)$.. In stochastic thermodynamics, formulating the zeroth law becomes more delicate for integer quantities.

Let us first consider this: when we connect a system A to B and B to C, then connection A to C only leads to no change if that connection does not alter the modulo states $N_A \mathrm{mod} \mathrm{n_A}$ and $N_C \mathrm{mod} \mathrm{n_C}$. Let us now introduce a connection between A and C that exchanges m particles at a time. If $m \mathrm{mod} \mathrm{N_A} = 0$ and $m \mathrm{mod} \mathrm{N_C} = 0$, this maintains the original equilibrium state. In this sense then, equilibrium between A and B and equilibrium between B and C implies equilibrium between A and C. In Fig. 10.5, such a situation is given.

A subtlety that arises is that the appropriate contact through which A and C would already be at equilibrium is hypothetical. Their actual exchange mechanisms or chemistries may be different. While we may maintain that there is transitivity and equilibrium, it does not follow that putting them in contact will maintain that equilibrium. Let us here suggest a formulation of the zeroth law that attempts to capture this subtlety

A zeroth law for integer quantities: If a system A is in equilibrium with a system B and A is in equilibrium with C, then B is in equilibrium with C for the appropriate subclass of contacts.

We here understand 'the appropriate subclass of contacts' as diachemical walls that impose the appropriate modulo constraints and fix the appropriate statistical moments that govern exchange. In the following section we will show that these two criteria are closely linked in general. We show that arbitrary passive processes performing mod 2 exchange (between ideal systems) must at least equalize the quantity $N(N-1)/V^2$. Most functional forms for current propensities do not allow for such an equilibration. For passive processes, such current propensities are forbidden by thermodynamics.

10.1.4 A black box performing mod 2 exchange

Suppose we have again have our reservoir coupled to a left and a right system L and R, but the left and right systems are now coupled by a black box: a mysterious, ill-specified subsystem (see 10.6). The following properties are known for the black box: it passively transports particles and preserves N mod 2. Its internal state structure can be completely different from E however.

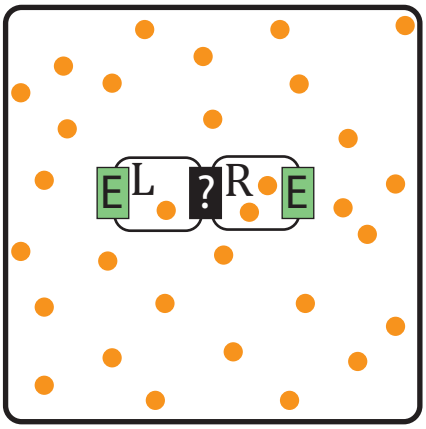


Figure 10.6: Setup for transport between compartments L and R, through a 'black box' labellled with an interrogation mark '?'. L and R are linked to a large reservoir. While the exact nature of the black box is unknown, it is known that it transports 2A at a time and operates passively.

Let us first consider a black box with a left-to-right flux rate obeying some function $J^{LR}(N_L)$, where N_L corresponds to the number of species A in the left compartment. We will refer to black boxes that obey this rule as 'simple' black boxes (in anticipation of 'complex' black boxes, for which currents do not only depend on concentrations in one compartment).

For a finite state space (finite number of possible N_L populations), we can always expand J^{LR} in powers of N_L, since we can supply a polynomial coefficient for every state so that a full exact fit is

realized. The instantaneous flux rate in a given state N_L is then

$$J^{LR}(N_L) = a_0 + a_1 N_L + a_2 N_L^2 + ... = \sum_{k=0}^{\infty} a_k N_L^k \tag{10.79}$$

where $a_0, a_1, a_2, ...$ are constants. Note that by introducing this flux rate, we are implicitly implying an appropriate seperation of timescales, such that memory effects from the black box become vanishingly small.

For a stationary probability distribution of N_L, let us then write the average flux rate

$$\langle J^{LR}(N_L)\rangle = \sum_{N_L}^{\infty}\sum_{i=0}^{\infty} a_i N_L^i P(N_L) = a_0 + a_1\langle N_L\rangle + a_2\langle N_L^2\rangle + ... \tag{10.80}$$

Similarly, we can define an instantaneous right-to-left flux rate $J^{RL}(N_R)$, which averages to

$$\langle J^{RL}(N_R)\rangle = b_0 + b_1\langle N_R\rangle + b_2\langle N_R^2\rangle + ... \tag{10.81}$$

Let us now consider the case of empty compartments and reservoirs: in such a case there are no particles to transport, and all fluxes should cancel, which means $a_0 = b_0 = 0$. In the limit where transport between the reservoir and compartments is fast with respect to intercompartment transport, the steady state particle number distributions correspond to those derived in Sec. 10.1.1. We can therefore use Eqs. (10.26) and (10.23) to express the flux rates as powers of λ. If the left compartment contains an even number of particles, we find

$$\langle J^{LR}\rangle = \sum_{n=0}^{\infty} \lambda^{1+2n} \tanh(\lambda) \sum_{i=2n+1}^{\infty} S(i,2n+1)a_i + \sum_{n=1}^{\infty} \lambda^{2n} \sum_{i=2n}^{\infty} S(i,2n)a_i \tag{10.82}$$

whereas for an odd ($N \mathrm{mod} 2 = 1$) compartment, we have

$$\langle J^{LR}\rangle = \sum_{n=0}^{\infty} \lambda^{1+2n} \coth(\lambda) \sum_{i=2n+1}^{\infty} S(i,2n+1)a_i + \sum_{n=1}^{\infty} \lambda^{2n} \sum_{i=2n}^{\infty} S(i,2n)a_i \tag{10.83}$$

In stochastic thermodynamics, the second law requires that at equilibrium all currents vanish on average. In our system, there is a reservoir, two compartments and a black box. However, no net currents should be produced, since there is no affinity to do so (chemically, that would require connecting to a second reservoir [7]). The second law thus implies that the coefficients must be such that $\langle J^{LR}\rangle = \langle J^{RL}\rangle$ for any value of λ and any pair of mod 2 compartments (even-even, even-odd, odd-even, odd-odd). For an odd-even pair, we can write for odd powers of λ that

$$\forall m \in \{1,3,5,...\} \quad \coth(\lambda) \sum_{i=m}^{\infty} S(i,m)a_i = \tanh(\lambda) \sum_{i=m}^{\infty} S(i,m)b_i. \tag{10.84}$$

Which can only be true regardless of λ if odd-power contributions vanish

$$\forall m \in \{1,3,5,...\} \quad \sum_{i=m}^{\infty} S(i,m)a_i = \sum_{i=m}^{\infty} S(i,m)b_i = 0. \tag{10.85}$$

Irrespective of the choice of mod 2 compartments, the even powers of λ verify

$$\sum_{i=m}^{\infty} S(i,m)a_i = \sum_{i=m}^{\infty} S(i,m)b_i = \gamma_m \quad \forall m \in \{2,4,6,...\} \tag{10.86}$$

Which means the mean flux in either direction can be written as

$$\langle J^{LR}\rangle = \langle J^{RL}\rangle = \sum_{n=1}^{\infty} \gamma_{2n}\lambda^{2n} \tag{10.87}$$

If the instantaneous flux rate has a finite number of coefficients M, we can write the matrix equation

$$\boldsymbol{S}\cdot(\boldsymbol{a}-\boldsymbol{b})=0 \tag{10.88}$$

where $\boldsymbol{S}$ is an $M\times M$ matrix with entries $S(i,m)$ that are the Stirling numbers of the second kind, and $\boldsymbol{a},\boldsymbol{b}$ are vectors of the coefficients a_i,b_i. Since $\boldsymbol{S}$ is linearly independent, we have M coefficients that are fixed by M constraints, which implies $a_i=b_i$. This means that the forward and backward transport propensities are completely symmetric.

The constants verify the matrix equation

$$\boldsymbol{S}\cdot\boldsymbol{a}=\boldsymbol{\gamma} \tag{10.89}$$

where $\boldsymbol{\gamma}$ is a vector of γ_m values, which are 0 for odd m. The solutions for this matrix equation on the Stirling numbers of the second kind, is a linear combination of vectors $\boldsymbol{c}_n$ of signed Stirling numbers of the first kind $c(n,k)$

$$\boldsymbol{a}=\sum_{n=2}^{M}\beta_n\boldsymbol{c}_n \tag{10.90}$$

The signed Stirling numbers of the first kind occur e.g. in falling factorial expressions like $N(N-1)(N-2)(N-3)$, which can be expanded as

$$-6N+11N^2-6N^3+N^4=c(4,1)N+c(4,2)N^2+c(4,3)N^3+c(4,4)N^4 \tag{10.91}$$

Substituting Eq. (10.90) in Eq. (10.79), the instantaneous current that preserves N mod 2 must become

$$J(N)=\sum_{n=1}^{\infty}\beta_{2n}(N)_{2n}, \tag{10.92}$$

where $(N)_m$ denotes a falling factorial $N(N-1)....(N-m+1)$. In order to satisfy (10.87) irrespective of λ, we equalize the quantity $\langle(N)_m\rangle$ (more completely: $\langle(N)_m\rangle/V^m$) for each nonzero entry β_m. This is in accord with the equilibrium criterion (10.37).

In conclusion, Nmod2 preserving transport is highly constrained. In this section we assumed the outgoing current propensity can be written as a power law with a finite number of terms that only depend on the internal concentration. Under these conditions, the current propensity is exactly equal in either direction and must be written as a linear combination of even falling factorials of particle numbers.

A similar proof can be written for transport preserving Nmodm, which leads to a current of the form

$$J(N)=\sum_{n=1}^{\infty}\beta_{mn}(N)_{mn}. \tag{10.93}$$

10.1.5 A mod 2 preserving complex black box

In the previous section we imagined a black box, where the outgoing transport of a compartment was only a function of concentrations in that compartment. In principle, we can imagine this transport to be affected by the system state of the other compartment. To illustrate this, imagine a modified version of our transport enzyme, with an extra 'gating' binding step (see fig. 10.7).

Let us write the instantaneous current as a power law of particle numbers on both sides

$$J^{LR} = a_{0,0}+a_{1,0}N_L+a_{0,1}N_R+a_{1,1}N_LN_R+a_{2,0}N_L^2+a_{0,2}N_R^2+a_{1,1}N_LN_R+... \tag{10.94}$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}a_{i,j}N_L^iN_L^j \tag{10.95}$$

Physically, transport from left to right right requires the presence of particles in the left compartment. The left-to-right current should thus vanish for $N_L = 0$, which means

$$\forall m \;\; a_{0,m} = 0. \tag{10.96}$$

Similarly, we can write

$$\begin{aligned} J^{RL} &= b_{0,0} + b_{1,0}N_L + b_{0,1}N_R + b_{1,1}N_L N_R + b_{2,0}N_L^2 + b_{0,2}N_R^2 + ... && (10.97)\\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b_{i,j}N_L^i N_L^j && (10.98) \end{aligned}$$

where, due to the same argument, we have that

$$\forall m \;\; b_{m,0} = 0. \tag{10.99}$$

If the expansion has a finite number of terms such that its maximum power of λ is M, then we expect the number of coefficients per side to scale as $M(M+1)/2$. For two even compartments in contact, we have λ^{2n} and $\lambda^{2n}\tanh(\lambda)^2$ contributions for even powers and $\lambda^{2n+1}\tanh(\lambda)$ contributions for odd powers, which results in $3M/2$ constraints for even M (and $3(M-1)/2+1$ for odd M). If $M > 2$, there are more coefficients than constraints and consequently there can be asymmetric contributions to the black box current. This is somewhat reminiscent of Smulochovski's trapdoor [8], which supposedly can be opened from one side to facilitate transport in either direction. We will now consider a trapdoor-like molecular machine.

Molecular Trapdoors

As a simple example of a molecular trapdoor, consider a transport enzyme E with a graph given by Fig. 10.7

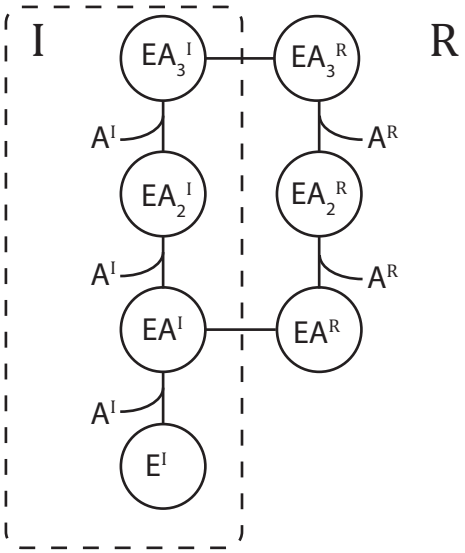


Figure 10.7: A small molecular trapdoor network that performs mod 2 transport. A transport process can only proceed if there is at least one A molecule in I, such that the state EA^I can be reached.

We suppose that each binding step is unfavorable, such that $P(E) \to 1$. Furthermore, we suppose that the slowest step is the transition between two EA_3 states, such that the master equation reaches a stationary distribution. Finally, we suppose that the transition $E \to EA \to EA_2 \to EA_3$ is

rapid with respect to particle exchange with the reservoir, but transition between two EA_3 states is slow. If timescales are separated in this way, we have an LR current

$$J^{LR} = k_h N_L(N_L-1)(N_L-2), \tag{10.100}$$

and an RL current

$$J^{RL} = k_h N_L N_R(N_R-1). \tag{10.101}$$

Using Eqs. (10.23) and (10.26) we can see that for any pair of odd and even mod 2 compartments, we have no net current, as

$$\langle J^{LR}\rangle = \langle J^{RL}\rangle. \tag{10.102}$$

However, we do have a trapdoor mechanism, in the sense that no current can occur without the initial binding of a particle in the left compartment. Let us now generalize this simplest of trapdoors, by considering a black box which first requires binding by a particle in the left compartment, according to

$$J^{LR} = N_L f(N_L), \tag{10.103}$$
$$J^{RL} = N_L g(N_R). \tag{10.104}$$

Let us now suppose we can write f and g as power laws

$$f(N_L) = a_0 + a_1 N_L + a_2 N_L^2 + ... \tag{10.105}$$
$$f(N_R) = b_0 + b_1 N_R + b_2 N_R^2 + ... \tag{10.106}$$

for which we can directly write $b_0 = 0$ since RL transport cannot occur in absence of R particles.

If the left compartment is even, we can use (10.23) to write for the average LR-current

$$\langle J^{LR}\rangle = \sum_{n=0}^{\infty} \lambda^{1+2n} \tanh(\lambda) \sum_{i=2n+1}^{\infty} S(i,2n+1)a_{i-1} + \sum_{n=1}^{\infty} \lambda^{2n} \sum_{i=2n}^{\infty} S(i,2n)a_{i-1}. \tag{10.107}$$

If the right compartment is odd, we can use (10.26) and (10.23) to write for the RL-current

$$\langle J^{RL}\rangle = \sum_{n=0}^{\infty} \lambda^{2+2n} \sum_{i=2n+1}^{\infty} S(i,2n+1)b_i + \sum_{n=1}^{\infty} \lambda^{2n+1} \tanh(\lambda) \sum_{i=2n}^{\infty} S(i,2n)b_i.$$

Similarly, if the right compartment is even, we can use (10.23) to write

$$\langle J^{RL}\rangle = \sum_{n=0}^{\infty} \lambda^{2+2n} \tanh^2(\lambda) \sum_{i=2n+1}^{\infty} S(i,2n+1)b_i + \sum_{n=1}^{\infty} \lambda^{2n+1} \tanh(\lambda) \sum_{i=2n}^{\infty} S(i,2n)b_i. \tag{10.108}$$

If we now match odd powers of (10.107),(10.108) and (10.108) (to annul the current), we obtain a vanishing coefficient for the first power λ^1

$$\sum_{i=1}^{\infty} a_{i-1} = 0, \tag{10.109}$$

and for higher powers λ^{2n+1}

$$\sum_{i=2n+1}^{\infty} S(i,2n+1)a_{i-1} = \sum_{i=2n}^{\infty} S(i,2n)b_i = \gamma_{2n+1}, \tag{10.110}$$

where γ_{2n+1} is a constant. For even powers λ^{2n}, we obtain

$$\sum_{i=2n}^{\infty} S(i,2n)a_{i-1} = \sum_{i=2n-1}^{\infty} S(i,2n-1)b_i = 0. \tag{10.111}$$

Let us again write eq. (10.89) for the modified vector $\boldsymbol{a}' = (a_0, a_1, a_2, ...)^T$ and $\boldsymbol{\gamma} = (0,0,\gamma_3,0,\gamma_5,..)^T$, to obtain

$$\boldsymbol{S}\cdot\boldsymbol{a}' = \boldsymbol{\gamma}. \tag{10.112}$$

For which we know that the solution can be written as

$$\boldsymbol{a}' = \sum_{n=1}^{M} \beta_{2n+1}\boldsymbol{c}_{2n+1} \tag{10.113}$$

Similarly, let us write for $\boldsymbol{b} = (b_1, b_2, b_3, ...)^T$ and $\boldsymbol{\gamma}' = (0, \gamma_3, 0, \gamma_5, ...)$ that

$$\boldsymbol{S}\cdot\boldsymbol{b} = \boldsymbol{\gamma}'. \tag{10.114}$$

Then, it follows that

$$\boldsymbol{b} = \sum_{n=1}^{M} \theta_{2n}\boldsymbol{c}_{2n}, \tag{10.115}$$

where each θ_{2n} is a constant, with the property

$$\theta_{2n} = \beta_{2n+1} \tag{10.116}$$

Upon substitution in Eqs. (10.105) and (10.106) we obtain the general form for a mod 2 trapdoor

$$J^{LR} = \sum_{n=1} \theta_{2n}(N_L)_{2n+1} \tag{10.117}$$

$$J^{RL} = \sum_{n=1} \theta_{2n}N_L(N_R)_{2n} \tag{10.118}$$

By the exact same arguments, we can derive a similar result for a mod m trapdoor, which yields

$$J^{LR} = \sum_{n=1} \theta_{mn}(N_L)_{mn+1} \tag{10.119}$$

$$J^{RL} = \sum_{n=1} \theta_{mn}N_L(N_R)_{mn}. \tag{10.120}$$

We can generalize the trapdoor binding step to any particular sequence of binding events by L and R particles to yield a binding step proportional to $(N_L)_k(N_R)_l$. Using the same derivation, it is then found for a modm trapdoor, that

$$J^{LR} = \sum_{n=1} \theta_{mn}(N_L)_{mn+k}(N_R)_l, \tag{10.121}$$

$$J^{RL} = \sum_{n=1} \theta_{mn}(N_L)_k(N_R)_{mn+l}. \tag{10.122}$$

The most general mod m trapdoor that we can then describe is a linear combination of all binding steps parametrized by k, l

$$J^{LR} = \sum_{n,k,l} \theta_{mn,k,l}(N_L)_{mn+k}(N_R)_l, \tag{10.123}$$

$$J^{RL} = \sum_{n,k,l} \theta_{mn,k,l}(N_L)_k(N_R)_{mn+l}. \tag{10.124}$$

Note that every *LR* term $(N_L)_{mn+k}(N_R)_l$ is associated with an *RL* term $(N_L)_k(N_R)_{mn+l}$. This observation can be readily understood when one performs a cycle decomposition[9] on a simple enzyme graph described by mass-action. However, the black box need not be an enzyme, it merely needs to be an object with polynomial current propensities.

A minimal argument

It should be stressed that we started with a 'black box'. By applying the second law, all current propensities of another form than Eq. (10.92) became forbidden. These current propensities follow a falling-factorial form, which is encountered in mass-action descriptions of chemistry. This can be understood when considering the Gibbs free energy change between any pair of states $\boldsymbol{n}, \boldsymbol{n}'$. This may happen in an arbitrary and rather elaborate way. However, (in the absence of hidden reactions and currents) the corresponding ΔG should still follow from Eq. (10.36). We must therefore expect the current propensities to be of a functional form that respects

$$\frac{J_{\boldsymbol{n}\to\boldsymbol{n}'}}{J_{\boldsymbol{n}'\to\boldsymbol{n}}} = \frac{z(\boldsymbol{n}')}{z(\boldsymbol{n})}\exp(-\Delta G^\circ). \tag{10.125}$$

This relation is true for all our examples, including the general asymmetric trapdoors (Eqs. (10.123), (10.124)). A similar result can be derived hold for other discrete conserved quantities, and this severely constrains the allowed functional forms of their currents.

In our derivation, we imposed a simplifying separation of timescales. A more complete treatment of transport through a black box will need to address the question of timescales, especially when perturbed from equilibrium. An interesting outlook will be to see how insights for black boxes can be united with thermodynamically consistent formulations of mesoscopic transport built up from a microscopic description[10, 11, 12].

10.2 Appendix: Toy Formose

Toy formose is a simplified scheme of a popular example of autocatalysis in the literature: the formose reaction[13]. In the genuine formose reaction, formaldehyde CH_2O (referred to as C1), reacts in an aldol reaction with an enolate. Under basic conditions, glycolaldehyde $OHCCH_2OH$ (C2), can be deprotonated to form such an enolate and subsequently react with C1 form glyceraldehyde

$$\mathrm{OHCCH_2OH} \overset{\mathrm{OH^-}}{\rightleftharpoons} \mathrm{OHCCHOH^-}, \tag{10.126}$$

$$\mathrm{OHCCHOH^- + H_2CO} \rightleftharpoons \mathrm{OHCCH(OH)CHOH^-}. \tag{10.127}$$

Glyceraldehyde does not form a very suitable enolate, it has to convert to 1,3-dihydroxyacetone (C3). It can do so by base-catalyzed tautomerization, but it was recently shown that the catalytic amount of Ca^{2+} that is added facilitates a hydride shift[14, 15], which can be considerably faster

$$\mathrm{OHCCH(OH)CH_2OH} \overset{\mathrm{Ca^{2+},OH^-}}{\rightleftharpoons} \mathrm{HOH_2CCOCH_2OH} \tag{10.128}$$

In turn, C3 then reacts to form a tetrulose, which can interconvert to form an aldose (C4). This aldose can, perform a base-catalyzed retro-aldol reaction, yielding two C2 molecules

$$\mathrm{OHCCH(OH)CH(OH)CH_2OH} \overset{\mathrm{OH^-}}{\rightleftharpoons} 2\mathrm{OHCCH_2OH}. \tag{10.129}$$

If interconversion is fast, or if we do not seek a detailed description, we can write a simplified scheme, which we will refer to as Toy Formose:

$$\mathrm{C_1 + C_2} \rightleftharpoons \mathrm{C_3}, \tag{10.130}$$

$$\mathrm{C_1 + C_3} \rightleftharpoons \mathrm{C_4}, \tag{10.131}$$

$$\mathrm{C_4} \rightleftharpoons 2\mathrm{C_2}. \tag{10.132}$$

Note that formaldehyde (C1) cannot form an enolate, which implies that pure C1 is not expected to perform the formose reaction. In practice, C2 and other formose species are present as a trace impurity in commercial formaldehyde [16].

It is important to appreciate that Toy Formose is an idealization that is far removed from the complex chemical network of typical formose reaction. First of all, the addition of C1 does not stop at C4, further additions are possible, forming C5, C6, C7 etc. each with more stereoisomer forms and tautomers. These larger species can also afford additional retro-aldol pathways, e.g. C6 may split in 2 C3, or C4 + C2.

We should also consider that reactive aldehyde groups in formose are often in a less reactive gem-diol form, which rapidly equilibrates in water with base

$$CH_2O + H_2O \rightleftharpoons CH_2(OH)_2. \tag{10.133}$$

Typically, 99.95% of C1 is in the gem-diol form, for larger aldehydes, the ratio gem-diol/aldehyde is often around 20:1 [17]. Starting at C4, the aldehyde group is reversibly removed because appreciable conversion occurs towards the cyclic form, e.g. erythrose is in a cyclic form 90 % of the time. Typically, longer sugars have considerably stabilized cyclic forms (the aldehyde fraction in glucose is estimated at 100 ppm [18]), whereas only the open form contains the carbonyl group necessary to perform the retro-aldol reaction.

Furthermore, formaldehyde reversibly forms polymeric species, poly(oxymethylenes), a reaction that is often represented as a dehydration of the gem-diol

$$2CH_2(OH)_2 \rightleftharpoons HOCH_2OCH_2OH + H_2O, \tag{10.134}$$

$$HO(CH_2O)nH + CH_2(OH)_2 \rightleftharpoons HO(CH_2O)_{n+1}H + H_2O. \tag{10.135}$$

The reaction is catalyzed by base [19], which from a mechanistic viewpoint is easier to see of we consider the base-catalyzed reaction with an aldehyde

$$[HO(CH_2O)_n]^- + CH_2O \rightleftharpoons [HO(CH_2O)_{n+1}]^-. \tag{10.136}$$

Such reactions can also form cyclic species, and larger species (e.g. C2) reversibly form dimers (in fact, C2 is sold as a dimer) and oligomers by such pathways.

Formaldehyde and larger aldehydes also engage in the irreversible Cannizzaro reaction (like the hydride shift in formose, this reaction is catalyzed by Ca^{2+})[17]. This reaction consumes an aldehyde and another carbonyl species, along with base, to form an alkanoate and an alcohol. For C1, we can e.g. write

$$CH_2O + CH_2(OH)_2 + OH^- \rightarrow HCOO^- + H_3COH + H_2O, \tag{10.137}$$

A visual indicator that the formose reaction has proceeded considerably , is the onset of a 'yellowing point', in which the initially colorless mixture obtains a yellow hue. This can be interpreted as the formation of products with conjugated double bonds, which can happen due to base-catalyzed dehydration,

$$RHOCHCHOHCHO \xrightarrow{OH^-} RCH{=}COHCHO + H_2O \tag{10.138}$$

The combination of these dynamic equilibria and irreversible reactions make formose a very interesting reaction to study. However, it performs poorly as a specific synthetic route to a single product with high yield, and its prebiotic relevance has been questioned in the context of ribose synthesis. In conceptual discussions on autocatalysis, formose is often represented as a simple example of autocatalysis. Formose is 'simple', only in the sense that we can use small building blocks (C1, C2), to acquire autocatalysis and that the simplest autocatalytic cycle can be condensed to the elegant Toy Formose representation.

In practice, the aldol and retro-aldol reactions in formose generate a much wider diversity of species, and the typical reaction conditions (high pH, catalytic Ca^{2+}, in water) favor a rich amount

of other reactions. To distinguish the hypothetical scheme (10.130)-(10.132) from the more realistic network observed in practice, we will refer to the former as Toy Formose. Toy Formose is a useful Toy model for the study of solitary autocatalytic reactions (Ch.5). We use the Toy model in several chapters, notably to explore the effects of minor decorations of the network.

10.2.1 Exchange between coupled droplets

Let us now write a simple model for Toy Formose in droplets, that are coupled through diffusion. For the chemical reactions in a droplet labeled i, we can write reaction rates

$$R^i_{1+2} = k_1^+ N^i_{C1} N^i_{C2}/V^i - k_1^- N^i_{C3}, \tag{10.139}$$
$$R^i_{1+3} = k_1^+ N^i_{C1} N^i_{C3}/V^i - k_1^- N^i_{C4}, \tag{10.140}$$
$$R^i_{2+2} = k_2^+ N^i_{C2} N^i_{C2}/V^i - k_2^- N^i_{C4}. \tag{10.141}$$

Where we attribute the same forward rate for the incorporation of $C1$ by $C2$ and $C3$. As noted before, in genuine formose these species are in dynamic equilibrium between tautomers and hydrates, which means effective rate constants for different steps can differ considerably[†] In our present discussion of Toy Formose such effects are neglected.

Exchange between compartment i and j is only considered non-negligible for solvent (H_2O), C_1 and C_2 [‡]

$$J^{i,j}_{C1} = P_{C1}\left(N^i_{C1}/V^i - N^j_{C1}/V^j\right), \tag{10.142}$$
$$J^{i,j}_{C2} = P_{C2}\left(N^i_{C2}/V^i - N^j_{C2}/V^j\right), \tag{10.143}$$
$$J^{i,j}_{H2O} = P_{H2O}\left(N^i_{H2O}/V^i - N^j_{H2O}/V^j\right). \tag{10.144}$$

Where P_k denotes the permeability for a species k, which contains contributions for partition equilibria, diffusion length and geometry to describe overall transfer between compartments. An approximation that works much better than it should is

$$P_k \approx \frac{D_k K \bar{A}}{L}, \quad \bar{A} = \frac{A^I A^{II}}{A^I + A^{II}} \tag{10.145}$$

where A^i denotes the surface of compartment i. A thermodynamically consistent expression for the exchange current should vanish when the chemical potentials are equal, and $\bar{A}$ is the simplest geometrical correction that respects this constraint.[§]

The volume for a compartment i is

$$V^i = \sum_k N^i_k v_k, \tag{10.146}$$

where v_k is the molar volume of a species k.

From which we can then write differential equations on the level of a pair of compartment I

[†]Experimental fitting of these parameters is currently under way. Performing these corrections is indispensable for a comprehensive model of genuine formose.

[‡]This was found experimentally in exchange experiments in the LBC lab, which will be detailed in an upcoming publication

[§]A more thorough treatment of the geometry and diffusion should take path integrals to account for all trajectories between compartments and their weights.

and II. For compartment I we then have

$$\frac{dN^I_{C1}}{dt} = -R^I_{1+2} - R^I_{1+3} - J^{I,II}_{C1}, \tag{10.147}$$
$$\frac{dN^I_{C2}}{dt} = -R^I_{1+2} - 2R^I_{2+2} - J^{I,II}_{C2}, \tag{10.148}$$
$$\frac{dN^I_{C3}}{dt} = -R^I_{1+3} + R^I_{1+2}, \tag{10.149}$$
$$\frac{dN^I_{C4}}{dt} = R^I_{1+3} + R^I_{2+2}, \tag{10.150}$$
$$\frac{dN^I_{\mathrm{H_2O}}}{dt} = -J^{I,II}_{\mathrm{H_2O}}. \tag{10.151}$$

and similarly, for compartment II we have

$$\frac{dN^{II}_{C1}}{dt} = -R^{II}_{1+2} - R^{II}_{1+3} + J^{I,II}_{C1}, \tag{10.152}$$
$$\frac{dN^{II}_{C2}}{dt} = -R^{II}_{1+2} - 2R^{II}_{2+2} + J^{I,II}_{C2}, \tag{10.153}$$
$$\frac{dN^{II}_{C3}}{dt} = -R^{II}_{1+3} + R^{II}_{1+2}, \tag{10.154}$$
$$\frac{dN^{II}_{C4}}{dt} = R^{II}_{1+3} + R^{II}_{2+2}, \tag{10.155}$$
$$\frac{dN^{II}_{\mathrm{H_2O}}}{dt} = J^{I,II}_{\mathrm{H_2O}}. \tag{10.156}$$

We study droplet exchange behavior by the following thought experiment: we start with a droplet I and II, with initial volumes V^{I}, V^{II}, and initial molecule numbers $\boldsymbol{N}^I(0), \boldsymbol{N}^{II}(0)$. After a time τ has passed, droplet I is split (one of the split droplets is removed), and droplet II is replaced with a copy of its initial state:

$$\boldsymbol{N}^I(\tau^+) = \frac{\boldsymbol{N}^I(\tau^-)}{2}, \tag{10.157}$$
$$\boldsymbol{N}^{II}(\tau^+) = \boldsymbol{N}^{II}(0). \tag{10.158}$$

in Sec. 3.4.7, (and repeated here) this protocol is performed using concentrations $\boldsymbol{C} = \{C_{C1}, C_{C2}, C_{C3}, C_{C4}, C_{H_2O}\}$, with $\boldsymbol{C}^I = \{3.0, 0.2, 0, 0, 48.2\}$, $\boldsymbol{C}^{II} = \{3.0, 0.0, 0, 0, 49.0\}$, and $V^I = 110pL$, $V^{II} = 70pL$.

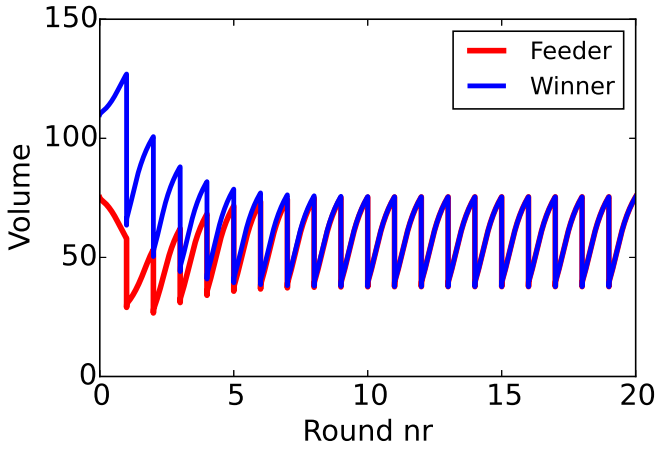


Figure 10.8: Periodic volume growth of an autocatalytic 'winner' compartment (red) and a 'feeder' compartment after one round of contact with a 'winner' (blue). Compartments are periodically placed in contact with a new feeder compartment.

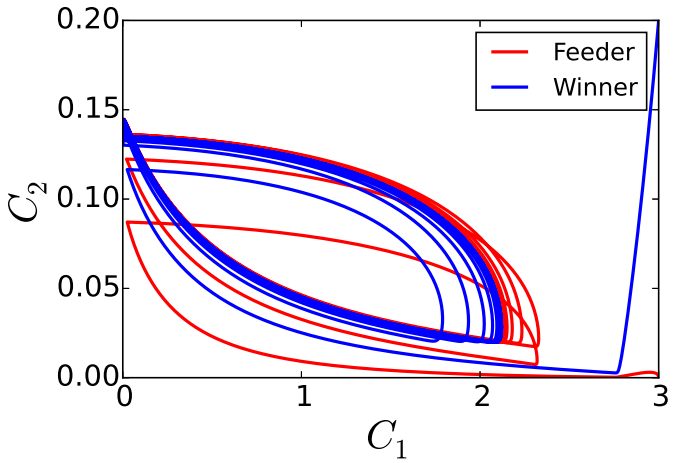


Figure 10.9: C_1 concentration as a function of C_2 concentration from $t = 0$ to 20τ, for the 'winner' compartment (blue) and a 'feeder' compartment (red). Both converge to the same limit cycle.

Rates for the toy model were $k_1^+ = 10.0\text{min}^{-1}$, $k_2^+ = 5.0\text{min}^{-1}$, $k_2^- = 0.1\text{min}^{-1}$, $P_{H_2O} = 0.66\text{min}^{-1}$, $P_{C_1} = 8.7$, $P_{C_2} = 0.4\text{min}^{-1}$.¶

As can be seen by the blue line in Fig. 10.8 and 10.9, a droplet that starts out as a 'feeder' droplet, can be propagated in subsequent rounds to become a 'winner' droplet. This can be attributed to the nonzero transport rate of C_2, which ensures that the reaction can also take off in droplet II.

10.3 Appendix: Purification in chemical networks

In Ch.4, we provided a physical link between 'information' and chemical composition, by showing how we can construct an engine that erases a pure enantiomer configuration to extract work. By running in reverse, work can be injected to drive a chemical purification. Such processes are at the heart of synthetic chemistry, where purifications are performed through a wide variety of processes. Here, we will be interested in how purification can be achieved by the network itself.

¶values based on preliminary measurements of diffusive transport.

The problem of classical synthetic limits will first be discussed on the level of a single reaction competing with side reactions. Subsequently, we will consider how these limits can be overcome on the level of a larger reaction network. Our objective is to give a small taste of the ways in which these networks differ from a single reaction and why that is pertinent for abiogenesis.

The origins of life is often perceived as a synthetic challenge that is addressed on the single-reaction level[20], with the notable exception of autocatalysis. A systems-level approach[21] is expected to be transformative to this approach, as it has already been for other branches of chemistry and industry.

10.3.1 The single-reaction case

Yield, purity, error

Suppose a synthesis step happens with a fractional yield

$$\eta = \frac{N_{final}}{N_{max}}, \tag{10.159}$$

where N_{final} is the final amount of product obtained, and N_{max} the theoretical maximum obtainable, based on stoichiometry. In chemistry, the chemical yield $Y(\%)$ is often expressed as a percentage

$$Y = \eta \cdot 100\%. \tag{10.160}$$

There can be various contributions for having $\eta < 1$, such as i) competing side reactions consume the substrate ii) thermodynamic conditions do not permit a complete conversion iii) insufficient reaction time iv) kinetic trapping of intermediates. The latter is particularly problematic for self-assembly of objects from a diversity of substrates, and has been referred to as a yield catastrophe[22][‖]. In linear synthesis, the product of one step is the reactant of the next step. After n steps of linear synthesis, only a fraction

$$\eta^{(n)} = \Pi_{i=1}^{n} \eta_i \tag{10.161}$$

of the starting material in step 1 is retained, underlining the need for a small number of efficient synthesis steps. The objective is to obtain the right product in high yield, but also to minimize the amount of undesirable side-products, which may be very detrimental.

In the following, we will consider the problem of forming the right product, from the viewpoint of chemical networks. The simplest network we will consider is

$$\mathrm{R} \overset{1}{\rightleftharpoons} \mathrm{S} + \mathrm{X} \overset{2}{\rightleftharpoons} \mathrm{W}. \tag{10.162}$$

Letting R denote the 'right' product, and W the 'wrong' product, we introduce the error fraction ϕ as the fraction of the wrong product produced

$$\phi = \frac{N_W}{N_W + N_R}, \tag{10.163}$$

which is expressed in terms of the total amount of W and R produced throughout the process.

In steady-state regimes, we can define a similar quantity, called the error rate ζ, which is the ratio of [24] production rates, or chemical currents J_W, J_R

$$\zeta = \frac{J_W}{J_W + J_R}, \tag{10.164}$$

[‖]Such a catastrophe becomes most pronounced in small systems: unfinished species mutually trap each other's scarce building blocks. Such catastrophes highlight the need for error-correcting machinery and chaperones[23] to overcome them.

For long times or if the system starts in a steady-state regime, we can write $\zeta = \phi$, since $N_W \to J_W t$ and $N_R \to J_R t$.

If the substrate S is completely consumed by these pathways, ζ can be related to the yield as $\eta = 1 - \phi$. For stationary currents, we can then write a 'rate yield' η_J based on the currents

$$\eta_J = 1 - \zeta = \frac{J_R}{J_W + J_R}. \tag{10.165}$$

Thermodynamic control vs kinetic control

The reaction network (10.162) corresponds to a very common situation in chemistry, in which two reactions compete for the consumption of a product. Oftentimes, these reactions may involve the same reactants, leading to a different products (see Fig. 10.10). A substantial number of named rules in chemistry dictate the preferential product in such a situation (e.g. Alder endo rule, Hofmann's rule, Zaitsev's rule, Bredt's rule, Woodward-Hoffmann rule, Markovnikov's rule, etc. [25, 17]).

For such a process, two limiting regimes can be identified: i) thermodynamic control, and ii) kinetic control.

Thermodynamic control

Thermodynamic control corresponds to the situation where species are acquired in accord with their Boltzmann weights, and thus follow their equilibrium constants. Supposing a solution of volume V in a closed system at temperature T, we can write

$$N_{\mathrm{R}} = K_1'(V) N_{\mathrm{S}} N_{\mathrm{X}}, \qquad N_{\mathrm{W}} = K_2'(V) N_{\mathrm{S}} N_{\mathrm{X}}. \tag{10.166}$$

where $K'(V)$ is an equilibrium constant in which all volume dependencies are absorbed. The ratio of equilibrium constants then verifies

$$\frac{K_2'}{K_1'} = \exp(-\beta(\mu_R^\circ - \mu_W^\circ)), \tag{10.167}$$

with μ_R° (resp. μ_R° the free energy of formation of R (resp. W), which is dominated by the thermodynamically stable product. The error fraction is then

$$\phi = \frac{1}{1 + \exp(-\beta(\mu_R^\circ - \mu_W^\circ))}. \tag{10.168}$$

As long as $\mu_W^\circ \gg \mu_R^\circ$, thermodynamic control is desirable. This is, however, not generally the case and often this regime may not provide the required level of purity.

Kinetic control

A reaction is under kinetic control, when its product composition is not determined by thermodynamic stability, but the relative rates of product formation. For this to be efficient in the network (10.162), products should not have the time to convert back to the reactant. This means reactions should be sufficiently irreversible, such that the timescales of forward reactions and backwards reactions are well-separated. We can then write for the rates of formation

$$\frac{dN_R}{dt} = k_R^+ N_{\mathrm{X}} N_{\mathrm{S}} - k_R^- N_{\mathrm{R}}, \qquad \frac{dN_W}{dt} = k_W^+ N_{\mathrm{X}} N_{\mathrm{S}} - k_W^- N_{\mathrm{W}}. \tag{10.169}$$

If the concentration of X is fixed (due to being buffered or overabundant), effective rate constants can be defined $\kappa_R = k_R^+ N_{\mathrm{X}}$, $\kappa_W = k_W^+ N_{\mathrm{X}}$. The typical time for an S molecule to be converted to an R (resp. W) is then $\tau_R^+ = 1/\kappa_R$ (resp. $\tau_W^+ = 1/\kappa_W$). Our reactions must occur on a shorter timescale than its reverse reaction

$$\tau_R^- \gg \tau_R^+, \quad \tau_W^- \gg \tau_W^+, \tag{10.170}$$

which changes the reaction scheme to a pair of irreversible reactions

$$\mathrm{R} \xleftarrow{1} \mathrm{S} + \mathrm{X} \xrightarrow{2} \mathrm{W}, \tag{10.171}$$

provided the reaction is performed on a timescale $\tau_r \ll \tau_R^-, \tau_W^-$. In this limit, the error fraction becomes

$$\phi = \frac{1}{1 + \kappa_R/\kappa_W} \tag{10.172}$$

By changing temperature T, experimental timescale τ_r, or the concentration of X, our model reaction can shift from thermodynamic control to kinetic control. The choice of reactants can be adapted to access the desired regime. For example, the formation of an enolate by deprotonation can occur swiftly with a small base, permitting the proton to go back-and-forth and the substrate to be deprotonated at different sites. A large, sterically hindered base will deprotonate more slowly, and will bias the most accessible deprotonation sites under kinetic control. In Fig 10.10 an example of these two attacks is given.

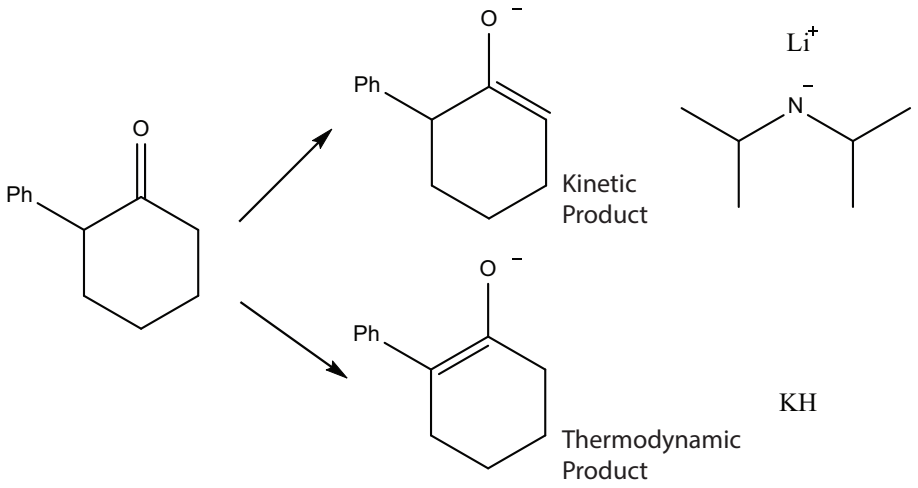


Figure 10.10: A kinetic product (top) and a thermodynamic product (bottom), obtained by abstracting a proton with a strong base. The proton that needs to be abstracted for the kinetic product is easily accessible, whereas the proton for the thermodynamic product is not. A large, bulky base (LDA, shown top right) almost exclusively gives the kinetic product (-70 C in THF). A small base, H^-, can access this proton to almost exclusively give the thermodynamic product (in THF).[26, 27, 17].

10.3.2 Kinetic discrimination

A closely related problem in kinetically controlled reactions is the discrimination between chemically similar reactants by the same reaction pathway. A discrimination process of particular prominence is the separation of two enantiomers via some asymmetric reaction pathway (e.g. via a chiral catalyst or reactant). Such a process of separating enantiomers is referred to as 'kinetic resolution' in the organic chemistry literature. Since enantiomers are similar in many regards, their resolution is challenging and this provides an elegant starting point for the study of kinetic discrimination.

Starting with a racemic mixture, consider the pair of reactions

$$\mathrm{D} + \mathrm{X} \xrightarrow{1} \mathrm{D}^*, \tag{10.173}$$

$$\mathrm{L} + \mathrm{X} \xrightarrow{2} \mathrm{L}^*. \tag{10.174}$$

In the following, we will designate D^* to be the desired product. To quantify the degree of discrimination between the two substrates, we introduce the error factor χ

$$\chi = \frac{k_L}{k} = \frac{k_L}{k_L + k_D}, \tag{10.175}$$

where k is the sum of rate constants. Fixing the concentration of X, the abundances of D^* and L^* are the solution of first-order differential equations

$$N_{D^*} = N_D^0(1-\exp(-k_D t)), \tag{10.176}$$
$$N_{L^*} = N_L^0(1-\exp(-k_L t)). \tag{10.177}$$

With N_D^0 (N_L^0) the initial amount of D (L). Let us from here on consider the case $N_D^0 = N_L^0$. Injecting these solutions in Eq. (10.163), the error fraction obeys

$$\phi = \frac{1-\exp(k\chi t)}{2-\exp(k\chi t)-\exp(k(1-\chi)t)}. \tag{10.178}$$

On short timescales, $t \ll 1/k_D$, the exponentials can be approximated linearly as $\exp(x) \approx 1+x$, to give

$$\phi \approx \chi. \tag{10.179}$$

The fractional yield η then becomes

$$\eta = \frac{2-\exp(-k_D t))-\exp(-k_L t))}{2} \tag{10.180}$$

While the yield η progressively increases, the error fraction ϕ increases, because the exhaustion of D slows down the desired reaction. Due to its much slower consumption of L, the side-reaction forming $L\cdot$ is exhausted in a much later stage. For kinetic resolution, a delicate trade-off thus exists between yield (Y) and purity and it quickly becomes imperative to have extremely low error factors, e.g. through chiral organometallic complexes or enzymatic catalysis.

10.3.3 Chemical networks beyond the single reaction

Dynamic kinetic resolution

Dynamic kinetic resolution is a recent improvement on kinetic resolution [28], in which a fast reaction is added that interconverts the products among which the discrimination is acting. For a pair of enantiomers, this is a racemization reaction, which is a fairly easy reaction to catalyze. A simplified reaction scheme is then

$$L^* \xleftarrow{1} L \underset{}{\overset{rac}{\rightleftharpoons}} D \xrightarrow{2} D^*. \tag{10.181}$$

The racemization reaction is the fastest reaction: $k_{rac} \gg k_D, k_L$. Since $\mu_L^\circ = \mu_D^\circ$, the racemization reaction ensures that $N_L = N_D$. The reactants L and D can directly be mapped to the reactant S in kinetic control, since the equations are equivalent:

$$\frac{dN_{D^*}}{dt} = k_D^+ N_X N_D - k_D^- N_D, \qquad \frac{dN_{L^*}}{dt} = k_L^+ N_X N_L - k_L^- N_L. \tag{10.182}$$

it follows that DKR (10.181) is equivalent to kinetic control in the conversion of a single species (10.171). For both, the error fraction does not depend on the yield, and obeys $\zeta = \chi$ at all times $t > 0$. In this sense, DKR is strictly better than kinetic resolution, for which the error fraction increases over time. In addition the yield is strictly better: kinetic resolution can at most acquire all

of the substance D to the desired substance D$*$. In DKR, the other half consisting of L can now also be converted to that substance.

This provides a first example of how a larger reaction networks can overcome the fundamental limits of single reactions**. Interestingly, this only required one extra reaction.

We will now turn to some examples of proofreading.

Degradative Proofreading

Our first instance of proofreading concerns the addition of a second reaction step, and a degradation step

$$\mathrm{S} \xrightarrow{1a} \mathrm{R}^* \xrightarrow{1b} \mathrm{R}, \qquad \mathrm{R}^* \xrightarrow{1x} \emptyset \tag{10.183}$$

$$\mathrm{S} \xrightarrow{2a} \mathrm{W}^* \xrightarrow{2b} \mathrm{W}, \qquad \mathrm{W}^* \xrightarrow{2x} \emptyset \tag{10.184}$$

An S molecule goes through a first reaction step to yield either R^* or W^*, yielding the production currents

$$J_{R^*} = k_{R*} N_S, \tag{10.185}$$

$$J_{W^*} = k_{W*} N_S. \tag{10.186}$$

From this, an error rate $\zeta_1 = k_{W^*}/(k_{W^*} + k_{R^*})$ is obtained. In turn, an R^* or W^* molecule is subjected to either a degradation

$$J_{R^*\emptyset} = k_\emptyset N_{R^*}, \tag{10.187}$$

$$J_{W^*\emptyset} = k_\emptyset N_{W^*}, \tag{10.188}$$

or a propagation to a final product,

$$J_R = k_R N_{R^*}, \tag{10.189}$$

$$J_W = k_W N_{W^*}. \tag{10.190}$$

We can then write the dynamics for $N_{\mathrm{R}^*}, N_{\mathrm{W}^*}$:

$$\frac{dN_{\mathrm{R}^*}}{dt} = J_{\mathrm{R}^*} - J_{\mathrm{R}^*\emptyset} - J_{\mathrm{R}}, \qquad \frac{dN_{\mathrm{W}^*}}{dt} = J_{\mathrm{W}^*} - J_{\mathrm{W}^*\emptyset} - J_{\mathrm{W}}. \tag{10.191}$$

Let us consider now a system where S is chemostatted at an abundance $\bar{N}_\mathrm{S}$, such that $N_{\mathrm{R}^*}, N_{\mathrm{W}^*}$ reach a steady-state:

$$N_{\mathrm{R}^*} = \frac{k_{R*} N_\mathrm{S}}{k_\mathrm{R} + k_\emptyset}, \qquad N_{\mathrm{W}^*} = \frac{k_{W*} N_\mathrm{S}}{k_\mathrm{W} + k_\emptyset} \tag{10.192}$$

The yield rate $\zeta^{(2)}$ for the two reactions combined is then

$$\zeta^{(2)} = \frac{J_W}{J_W + J_R} = \frac{k_\mathrm{W} \frac{k_{W*}}{k_\mathrm{W} + k_\emptyset}}{k_\mathrm{W} \frac{k_{W*}}{k_\mathrm{W} + k_\emptyset} + k_\mathrm{R} \frac{k_{R*}}{k_\mathrm{R} + k_\emptyset}}, \tag{10.193}$$

for which we are interested in two limiting regimes:

$$\zeta^{(2)} = \left(\frac{k_{W*}}{k_{W*} + k_{R*}} \right), \quad k_\emptyset \ll k_\mathrm{R}, k_\mathrm{W} \tag{10.194}$$

$$\zeta^(2) = \left(\frac{k_W k_{W*}}{k_W k_{W*} + k_\mathrm{R} k_{R*}} \right)^2, \quad k_\emptyset \gg k_\mathrm{R}, k_\mathrm{W}. \tag{10.195}$$

**Evidently, the racemization reaction should only act on substrates, and not on the products of the irreversible reactions, otherwise the effect of kinetic resolution will be undone. Alternatively, one can remove the products from the reactive phase rapidly after formation

Interestingly, when degradation is considerable (Eq. (10.195)), the selectivity can increase. In particular, if $k_{\mathrm{R}}k_{R*} \gg k_{\mathrm{W}}k_{W*}$, the latter regime approaches

$$\zeta^{(2)} \to \frac{k_{\mathrm{W}}}{k_{\mathrm{R}}}\frac{k_{W*}}{k_{R*}}, \tag{10.196}$$

which means the error is reduced by the product of two selectivities instead of one.

This principle can be repeated. For simplicity, let $k_{W*} = k_W$, $k_{R*} = k_R$. Let us now consider another pair of reactions

$$\mathrm{R} \xrightarrow{1c} \mathrm{R}^{*}*, \tag{10.197}$$

$$\mathrm{W} \xrightarrow{2c} \mathrm{W}^{*}*. \tag{10.198}$$

which are again in competition with a degradation process occurring at a rate $k_\emptyset$. The yield rate for the three reactions can then be shown to become

$$\zeta^{(3)} = \frac{J_W \frac{k_W}{k_W+k_\emptyset}}{J_W \frac{k_W}{k_W+k_\emptyset} + J_R \frac{k_R}{k_R+k_\emptyset}}. \tag{10.199}$$

which for $k_\emptyset \gg k_{\mathrm{R}} \gg k_{\mathrm{W}}$ approaches

$$\zeta^{(3)} \to \left(\frac{k_{\mathrm{W}}}{k_{\mathrm{R}}}\right). \tag{10.200}$$

More generally, $n-1$ degradation steps then lead to

$$\zeta^{(n)} \to \left(\frac{k_{\mathrm{W}}}{k_{\mathrm{R}}}\right). \tag{10.201}$$

Such a process can in principle be repeated to reach an arbitrarily low error rate. Evidently, however, there is another problem to consider: the fractional yield η will go down with every step, especially if degradation is the dominant reaction. Indeed, the initial current J_{R^*} reduced by a fraction $k_{\mathrm{R}}/(k_\emptyset + k_{\mathrm{R}})$, leading the fractional yield $\eta^(n)$ to exponentially decrease

$$\eta^{(n+1)}/\eta^{(n)} \to \left(\frac{k_{\mathrm{R}}}{k_\emptyset + k_{\mathrm{R}}}\right). \tag{10.202}$$

There is thus a tradeoff between yield for purity. Let us here remark that the processes used here do not need to be chemical reactions. Degradation can e.g. be achieved through dilution, and kinetic selection can also take place for other processes[29]. In Ref. [30] for example, kinetic resolution between ions is achieved when they traverse a pore, by their different diffusion constants.

Dissipative Recycling

Let us now turn to another simple network exploiting degradation to improve kinetic control and which recovers material via recycling. Such a system was recently discovered in experiments attempting to ligate RNA in stereoselective fashion in a prebiotic chemical setting [31]. An example of such a network for the kinetic resolution is shown in Fig. 10.11.

An activated chemical species A^* reacts with W, either to form the right product R, or the wrong product W. The products can degrade, forming the nonactive species A. Subsequently, A

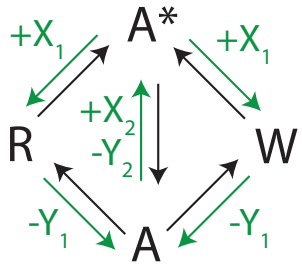


Figure 10.11: A schematic illustration of dissipative recycling. Forward cycle shown in green, along with reservoir compounds consumed X_1, X_2 or released Y_1, Y_2.

must be reactivated to form A$*$ and start the cycle anew

$$\mathrm{A}^* + \mathrm{X}_1 \overset{\mathrm{R}_1}{\rightleftharpoons} \mathrm{R}, \qquad R_{R1} = k^+_{R1} N_{\mathrm{A}^*} N_{\mathrm{X}_1} - k^-_{R1} N_{\mathrm{R}} \tag{10.203}$$

$$\mathrm{R} \overset{\mathrm{R}_2}{\rightleftharpoons} \mathrm{A} + \mathrm{Y}_1, \qquad R_{R2} = k^+_{R2} N_{\mathrm{R}} - k^-_{R2} N_{\mathrm{A}} N_{\mathrm{Y}_1} \tag{10.204}$$

$$\mathrm{A} + \mathrm{X}_2 \overset{\mathrm{A}^*}{\rightleftharpoons} \mathrm{A}^* + \mathrm{Y}_2, \qquad R_{A^*} = k^+_{A^*} N_{\mathrm{A}} N_{\mathrm{X}_2} - k^-_{A*} N_{\mathrm{A}^*} N_{\mathrm{Y}_2} \tag{10.205}$$

$$\mathrm{A}^* + \mathrm{X}_1 \overset{\mathrm{W}_1}{\rightleftharpoons} \mathrm{W}, \qquad R_{W1} = k^+_{W1} N_{\mathrm{A}^*} N_{\mathrm{X}_1} - k^-_{W1} N_{\mathrm{W}} \tag{10.206}$$

$$\mathrm{W} \overset{\mathrm{W}_2}{\rightleftharpoons} \mathrm{A} + \mathrm{Y}_1, \qquad R_{W2} = k^+_{W2} N_{\mathrm{W}} - k^-_{W2} N_{\mathrm{A}} N_{\mathrm{Y}_1} \tag{10.207}$$

This reaction network has to be supplied with reactants X_1, X_2 and produces waste products Y_1, Y_2 that are removed. This is achieved by exchange with chemostats, which fix their concentrations. At steady state, we can then write

$$R_{R_1} = R_{R_2}, \qquad R_{W_1} = R_{W_2}, \qquad R_{A*} = R_{R_1} + R_{W_1}. \tag{10.208}$$

Let us decompose R in its forward part and backward part:

$$R = R^+ - R^-, \qquad R^+_{R_1} = k^+_{R_1} N_{\mathrm{A}^*} N_{\mathrm{X}_1}, \qquad R^-_{R_1} = k^-_{R_1} N_{\mathrm{R}} \tag{10.209}$$

Choosing these concentrations such that all forward reactions are dominant: $k^+_{R_1} \gg k^-_{R_1}$, we can write

$$k^+_{R_1} N_{\mathrm{A}^*} N_{\mathrm{X}_1} = k^+_{R_2} N_{\mathrm{R}}, \tag{10.210}$$

$$k^+_{W_1} N_{\mathrm{A}^*} N_{\mathrm{X}_1} = k^+_{W_2} N_{\mathrm{W}}, \tag{10.211}$$

$$k^+_{A^*} N_{\mathrm{A}} N_{\mathrm{X}_2} = k^+_{W_2} N_{\mathrm{W}} + k^+_{R_2} N_{\mathrm{R}}. \tag{10.212}$$

The conservation law $N_R + N_A + N_{A^*} + N_W = N^\circ_A$ can then be used to yield the steady-state concentrations of R and W

$$N_R = \frac{k^+_{R_1} N_{\mathrm{X}_1}}{k^+_{R_2}} \frac{N^\circ_A}{1 + \frac{k^+_{R_1} N_{\mathrm{X}_1}}{k^+_{R_2}} + \frac{k^+_{W_1} N_{\mathrm{X}_1}}{k^+_{W_2}} + \frac{k^+_{R_1} N_{\mathrm{X}_1}}{k^+_{A^*} N_{\mathrm{X}_2}} + \frac{k^+_{W_1} N_{\mathrm{X}_1}}{k^+_{A^*} N_{\mathrm{X}_2}}}, \tag{10.213}$$

$$N_W = \frac{k^+_{W_1} N_{\mathrm{X}_1}}{k^+_{W_2}} \frac{N^\circ_A}{1 + \frac{k^+_{R_1} N_{\mathrm{X}_1}}{k^+_{R_2}} + \frac{k^+_{W_1} N_{\mathrm{X}_1}}{k^+_{W_2}} + + \frac{k^+_{R_1} N_{\mathrm{X}_1}}{k^+_{A^*} N_{\mathrm{X}_2}} + \frac{k^+_{W_1} N_{\mathrm{X}_1}}{k^+_{A^*} N_{\mathrm{X}_2}}}. \tag{10.214}$$

The error rate is then

$$\zeta = \frac{\frac{k^+_{W_1}}{k^+_{W_2}}}{\frac{k^+_{R_1}}{k^+_{R_2}} + \frac{k^+_{W_1}}{k^+_{W_2}}} = \frac{1}{1 + \frac{k^+_{W_2} k^+_{R_1}}{k^+_{W_1} k^+_{R_2}}}. \tag{10.215}$$

There are two reaction steps in which discrimination is possible. For the first, we can write the hypothetical error for the processes under kinetic control

$$\chi = \frac{k^+_{W_1}}{k^+_{R_1} + k^+_{W_1}}, \tag{10.216}$$

For the second, it would be erroneous to degrade the right product, and it is preferable to characterize the error as

$$\rho = \frac{k^+_{R_2}}{k^+_{R_2} + k^+_{W_2}}. \tag{10.217}$$

The error rate can, again, go beyond the classical kinetically controlled regime, by the appropriate combination of both processes. If $\chi \ll 1, \rho \ll 1$, $\zeta \to \chi\rho$, which is reminiscent of the χ^2-regime found for degradative proofreading.

Kinetic resolution through dissipative recycling

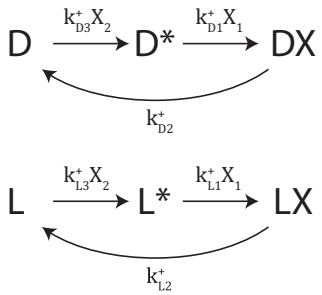


Figure 10.12: A schematic illustration of dissipative recycling for the resolution of enantiomers.

We can repeat this exercise for kinetic discrimination of two similar, uncoupled species (see Fig. 10.12). Let us again consider a pair of enantiomers L and D, of which we desire to maximize the production of DX. Let us write a system with a functionalization, degradation (e.g. hydrolysis) and reactivation reaction

$$\mathrm{D}^* + \mathrm{X}_1 \overset{D_1}{\rightleftharpoons} \mathrm{DX}, \qquad R_{D_1} = k^+_{D_1} N_{\mathrm{D}^*} N_{\mathrm{X}_1} - k^-_{D_1} N_{\mathrm{DX}} \tag{10.218}$$

$$\mathrm{DX} \overset{D_2}{\rightleftharpoons} \mathrm{D} + \mathrm{Y}_1, \qquad R_{D_2} = k^+_{D_2} N_{\mathrm{DX}} - k^-_{D_2} N_{\mathrm{D}} N_{\mathrm{Y}_1} \tag{10.219}$$

$$\mathrm{D} + \mathrm{X}_2 \overset{D_3}{\rightleftharpoons} \mathrm{D}^* + \mathrm{Y}_2, \qquad R_{D_3} = k^+_{D_3} N_{\mathrm{D}} N_{\mathrm{X}_2} - k^-_{D_3} N_{\mathrm{D}^*} N_{\mathrm{Y}_2}. \tag{10.220}$$

We define the analogous set of reactions for the L species

$$\mathrm{L}^* + \mathrm{X}_1 \overset{L_1}{\rightleftharpoons} \mathrm{LX}, \qquad R_{L1} = k^+_{L_1} N_{\mathrm{L}^*} N_{\mathrm{X}_1} - k^-_{L_1} N_{\mathrm{LX}} \tag{10.221}$$

$$\mathrm{LX} \overset{L_2}{\rightleftharpoons} \mathrm{L} + \mathrm{Y}_1, \qquad R_{L_2} = k^+_{L_2} N_{\mathrm{LX}} - k^-_{L_2} N_{\mathrm{L}} N_{\mathrm{Y}_1} \tag{10.222}$$

$$\mathrm{L} + \mathrm{X}_2 \overset{L_3}{\rightleftharpoons} \mathrm{L}^* + \mathrm{Y}_2, \qquad R_{L_3} = k^+_{L_3} N_{\mathrm{L}} N_{\mathrm{X}_2} - k^-_{L_3} N_{\mathrm{L}^*} N_{\mathrm{Y}_2}. \tag{10.223}$$

In the absence of racemization reactions, the system of equations to be solved for a steady state is solved easily in the irreversible reaction regime explored before:

$$R^+_{D_1} = R^+_{D_2} = R^+_{D_3}, \quad R^+_{L_1} = R^+_{L_2} = R^+_{L_3}. \tag{10.224}$$

Using the conservation laws

$$N_{L^\circ} = N_{\mathrm{L}} + N_{\mathrm{LX}} + N_{\mathrm{L}^*}, \qquad N_{D^\circ} = N_{\mathrm{D}} + N_{\mathrm{DX}} + N_{\mathrm{D}^*} \tag{10.225}$$

We obtain

$$N_{DX} = N_D^\circ \frac{\kappa_{D_3}\kappa_{D_1}}{(\kappa_{D_1}+\kappa_{D_2})\kappa_{D_3}+\kappa_{D_2}\kappa_{D_1}} = N_D^\circ \frac{1}{1+\kappa_{D_2}/\kappa_{D_1}+\kappa_{D_2}/\kappa_{D_3}}, \tag{10.226}$$

$$N_{LX} = N_L^\circ \frac{\kappa_{L_3}\kappa_{L_1}}{(\kappa_{L_1}+\kappa_{L_2})\kappa_{L_3}+\kappa_{L_2}\kappa_{L_1}} = N_L^\circ \frac{1}{1+\kappa_{L_2}/\kappa_{L_1}+\kappa_{L_2}/\kappa_{L_3}}. \tag{10.227}$$

where $\kappa_{L_1} = k^+_{L_1} N_{\mathrm{X}_1}$, $\kappa_{L_2} = k^+_{L_2}$ and $\kappa_{L_3} = k^+_{L_3} N_{\mathrm{X}_2}$. Similarly, $\kappa_{D_1} = k^+_{D_1} N_{\mathrm{X}_1}$, $\kappa_{D_2} = k^+_{D_2}$ and $\kappa_{D_3} = k^+_{D_3} N_{\mathrm{X}_2}$. Setting $N_L^\circ = N_D^\circ$, we obtain an error rate

$$\zeta = \frac{1+\kappa_{D_2}/\kappa_{D_1}+\kappa_{D_2}/\kappa_{D_3}}{2+\kappa_{D_2}/\kappa_{D_1}+\kappa_{D_2}/\kappa_{D_3}+\kappa_{L_2}/\kappa_{L_3}+\kappa_{L_2}/\kappa_{L_1}} \tag{10.228}$$

Although there are three reactions at which discrimination can take place, we no longer have the product of three ratios of rates. This can be understood untuitively, when we consider the lifetimes of molecules in their different forms:

$$\tau_{D_1} = \frac{1}{\kappa_{D_1}}, \qquad \tau_{D_2} = \frac{1}{\kappa_{D_2}}, \qquad \tau_{D_3} = \frac{1}{\kappa_{D_3}}, \tag{10.229}$$

and idem for L molecules. The species of interest, DX, spends a typical time τ_{D_2} in that state before being degraded to D. D then spends a typical time τ_{D_3} to be reactivated, and D^* spends a time τ_{D_1} to be functionalized to D. From this argument, we find that D and L molecules respectively spend a fraction

$$p(\mathrm{DX}) = \frac{\tau_{D_2}}{\tau_{D_1}+\tau_{D_2}+\tau_{D_3}}, \qquad p(\mathrm{LX}) = \frac{\tau_{L_2}}{\tau_{L_1}+\tau_{L_2}+\tau_{L_3}}, \tag{10.230}$$

of their time as a DX, respectively LX, molecule. Noting that

$$\zeta = \frac{p(\mathrm{LX})}{p(\mathrm{DX})+p(\mathrm{LX})} = \frac{\tau_{L_2}}{\tau_{D_2}+\tau_{L_2}} \frac{\tau_{D_1}+\tau_{D_2}+\tau_{D_3}}{\tau_{D_1}+\tau_{D_2}+\tau_{D_3}+\tau_{L_1}+\tau_{L_2}+\tau_{L_3}}, \tag{10.231}$$

we see that the time occupation argument implies that reactivation and degradation play a similar role, their contribution is additive. We obtain the following limiting regimes of interest

$$\zeta \to 1, \qquad \kappa_{L_2} \ll \kappa_{L_3}, \kappa_{L_1}. \quad \kappa_{D_2} \ll \kappa_{D_3}, \kappa_{D_1}, \tag{10.232}$$

$$\zeta \to \frac{\tau_{L_2}}{\tau_{D_2}+\tau_{L_2}} \frac{\tau_{D_1}}{\tau_{D_1}+\tau_{L_1}}, \qquad \kappa_{L_1} \ll \kappa_{L_3}, \kappa_{L_2}. \quad \kappa_{D_1} \ll \kappa_{D_3}, \kappa_{D_2}, \tag{10.233}$$

$$\zeta \to \frac{\tau_{L_2}}{\tau_{D_2}+\tau_{L_2}} \frac{\tau_{D_3}}{\tau_{D_3}+\tau_{L_3}}, \qquad \kappa_{L_3} \ll \kappa_{L_1}, \kappa_{L_2}. \quad \kappa_{D_3} \ll \kappa_{D_1}, \kappa_{D_2}. \tag{10.234}$$

Dissipative recycling can thus access regimes where the kinetic selectivity of two reactions is exploited, without suffering degradation. Nevertheless, this selectivity does not come for free: the reaction needs a fuel to do so.

Every sequence of reactions $\mathrm{L_1L_2L_3}$ or $\mathrm{D_1D_2D_3}$ is accompanied by the net transfer of $\mathrm{X_1X_2}$ to the system from a reservoir, conversion to $\mathrm{Y_1Y_2}$, and their subsequent ejection in a reservoir

$$\mathrm{X_1 + X_2 \to Y_1 + Y_2}. \tag{10.235}$$

The free energy change accompanying this process is

$$\Delta\mu = \mu^\circ_{Y_1} + \mu^\circ_{Y_2} - \mu^\circ_{X_1} - \mu^\circ_{X_2} - k_bT \ln \frac{N_{X_1}N_{X_2}}{N_{Y_1}N_{Y_2}} = \Delta\mu^\circ - k_bT \ln \frac{N_{X_1}N_{X_2}}{N_{Y_1}N_{Y_2}}. \tag{10.236}$$

In our present evaluation, we have assumed that all reactions are irreversible, which means $\Delta\mu \ll 0$. This is generally necessary to access to proofreading regimes. A detailed analysis of this tradeoff is

beyond the scope of this discussion, the present objective is only to demonstrate how extra reactions modify the limitations encountered in synthesis.

As a strategy to obtain DX, dissipative recycling has the drawback that reservoirs need to keep feeding the reaction. $\Delta\mu$ can be seen as a maintenance cost, that needs to be paid per D molecule on a timescale $\tau_D = \tau_{D1} + \tau_{D2} + \tau_{D3}$. If the timescale τ_r of the experiment or physical process using DX is much longer ($\tau_r \gg \tau_D$), then this strategy quickly becomes very expensive.

The biological process of kinetic proofreading solves this problem by abstracting functionalized molecules DX, LX from the mixture.

10.3.4 Kinetic Proofreading

Kinetic proofreading was pioneered by Hopfield [32] and subsequently expanded upon by Ninio [33] as a branch of mechanisms to account for the high fidelity of biochemical reactions. These fidelities could neither be explained in terms of thermodynamic control nor kinetic control. Moreover, networks can have multiple proofreading steps, to accumulate selectivity. It was recently found that tRNA [34] discrimination in bacteria utilizes at least two proofreading steps, and it is thought that proofreading steps in biological networks often come in pairs.

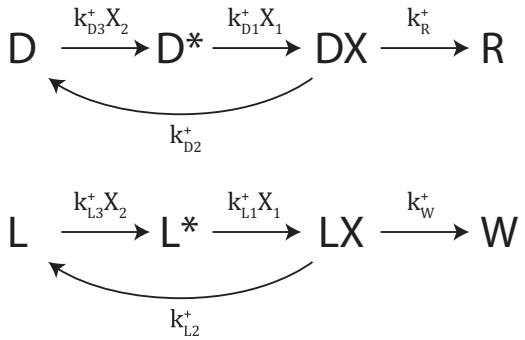


Figure 10.13: Simplified representation of resolution by proofreading. Reactions are drawn as irreversible.

The original Hopfield model can be otained by adding an irreversible propagation reaction to the dissipative recycling scheme or a recycling path to the degradative proofreading scheme. Let us here consider these decorations for a kinetic resolution process

$$\mathrm{D}^* + \mathrm{X}_1 \overset{D_1}{\rightleftharpoons} \mathrm{DX}, \qquad R_{D_1} = k^+_{D_1} N_{\mathrm{D}^*} N_{\mathrm{X}_1} - k^-_{D_1} N_{\mathrm{DX}}, \tag{10.237}$$

$$\mathrm{DX} \overset{R}{\rightleftharpoons} \mathrm{R}, \qquad R_{\mathrm{R}} = k^+_{\mathrm{R}} N_{\mathrm{DX}} - k^-_{\mathrm{R}} N_{\mathrm{R}}, \tag{10.238}$$

$$\mathrm{DX} \overset{D_2}{\rightleftharpoons} \mathrm{D} + \mathrm{Y}_1, \qquad R_{D_2} = k^+_{D_2} N_{\mathrm{DX}} - k^-_{D_2} N_{\mathrm{D}} N_{\mathrm{Y}_1}, \tag{10.239}$$

$$\mathrm{D} + \mathrm{X}_2 \overset{D_3}{\rightleftharpoons} \mathrm{D}^* + \mathrm{Y}_2, \qquad R_{D_3} = k^+_{D_3} N_{\mathrm{D}} N_{\mathrm{X}_2} - k^-_{D_3} N_{\mathrm{D}^*} N_{\mathrm{Y}_2}, \tag{10.240}$$

$$\mathrm{L}^* + \mathrm{X}_1 \overset{L_1}{\rightleftharpoons} \mathrm{LX}, \qquad R_{L_1} = k^+_{L_1} N_{\mathrm{L}^*} N_{\mathrm{X}_1} - k^-_{L_1} N_{\mathrm{LX}}, \tag{10.241}$$

$$\mathrm{DX} \overset{W}{\rightleftharpoons} \mathrm{W}, \qquad R_{\mathrm{W}} = k^+_{\mathrm{W}} N_{\mathrm{DX}} - k^-_{\mathrm{W}} N_{\mathrm{W}}, \tag{10.242}$$

$$\mathrm{LX} \overset{L_2}{\rightleftharpoons} \mathrm{L} + \mathrm{Y}_1, \qquad R_{L_2} = k^+_{L_2} N_{\mathrm{LX}} - k^-_{L_2} N_{\mathrm{L}} N_{\mathrm{Y}_1}, \tag{10.243}$$

$$\mathrm{L} + \mathrm{X}_2 \overset{L_3}{\rightleftharpoons} \mathrm{L}^* + \mathrm{Y}_2, \qquad R_{L_3} = k^+_{L_3} N_{\mathrm{L}} N_{\mathrm{X}_2} - k^-_{L_3} N_{\mathrm{L}^*} N_{\mathrm{Y}_2}. \tag{10.244}$$

For this proofreading scheme, we can write conservation laws for the two compounds we are resolving

$$N^\circ_{\mathrm{D}} = N_{\mathrm{D}} + N_{\mathrm{D}^*} + N_{\mathrm{R}} + N_{\mathrm{DX}}, \qquad N^\circ_{\mathrm{L}} = N_{\mathrm{L}} + N_{\mathrm{L}^*} + N_{\mathrm{W}} + N_{\mathrm{LX}}. \tag{10.245}$$

Having the final propagation reaction be far slower than the rest of the reactions $k_R^+ \ll k_{D_1}^+, k_{D_2}^+, k_{D_3}^+$, we can use the steady-state solution for dissipative recycling for DX, LX.

$$N_{\mathrm{DX}} = (N_{\mathrm{D}}^\circ - N_{\mathrm{R}}) \frac{\kappa_{D_1}\kappa_{D_3}}{\kappa_{D_2}\kappa_{D_3} + \kappa_{D_1}\kappa_{D_3} + \kappa_{D_1}\kappa_{D_2}}, \tag{10.246}$$

$$N_{\mathrm{LX}} = (N_{\mathrm{L}}^\circ - N_{\mathrm{W}}) \frac{\kappa_{L_1}\kappa_{L_3}}{\kappa_{L_2}\kappa_{L_3} + \kappa_{L_1}\kappa_{L_3} + \kappa_{L_1}\kappa_{L_2}}. \tag{10.247}$$

The production rates of the right product R and wrong product W is then

$$R_R^+ = \kappa_R N_{\mathrm{DX}}, \tag{10.248}$$

$$R_W^+ = \kappa_W N_{\mathrm{LX}}. \tag{10.249}$$

From which we find

$$N_{\mathrm{DX}} = N_{\mathrm{D}}^\circ \exp(-\kappa_R t) \frac{\kappa_{D_1}\kappa_{D_3}}{\kappa_{D_2}\kappa_{D_3} + \kappa_{D_1}\kappa_{D_3} + \kappa_{D_1}\kappa_{D_2}}, \tag{10.250}$$

$$N_{\mathrm{LX}} = N_{\mathrm{L}}^\circ \exp(-\kappa_W t) \frac{\kappa_{L_1}\kappa_{L_3}}{\kappa_{L_2}\kappa_{L_3} + \kappa_{L_1}\kappa_{L_3} + \kappa_{L_1}\kappa_{L_2}}. \tag{10.251}$$

On short timescales, using $N_{\mathrm{L}}^\circ = N_{\mathrm{D}}^\circ$ this gives an error fraction

$$\zeta = \frac{\kappa_W \frac{\kappa_{L_1}\kappa_{L_3}}{\kappa_{L_2}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_2}}}{\kappa_W \frac{\kappa_{L_1}\kappa_{L_3}}{\kappa_{L_2}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_2}} + \kappa_R \frac{\kappa_{D_1}\kappa_{D_3}}{\kappa_{D_2}\kappa_{D_3}+\kappa_{D_1}\kappa_{D_3}+\kappa_{D_1}\kappa_{D_2}}} = \frac{1}{1 + \frac{\kappa_{\mathrm{R}}}{\kappa_{\mathrm{W}}} \frac{\kappa_{D_1}\kappa_{D_3}}{\kappa_{L_1}\kappa_{L_3}} \frac{\kappa_{L_2}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_3}+\kappa_{L_1}\kappa_{L_2}}{\kappa_{D_2}\kappa_{D_3}+\kappa_{D_1}\kappa_{D_3}+\kappa_{D_1}\kappa_{D_2}}}, \tag{10.252}$$

which has a number of straightforward regimes

$$\zeta \to \frac{1}{1+\frac{\kappa_{\mathrm{R}}\kappa_{D1}\kappa_{L2}}{\kappa_{\mathrm{W}}\kappa_{L1}\kappa_{D2}}}, \qquad \frac{\kappa_{L2}\kappa_{L3}}{\kappa_{D2}\kappa_{D3}} \gg \frac{\kappa_{L2}\kappa_{L1}}{\kappa_{D2}\kappa_{D1}}, \frac{\kappa_{L1}\kappa_{L3}}{\kappa_{D3}\kappa_{D1}}, \tag{10.253}$$

$$\zeta \to \frac{1}{1+\frac{\kappa_{\mathrm{R}}\kappa_{L2}\kappa_{D3}}{\kappa_{\mathrm{W}}\kappa_{D2}\kappa_{L3}}}, \qquad \frac{\kappa_{L2}\kappa_{L1}}{\kappa_{D2}\kappa_{D1}} \gg \frac{\kappa_{L1}\kappa_{L3}}{\kappa_{D3}\kappa_{D1}}, \frac{\kappa_{L2}\kappa_{L3}}{\kappa_{D2}\kappa_{D3}}, \tag{10.254}$$

$$\zeta \to \frac{1}{1+\frac{\kappa_{\mathrm{R}}}{\kappa_{\mathrm{W}}}}, \qquad \frac{\kappa_{L1}\kappa_{L3}}{\kappa_{D3}\kappa_{D1}} \gg \frac{\kappa_{L2}\kappa_{L1}}{\kappa_{D2}\kappa_{D1}}, \frac{\kappa_{L2}\kappa_{L3}}{\kappa_{D2}\kappa_{D3}}. \tag{10.255}$$

Two of the regimes identified here involve the ratios of three different rate constants, a considerable improvement over the single ratio in simple kinetic resolution.

On longer timescales, we need to take into account the depletion of reactants and we can write

$$\zeta = \frac{1}{1 + \frac{\kappa_{D1}\kappa_{D3}}{\kappa_{L1}\kappa_{L3}} \frac{\kappa_{L2}\kappa_{L3}+\kappa_{L1}\kappa_{L3}+\kappa_{L1}\kappa_{L2}}{\kappa_{D2}\kappa_{D3}+\kappa_{D1}\kappa_{D3}+\kappa_{D1}\kappa_{D2}} (1 - \exp((\kappa_{\mathrm{R}} - \kappa_{\mathrm{W}})t))} \tag{10.256}$$

As before, this error fraction will quickly rise when DX is being exhausted. If, however, a rapid racemization reaction is introduced (see Fig. 10.14) that interconverts D and L,

$$\mathrm{D} \overset{rac}{\rightleftharpoons} \mathrm{L} \tag{10.257}$$

the error fraction will remain stable at the fixed value given by Eq. (10.252), provided $k_{rac} \gg \kappa_{L_1}, \kappa_{L_2}, \kappa_{L_3}, \kappa_{D_1}, \kappa_{D_2}, \kappa_{D_3}, \kappa_W, \kappa_R$.

Since the only final products are R and W and everything is recycled, we can define a yield for the current

$$\eta_R = \frac{R_R^+}{R_W^+ + R_R^+} = 1 - \zeta. \tag{10.258}$$

Due to the recycling, the reduced error ζ also improves the yield of desired product. To reach this optimal regime, k_R^+ needs to be relatively slow, which means the system performs several dissipative cycles, which work best in a strongly irreversible regime, which is discussed in detail in Ref. [24]. The exact behavior is shaped by delicate tradeoffs between speed, dissipation and accuracy.

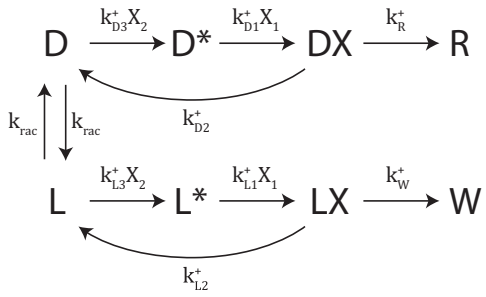


Figure 10.14: Simplified representation of dynamic kinetic resolution with proofreading. Apart from racemization, reactions are drawn as irreversible.

Multiple proofreading steps

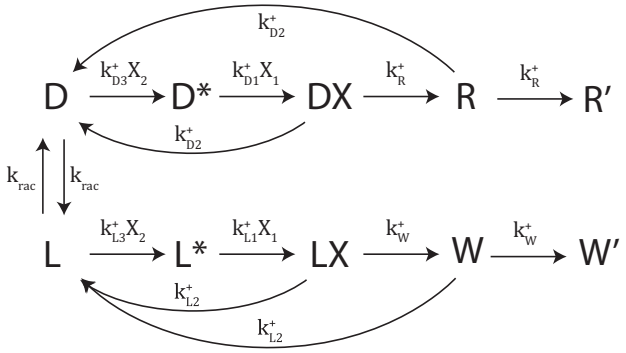


Figure 10.15: A network performing two proofreading steps. In order to correct the error, an incorrect species has to be brought all the way to D or L, here performed in a single step.

In Fig. 10.15 a two-step proofreading scheme is depicted. In the second step, a fraction χ_R (resp. χ_W) of the current J_R (resp. J_W) is converted to a current of R', $J_{R'}$ (resp. W', $J_{W'}$)

$$J_{R'} = \frac{k_R^+}{k_R^+ + k_{D2}^+} J_R = \chi_R J_R, \tag{10.259}$$

$$J_{W'} = \frac{k_W^+}{k_W^+ + k_{L2}^+} J_W = \chi_W J_W. \tag{10.260}$$

From this, we can write a steady-state error rate $\zeta^{(2)}$ for two proofreading steps, using $\zeta^{(1)}$ from the scheme with one proofreading step

$$\zeta^{(2)} = \frac{J_{W'}}{J_{R'} + J_{W'}} = \frac{\chi_W \zeta^{(1)} J}{\chi_R (1-\zeta^{(1)}) J + \chi_W \zeta^{(1)} J}. \tag{10.261}$$

where J denotes a total flux

$$J = J_R + J_W. \tag{10.262}$$

Let us now define ρ as the ratio of χ_W, χ_R

$$\rho = \frac{\chi_W}{\chi_R}, \tag{10.263}$$

So that we can rewrite $\zeta^{(2)}$ as

$$\zeta^{(2)} = \frac{\rho \zeta^{(1)}}{1 - \zeta^{(1)} + \rho \zeta^{(1)}} = \zeta^{(1)} \zeta^{(1b)} \tag{10.264}$$

Where $\zeta^{(1b)}$ is an error rate for the second step.

By this extra proofreading step, we have reduced the error rate with a factor $\zeta^{(1b)}$. For $\zeta^{(1)} \ll 1$, $\rho \ll 1$, we have $\zeta^{(1b)} \to \rho$. Had we added n proofreading steps of this kind, we would improve the resolution by a factor ρ every time, such that

$$\zeta^{(n+1)} \to \zeta^{(1)} \rho^n. \tag{10.265}$$

These extra steps come at a cost, however.

With extra steps comes an increase in processing time. The extra proofreading step sends a fraction of $1-\chi_R$ correct molecules back to the starting position, which means it takes on average $1/\chi_R$ full attempts to reach R′. Letting τ_R be the typical time to finish the first proofreading step, we then have a typical waiting time

$$\tau^{(2)}_{KPR} = \frac{\tau_R}{\chi_R} + \tau'_R. \tag{10.266}$$

For n proofreading steps sending back $1-\chi_R$ correct molecules, we have, roughly

$$\frac{\tau^{(n+1)}_{KPR}}{\tau^{(n)}_{KPR}} \approx \frac{1}{\chi_R} \tag{10.267}$$

As the reactions are irreversible, we need to expend free energy to fuel every proofreading step, which leads to escalating costs.

A detailed characterization of the tradeoff between accuracy, dissipation and processing speed for various proofreading schemes was discussed by Rao and Peliti [24], to which we refer for a more in-depth discussion. For our present purposes, it is important to know that this trade-off exists and that it follows from a network structure that sends incorrect species back.

Towards repertoires of functional networks

The ways in which we do chemistry are changing. More elaborate chemical networks exploit properties of molecular collectives, allowing to go beyond classical synthetic limits. By making reactions dynamic (e.g. dynamic covalent and noncovalent chemistry), large libraries of interconverting compounds can be generated, which can be screened and selected [35]. Such ensembles can self-sort and self-assemble into functional structures such as molecular cages, self-healing polymers and responsive materials [36]. By making kinetic resolution dynamic (DKR, e.g. through a racemization catalyst[28]) or by abstracting a competing substrate by complexation (e.g. via a molecular cage[37]), important limitations in yield, purity and efficiency are overcome. By introducing degradation and regeneration steps, chemical products can be enriched and purified in situ, used recently to abiotically enrich 5'-3' over 5'-2' bonds in RNA[31]. Since its inception by Kondepudi et al, autocatalytic pathways became a topic of interest for origins-of-life questions regarding chirality. Autocatalytic recycling in crystallization and redissolution (Viedma Ripening [38]) is now employed on industrial scales to form enantiopure compounds[39], rapidly overturning classical purifications of chiral drugs.

In biology, such unorthodox ways of doing chemistry are commonplace. Biological systems exploit dynamic reactions to exchange and store sugars into compartmentalized polymers, acting as a self-organized reservoir of sugar monomers [40]. In kinetic proofreading, kinetically controlled intermediates are degraded[32, 33]. A second, slower reaction acting on these intermediates can then exert another round of kinetic control. In nature, multiple error-correction steps are coupled with selective enzymes, to achieve the high fidelities required for e.g. protein synthesis and genetic replication[34, 41]. Sequential irreversible reactions introduce memory in waiting times, which is exploited in chemical clocks, checkpoint control, and sensing[42, 43, 44].

The present section has attempted to give a little taste for how small reaction networks lead to interesting behavior. The building blocks for such networks are often counterintuitive: racemization achieves the exact opposite of resolution, yet it can be exploited in DKR to improve the resolution. Dissipatively converting product compounds back to substrates seems wasteful, but in proofreading schemes it achieves error correction (Before proofreading was understood, such reactions were indeed interpreted as wasteful side-reactions[32]).

For the Origins of Life and prebiotic chemistry, it is of the essence to get a firm understanding of these subtleties. It may have taken a while for extant biomolecules and macromolecules to enter into abiogenesis. Functional reaction networks, however, are far more general. They may have been the bread and butter of abiogenesis, by providing some of the earliest plausible features that chemical evolution could have selected for.

10.4 Dissipative sequence exploration: Cycle decomposition

To enumerate the number of independent cycles in a reaction network, one can use the corresponding stoichiometry matrix [7, 45] to show that

$$r+\ell=c+s, \tag{10.268}$$

where:
r: the number of possible reactions
ℓ: the number of conservation laws
c: the number of independent cycles
s: the number of non-chemostatted species
We reiterate that a cycle is a linear combination of reactions that leaves the system unchanged. We now introduce L, which denotes the largest number of monomers a polymer can contain. Let L be arbitrarily large. Then, for even L, there are: L activation reactions, as there exist L types of unactivated polymer.
$3/4L^2-1/2L$ ligation reactions, which is the number of unique pairs $(\mathbf{n}^*,\mathbf{m})$ and $(\mathbf{n}^*,\mathbf{m}^*)$ we can make (noting that $(\mathbf{n}^*,\mathbf{m}^*)=(\mathbf{m}^*,\mathbf{n}^*)$), for which $n+m\leq L$.
$3/4L^2-1/2L$ hydrolysis reactions, which is the number of unique pairs $(\mathbf{n},\mathbf{m})$ and $(\mathbf{n},\mathbf{m}^*)$ we can make, for which $n+m\leq L$.
For uneven L, we have $3/4L^2-1/2L-1/4$ ligation reactions and $3/4L-1/2L-1/4$ hydrolysis reactions. Thus: $r=3/2L^2$ for even L (resp. $r=3/2L^2-1/2$ for uneven L). We only have conservation of monomer units: $l=1$. We have activated and unactived polymers up to length L: $s=2L$. This yields $c=3/2L^2-2L+1$ (resp $c=3/2L^2-2L+1/2$ for uneven). The smallest cycles can be subdivided in three types:
I: $\mathbf{n}^*+\mathbf{m}\rightarrow[\mathbf{n}+\mathbf{m}]\rightarrow\mathbf{n}+\mathbf{m}\rightarrow\mathbf{n}^*+\mathbf{m}$
II: $\mathbf{n}^*+\mathbf{m}^*\rightarrow[\mathbf{n}+\mathbf{m}]^*\rightarrow\mathbf{n}+\mathbf{m}^*\rightarrow\mathbf{n}^*+\mathbf{m}^*$
III: $\mathbf{n}^*+\mathbf{m}\rightarrow[\mathbf{n}+\mathbf{m}]\rightarrow[\mathbf{n}+\mathbf{m}]^*\rightarrow\mathbf{n}^*+\mathbf{m}$
where each cycle can be identified with (j,n,m), where $j\in\{I,II,III\}$.We have compacted the reaction notation by only including polymer species. For example, cycle (I,n,m) reads as: ligation of $\mathbf{n}^*+\mathbf{m}$, followed by hydrolysis yielding $\mathbf{n}+\mathbf{m}$, followed by activation of $\mathbf{n}$. Note that these cycles are not all linearly independent, as we have

$$(I,n,m)=(I,m,n)-(II,m,n)-(III,m,n)+(II,n,m)+(III,n,m).\ \ \forall n\neq m \tag{10.269}$$

Up to $L=3$, these cycles describe the system completely. Starting at $L=4$, we also see the appearance of larger cycles of the form:
IV: $\mathbf{n}^*+\mathbf{m}^*+\mathbf{k}^*\rightarrow[\mathbf{n}+\mathbf{m}]^*+\mathbf{k}^*\rightarrow[\mathbf{n}+\mathbf{m}+\mathbf{k}]^*\rightarrow[\mathbf{n}+\mathbf{k}]^*+\mathbf{m}\rightarrow\mathbf{n}^*+\mathbf{m}+\mathbf{k}\rightarrow\mathbf{n}^*+\mathbf{m}^*+\mathbf{k}\rightarrow\mathbf{n}^*+\mathbf{m}^*+\mathbf{k}^*$
For our purposes, the exact choice of linearly independent cycles is not important.

Since activation increases the number of activated species by one and ligation decreases the number of activated species by one, we need one ligation reaction per activation reaction to construct a cycle. Similarly, ligation decreases the total number of species by one, and hydrolysis increases the total number of species by one. It follows that any cycle we construct must contain $3n$ reactions, with n of each type.

10.4.1 Stability of steady state solution

In order to assess the stability of the fixed point, we follow an approach similar to Ref. [24]. We first define the Lyapunov function

$$L = \sum_n N_n^A \ln \frac{N_n^A}{\bar{N}_n^A} + \sum_n N_n^D \ln \frac{N_n^D}{\bar{N}_n^D} + (N - \bar{N}). \tag{10.270}$$

where $\bar{N}_n^A$ denotes the steady state concentration of $\mathbf{n}^*$. This Lyapunov function has the time derivative

$$\frac{dL}{dt} = \sum_n \frac{dN_n^A}{dt} \ln \frac{N_n^A}{\bar{N}_n^A} + \sum_n \frac{dN_n^D}{dt} \ln \frac{N_n^D}{\bar{N}_n^D}, \tag{10.271}$$

Let us now consider a small perturbation[††] ε, either applied at $\mathbf{n}^*$

$$N_m^A = \bar{N}_m^A + \varepsilon \delta_m^n, \tag{10.272}$$

or at $\mathbf{n}$

$$N_m^D = \bar{N}_m^D + \varepsilon \delta_m^n. \tag{10.273}$$

Upon substitution, we obtain for $\mathbf{n}^*$

$$\begin{aligned} \frac{dL}{dt} &= -\frac{\varepsilon^2}{\bar{N}_n^A}[k_{act}^- N_{XH} \bar{N}_n^A + k_{lig}^- (n-1) N_{YOH} \bar{N}_n^A \\ &+ k_{hyd}^- N^D N_n^A + k_{lig}^+ (\bar{N}_n^A + \bar{N}_n) \\ &+ k_{hyd}^+ (n-1) N_{H_2O}], \end{aligned} \tag{10.274}$$

and for $\mathbf{n}$ we obtain

$$\begin{aligned} \frac{dL}{dt} &= -\frac{\varepsilon^2}{\bar{N}_n^A}[k_{act}^+ N_{XY} \bar{N}_n^A + k_{lig}^- (n-1) N_{YOH} \bar{N}_n^D \\ &+ k_{hyd}^- (N + N^D) N_n^D + k_{lig}^+ \bar{N}^A \\ &+ k_{hyd}^+ (n-1) N_{H_2O}]. \end{aligned} \tag{10.275}$$

Either way, $L < 0$, irrespective of the sign of ε, which ensures the stability of the fixed point.

10.4.2 Optimal Exploration fraction

Starting from an exponential length distribution, we can write

$$N_n = Q^{n-1} N_1, \tag{10.276}$$

[††]For this argument to be more rigorous, we are implicitly assuming we introduced a chemostat such that N, M are allowed to fluctuate. Alternatively, one can introduce multiple perturbations that collectively fix N, M, to arrive at the same conclusion.

where $Q = 1 - N/M$ parametrizes the steepness of the exponential. We can immediately write

$$N = \frac{N_1}{1-Q}, \tag{10.277}$$

$$M = \frac{N_1}{(1-Q)^2}. \tag{10.278}$$

To quantify the number of cycles in which we explore a species of length n, we look at steady state hydrolysis. Due to Kirchoffs law, this degradation equals the formation and thus exploration. As we assume all bonds are equally prone to hydrolysis, we are interested in the fraction of the total bonds that are present in species of length n

$$\varepsilon_n = \frac{k_{hyd}(n-1)N_n}{k_{hyd}(M-N)} = \frac{(n-1)Q^{n-1}}{\frac{1}{(1-Q)^2} - \frac{1}{1-Q}}. \tag{10.279}$$

In order to maximize exploration, we need to find the Q that maximizes ε_n

$$\frac{d\varepsilon_n}{dQ} = 0. \tag{10.280}$$

Which can be shown to yield

$$Q = \frac{n-2}{n}. \tag{10.281}$$

Upon substitution in Eq. (10.279) this yields

$$\varepsilon_n = \frac{4(n-2)^{n-2}(n-1)}{n^n}. \tag{10.282}$$

We note that

$$\lim_{n\to\infty} \left(\frac{n-2}{n}\right)^{n-2} = e^2, \tag{10.283}$$

so that for $n \gg 1$ we can write

$$\varepsilon_n \approx \frac{4}{e^2 n}. \tag{10.284}$$

Although the length distribution decreases exponentially, we still obtain appreciable exploration, as long as the distribution is sufficiently tailored to the desired length. For a given length n, this ideally corresponds to a degree of concatenation of half that length

$$\chi = \frac{1}{1-Q} = \frac{n}{2}. \tag{10.285}$$

As an example, suppose we would wish to explore sequences of length 100. For an exponential distribution, the ideal average length is then 50, and the fraction of cycles exploring this length is $\varepsilon_{100} \approx 0.0054$. If $\chi = 100$ instead, we have $\varepsilon_{100} \approx 0.0037$.

10.5 Appendix: Population-level noise from a single individual

Let us consider an age-dependent renewal process, in which the probability density of branching at age t is given by $f(t)$, and upon branching, the probability of having k offspring is given by

ϕ_k (assumed to be age-independent for simplicity). We would like to evaluate the behavior of the number $N(t)$ of individuals at time t. Let us define the function $h(s)$ by

$$h(s) = \sum_{k=0}^{\infty} \phi_k s^k. \tag{10.286}$$

We also define the generating function for the process $N(t)$ by

$$G(s,t) = \sum_{k=0}^{\infty} p_k(t) s^k, \tag{10.287}$$

where $p_k(t)$ is the probability that $N(t) = k$. We assume that $p_k(0) = \delta_{k1} = p_k^0$, i.e., that we start from a single object. We can then evaluate $p_k(t)$ by adding the probability that no branching has occurred between 0 and t, which is given by $1 - \int_0^t dt\ f(t) = 1 - F(t)$, with the effect of the first branching at time u, such that $0 < u < t$. We obtain

$$p_k(t) = p_k^0 (1 - F(t)) + \sum_l \phi_l \int_0^t du\ f(u) \sum_{\{n_i\}} \delta_{\sum_{i=1}^{\ell} n_i, k} \prod_{i=1}^{\ell} p_{n_i}(t-u). \tag{10.288}$$

Multiplying by s^k and summing, we obtain

$$G(s,t) = s\,(1 - F(t)) + \int_0^t du\ f(u)\ h\,(G(s,t-u))\,. \tag{10.289}$$

Taking the derivative with respect to s at $s = 1$, we obtain the following equation for the average $\mu(t) = \sum_k k\, p_k(t)$:

$$\mu(t) = 1 - F(t) + m \int_0^t du\ \mu(t-u) f(u), \tag{10.290}$$

where $m = h'(1) = \sum_k k \phi_k$ is the average number of daughters upon branching.

To solve this equation in the limit $t \to \infty$, let us multiply both sides by $e^{-\alpha t}$ and take the limit. Since $\lim_{t\to\infty} F(t) = 1$, we obtain

$$\begin{aligned} \mu^* &= \lim_{t\to\infty} \mu(t) e^{-\alpha t} = \lim_{t\to\infty} \int_0^{\infty} du\ m\mu(t-u)\ e^{-\alpha(t-u)} e^{-\alpha u}\ f(u) \\ &= \mu^* m \int_0^{\infty} du\ e^{-\alpha u} f(u). \end{aligned} \tag{10.291}$$

This equation allows for a solution different from 0 and ∞ if α is chosen to satisfy

$$m \int_0^{\infty} du\ e^{-\alpha u} f(u) = 1. \tag{10.292}$$

Then, making use of a result by Smith [46], we obtain

$$\frac{1}{\mu^*} = \frac{\alpha m^2}{m-1} \int_0^{\infty} du\ u e^{-\alpha u} f(u). \tag{10.293}$$

As a consequence, we have

$$\mu(t) \approx \mu^* e^{\alpha t}. \tag{10.294}$$

In the case we are considering we have

$$f_L(t) = \frac{1}{\Gamma(L)} \kappa^L t^{L-1} e^{-\kappa t}, \tag{10.295}$$

and $m = 2$, which yields

$$\alpha = \kappa(2^{1/L} - 1) \approx \frac{\kappa \ln 2}{L}, \tag{10.296}$$

giving, as long as $L \gg 1$,

$$\mu(t) \approx \frac{2^{\kappa t/L}}{2\ln 2}. \tag{10.297}$$

We can use this framework to also evaluate higher moments of the population size, and from that obtain the coefficient of variation of the population size which characterizes the amplitude of the noise. Let us denote the second derivative of the generating function with respect to s by ζ

$$\zeta(t) = \left.\frac{d^2 G(s,t)}{d^2 s}\right|_{s=1} = \sum_{k=1}^{\infty} (k(k-1)) p_k(t). \tag{10.298}$$

At large times, $\zeta(t) \approx \zeta^* e^{2\alpha t}$. The variance of the population size σ^2 follows from the standard relation:

$$\sigma^2 = \zeta + \mu - \mu^2 \simeq \zeta - \mu^2. \tag{10.299}$$

For the specific case we are considering, we find

$$\zeta^* = \frac{2\mu^{*2}}{(2^{\frac{L+1}{L}} - 1)^L - 2}. \tag{10.300}$$

After extracting the leading contribution in the large L limit, we find:

$$\frac{\sigma}{\mu} \approx \frac{\sqrt{2}\ln(2)}{\sqrt{L}}, \tag{10.301}$$

which is numerically close to $1/\sqrt{L}$ since $\sqrt{2}\ln 2 = 0.980.. \approx 1$.

10.5.1 Population-level noise from n individuals

If we start from n individuals rather than just one, we can write the probability to have k individuals at time t, $p_k^{(n)}(t)$, in terms of the subpopulations generated by n single individuals,

$$p_k^{(n)}(t) = \sum_{\{m_1,\ldots,m_n\}} \delta_{\sum_j m_j,k} \prod_{j=1}^{n} p_{m_j}^{(1)}(t). \tag{10.302}$$

Here, $p_k^{(1)}(t)$ denotes the probability of having a population size of k at time t, starting from one individual, which was considered in 10.5. Note that we have added an additional superscript (1) to the notation used in 10.5 to emphasize the initial condition. From this equation, the new generating function follows :

$$G^{(n)}(s,t) = [G^{(1)}(s,t)]^n. \tag{10.303}$$

From this equation, we obtain the average,

$$\mu^{(n)}(t) = n\mu^{(1)}(t), \tag{10.304}$$

which expresses the average with n initial strands in terms of the average with one initial strand. For the second moment, we obtain

$$\zeta^{(n)} = n(n-1)[\mu^{(1)}]^2 + n\zeta^{(1)}, \tag{10.305}$$

We can then extract $\sigma^{(n)}$ by using Eq. (10.299), which yields

$$\begin{aligned}[\sigma^{(n)}]^2 &= n(n-1)[\mu^{(1)}]^2 + n\mu^{(1)} - n^2[\mu^{(1)}]^2 + n\zeta^{(1)} \\ &= n\zeta^{(1)} + n\mu^{(1)}(1-\mu^{(1)}).\end{aligned} \tag{10.306}$$

Together with Eq. (10.304), this leads to

$$\frac{\sigma^{(n)}}{\mu^{(n)}} \simeq \frac{\sqrt{\zeta^{(1)} - [\mu^{(1)}]^2}}{\sqrt{n}\mu^{(1)}} = \frac{\sigma^{(1)}}{\sqrt{n}\mu^{(1)}} \tag{10.307}$$

which is the coefficient of variation found previously for a single individual in the initial condition, divided by $\sqrt{n}$ as expected for the growth from independent individuals. This confirms the scaling found in Eq. (8.77).

Bibliography

Articles

[1] Christopher Jarzynski. “Nonequilibrium equality for free energy differences”. In: *Phys. Rev. Lett.* 78 (1997), p. 2690.

[5] Punyabrata Pradhan, Christian P Amann, and Udo Seifert. “Nonequilibrium Steady States in Contact : Approximate Thermodynamic Structure and Zeroth Law for Driven Lattice Gases”. In: *Phys. Rev. Lett.* 105 (2010), p. 150601.

[6] Manabendra N Bera et al. “Generalized laws of thermodynamics in the presence of correlations”. In: *Nat. Commun.* 8 (2017), p. 2180.

[7] Matteo Polettini and Massimiliano Esposito. “Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws”. In: *J. Chem. Phys.* 141.2 (2014), pp. –.

[8] Marian Smulochowski. “Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene”. In: *Phys. Z.* 2.1 (1927), pp. 226–251.

[9] J. Schnakenberg. “Network theory of microscopic and macroscopic behavior of master equation systems”. In: *Rev. Mod. Phys.* 48 (1976), p. 571.

[10] David Andrieux and Pierre Gaspard. “Fluctuation theorem for transport in mesoscopic systems”. In: *J. Stat. Mech.* 2006 (2006), P01011.

[11] David Andrieux and Pierre Gaspard. “Stochastic approach and fluctuation theorem for ion transport”. In: *J. Stat. Mech.* 2009 (2009), P02057.

[12] Jiayin Gu and Pierre Gaspard. “Stochastic approach and fluctuation theorem for charge transport in diodes”. In: *Phys. Rev. E.* 97 (2018), p. 052138.

[13] A. Butlerow. “Bildung einer zuckerartigen Substanz durch Synthese”. In: *Justus Liebigs Ann. Chem.* 120 (1861), pp. 295–298.

[14] Chandrakumar Appayee and Ronald Breslow. “Deuterium Studies Reveal a New Mechanism for the Formose Reaction Involving Hydride Shifts”. In: *J. Am. Chem. Soc.* 136 (2014), pp. 3720–3723.

[15] Liang Cheng, Charles Doubleday, and Ronald Breslow. “Evidence for tunneling in base-catalyzed isomerization of glyceraldehyde to dihydroxyacetone by hydride shift under formose conditions”. In: *Proc. Natl. Acad. Sci. U. S. A.* 112.14 (2015), pp. 4218–4220.

[16] R.F Socha, A.H Weiss, and M.M Sakharov. “Autocatalysis in the Formose Reaction”. In: *React. Kinet. Catal. Lett.* 14.2 (1980), pp. 119–128.

[18] Jason P Dworkin and Stanley L Miller. “A kinetic estimate of the free aldehyde content of aldoses”. In: *Carbohydr. Res.* 329 (2000), pp. 359–365.

[19] Michal Rivlin, Uzi Eliav, and Gil Navon. “NMR studies of the equilibria and reaction rates in aqueous solutions of formaldehyde”. In: *J. Phys. Chem. B* 119.12 (2015), pp. 4479–4487.

[20] Steven A Benner. “Prebiotic plausibility and networks of paradox-resolving independent models”. In: *Nat. Commun.* 9 (2018), p. 5173.

[21] Kepa Ruiz-mirazo, Carlos Briones, and De Escosura. “Prebiotic Systems Chemistry : New Perspectives for the Origins of Life”. In: *Chem. Rev.* 114 (2014), pp. 285–366.

[22] Florian M. Gartner et al. “Stochastic Yield Catastrophes and Robustness in Self-Assembly”. In: (2019). arXiv: 1905.09912.

[23] Basile Nguyen et al. “Thermodynamic bounds on the ultra- and infra-affinity of Hsp70 for its substrates Thermodynamic bounds on the ultra- and infra-affinity of Hsp70 for its”. In: *Biophys. J.* 113.June (2017), pp. 362–370. arXiv: arXiv:1702.01649v2.

[24] R. Rao and L. Peliti. “Thermodynamics of accuracy in kinetic proofreading: dissipation and efficiency trade-offs”. In: *J. Stat. Mech.* 2015.6 (2015), p. 06001.

[26] E Vedejs. “Method for direct hydroxylation of enolates. Transition metal peroxide reactions”. In: *J. Am. Chem. Soc.* 96.18 (1974), pp. 5944–5946.

[27] Hans J. Reich, James M. Renga, and Ieva L. Reich. “Organoselenium chemistry. Conversion of ketones to enones by selenoxide syn elimination”. In: *J. Am. Chem. Soc.* 97.19 (1975), pp. 5434–5447.

[28] H. Pellissier. “Dynamic Kinetic Resolution”. In: *Tetrahedron* 59.42 (2003), p. 8291.

[29] Cato Sandford, Daniel Seeto, and Alexander Y Grosberg. “Active sorting of particles as an illustration of the Gibbs mixing paradox”. In: (2018). arXiv: 1705.05537v2.

[30] Sophie Marbach, Lydéric Bocquet, and Sophie Marbach. “Active sieving across driven nanopores for tunable selectivity Active sieving across driven nanopores for tunable selectivity”. In: *J. Chem. Phys.* 146 (2017), p. 154701.

[31] Angelica Mariani and John D. Sutherland. “Non-Enzymatic RNA Backbone Proofreading through Energy-Dissipative Recycling”. In: *Angew. Chemie - Int. Ed.* 56.23 (2017), pp. 6563–6566.

[32] J. J. Hopfield. “Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes requiring high Specificity”. In: *Proc. Natl. Acad. Sci. U.S.A.* 71.10 (1974), pp. 4135–4139.

[33] J Ninio. “Kinetic amplification of enzyme discrimination”. In: *Biochimie* 57.5 (1975), pp. 587–595.

[34] K.W. Ieong et al. “Two proofreading steps in genetic code translation”. In: *Proc. Natl. Acad. Sci. U.S.A.* 113.48 (2016), pp. 13744–13749.

[35] Yinghua Jin et al. “Recent advances in dynamic covalent chemistry”. In: *Chem Soc Rev* 42 (2013), pp. 6634–6654.

[37] S Liu et al. “Kinetic Resolution of constitutional isomers controlled by selective protection inside a supramolecular nanocapsule”. In: *Nat. Chem.* 2 (2010), pp. 847–852.

[38] C. Viedma. “Chiral Symmetry Breaking During Crystallization: Complete Chiral Purity Induced by Nonlinear Autocatalysis and Recycling”. In: *Phys. Rev. Lett.* 94.6 (2005), p. 065504.

[39] I. Baglai et al. “A Viedma Ripening route to an enantiopure building block for Levetiracetam and Brivaracetam”. In: *Org. Biomol. Chem.* 17 (2019), pp. 35–38.

[40] Ö Kartal et al. “Carbohydrate-active enzymes exemplify entropic principles in metabolism”. In: *Mol. Syst. Biol.* 1 (7), p. 542.

[41] Pierre Gaspard. “Kinetics and thermodynamics of DNA polymerases with exonuclease proofreading”. In: *Phys. Rev. E* 93.4 (Apr. 2016), p. 42420.

[42] H Wierenga, P.R ten Wolde, and N. B. Becker. “Quantifying fluctuations in reversible enzymatic cycles and clocks”. In: *Phys. Rev. E* 97.4 (2018), p. 042404.

[43] T.W. McKeithan. “Kinetic proofreading in T-cell receptor signal transduction”. In: *Proc. Natl. Acad. Sci. U.S.A.c* 92.11 (1995), pp. 5042–5046.

[44] S.G. Das, M. Rao, and G Lyengar. “Universal lower bound on the free-energy cost of molecular measurement”. In: *Phys. Rev. Es* 95.6 (2017), p. 062410.

[45] Riccardo Rao and Massimiliano Esposito. “Nonequilibrium Thermodynamics of Chemical Reaction Networks : Wisdom from Stochastic Thermodynamics”. In: *Phys. Rev. X* 6 (2016), p. 041064.

[46] W L Smith. “Asymptotic Renewal Theorems”. In: *Proc. Roy. Soc. Edinburgh.* 64.1 (1953), pp. 9–48.

Books

[2] Théophile De Donder. *L’Affinité*. Oarus: Gauthier-Villars, 1927, p. 43.

[3] A.B. Pippard. *Elements of classical thermodynamics for advanced students of physics*. Cambridge University Press, Cambridge, UK, 1966.

[4] M. Bailyn. *A Survey of Thermodynamics*. New York: American Institute of Physics Press, 1994.

[17] Francis A. Carey and Richard J. Sundberg. *Advanced Organic Chemistry Part A: Structure and Mechanisms*. New York: Springer, 2006.

[25] T.W. Graham Solomons. *Organic Chemistry*. Sixth edit. New York: John Wiley & Sons, 1995.

[36] W Zhang and Y Jin, eds. *Dynamic Covalent Chemistry: Principles, Reactions, and Applications*. 2018.

RÉSUMÉ

Le domaine des Origines de la vie cherche à expliquer l'abiogenèse : comment la matière abiotique peut se transformer en vivant. Ces dernières décennies ont vu l'élaboration de plusieurs scénarios prébiotiques : hypothèses sur la place, la chimie et les mécanismes physiques de l'abiogenèse. Dans cette thèse, nous introduisons des cadres rigoureux, pour l'étude systématique des aspects physiques de l'abiogenèse. Ces cadres s'appuient sur des avancements en thermodynamique hors d'équilibre, réseaux de réactions chimiques et sélection sur plusieurs niveaux. Ils soulignent la cohérence thermodynamique et la structure fondamentale de la chimie. Ch.1 donne une introduction critique au domaine des origines de la vie, soulignant ce que nous savons, ce que les scénarios supposent et les développements historiques qui ont façonné la pensée actuelle. Ch.2 présente le cadre théorique des réseaux chimiques. Nous élargissons le cadre avec de nouveaux outils et un critère pour une description universelle de la chimie : la non-ambiguïté. Ch.3 décrit des moyens pour rendre les réseaux "ouverts" : chimiostats, CSTR, transfert en série, couplage diffusif aux compartiments. Le concept de chimiostat est étendu au "chimiostat composite", qui fixe une combinaison d'espèces. Nous corrigeons la loi zéro de la thermodynamique, qui, en thermodynamique stochastique, peut être brisée pour les quantités entières conservées. Ch.4 illustre le concept d'information dans les réseaux chimiques à l'aide d'un moteur qui extrait du travail en racémisant des énantiomères purs. On passe ensuite au procédé inverse : la purification des compositions. On présente des réseaux qui améliorent la purification, avec différents compromis. Ch.5 présente la dérivation des critères universels pour la catalyse et l'autoréplication pour les réseaux chimiques non-ambigus. L'échange entre phases et compartiments (diffusion, évaporation, etc.) conduit à l'émergence de nouvelles formes d'autocatalyse à compartiments multiples. Nous passons en revue différents cadres théoriques pour l'évolution chimique en Ch.6. Ces cadres se concentrent sur des chimies et des structures de réseaux spécifiques, et dépendent sensiblement du niveau de description. Ces approches peuvent être réunis et étendues, ils sont compris dans notre cadre général pour l'autocatalyse. On décrit des aspects thermodynamiques et structuraux de l'évolution autocatalytique, qui se produit par des processus de branchement aux taux microscopiques. L'extension à l'autocatalyse à compartiments multiples conduit à des comportements écologiques (syntrophie, parasitisme, coopération). Ch.7 étudie la formation de copolymères longes (par adsorption, recombinaison, ligature chimiquement activé). On trouve les coûts thermodynamiques pour la génération dissipative des séquences aléatoires. Ceci introduit des bornes énergétiques sur les scénarios qui reposent sur l'apparition des structures rares. Ch.8 présente un cadre statistique pour la compartimentation transitoire. Cette forme de sélection multi-niveaux n'a pas de lignées : un compartiment survivant disparaît après sa croissance et sélection. Cela fait que son contenu peut être multiplié par plus qu'un facteur de 2 (expérimentalement : $> 10^6$). Ce mécanisme s'est avéré capable de surmonter les invasions parasitaires, d'induire une coopération et d'abaisser les seuils d'erreur. Le bruit de composition suit de la cinétique de croissance. La polymérisation peut réduire considérablement ce bruit, ce qui peut favoriser sa sélection. À ce niveau de description, les parasites donnent lieu à une nouvelle catastrophe lié à la complexation. Ch.9 présente un nouveau scénario, fondé sur des mécanismes. Le scénario est basé sur la structure de la chimie, l'autocatalyse et sélection à plusieurs niveaux, mais ne spécifie pas de molécules : elles peuvent être introduites a posteriori. Nous fournissons une base sur laquelle des scénarios rigoureux pour l'avenir peuvent être construits.

MOTS CLÉS

Thermodynamique, Information, Réseaux Chimiques, Origine de la Vie, Compartiments, Évolution Chimique.

ABSTRACT

The academic field of Origins of Life seeks to explain abiogenesis: how abiotic matter can be transformed to living systems. Recent decades have seen a substantial development of prebiotic scenarios: hypotheses on the place, chemistry and physical mechanisms of abiogenesis. In this thesis, we introduce rigorous frameworks, for the systematic study of physical aspects of abiogenesis. These frameworks build upon recent insights in nonequilibrium thermodynamics, chemical reaction networks and group selection. They stress thermodynamic consistency and the fundamental structure of chemistry. In Ch.1, a critical introduction to the field of origins of life is given, highlighting what we truly know, what popular scenarios assume and historical developments that have shaped the current thinking. In Ch.2, the theoretical framework of chemical networks is introduced. We extend the framework with new tools and a criterion for a universal description of chemistry: nonambiguity. In Ch.3, we describe ways to make networks 'open': chemostats, CSTR, serial transfer, diffusive coupling to compartments. The concept of chemostats is extended to 'composite chemostats', which chemostats combinations of species. We correct the zeroth law of thermodynamics, which is shown to be violated for conserved integer quantities in stochastic thermodynamics. In Ch.4, we illustrate the concept of information in chemical networks, using a single-molecule information engine that also works in the macroscopic limit, and that extracts work from the racemization of enantiomerically pure molecules. We then move to the opposite process: purifying compositions. We illustrate a variety of chemical networks that achieve purification, and we discuss their tradeoffs. In Ch.5, we derive universal criteria for catalysis and self-replication for unambiguous chemical networks. The addition of exchange processes between phases and compartments (diffusion, evaporation, partitioning etc.) leads to emergent new forms of multicompartment autocatalysis. In Ch.6, we review the concept of chemical evolution and some of the frameworks developed for it. These frameworks focus on specific chemistries and network structures, and we show that their interpretation critically hinges on the level of coarse graining. These approaches, often treated as mutually exclusive, are united, extended and encompassed by our general framework for autocatalysis. We study structural and thermodynamic aspects of autocatalytic evolution in a single reactor, which occurs by branching processes built up from microscopic rates. The extension to multicompartment autocatalysis leads to new emergent ecological behavior (syntrophy, parasitism), favoring cooperation and spatial confinement. In Ch.7, we study the thermodynamics of making long polymers in various out-of-equilibrium situations (adsorption, recombination reactions, chemically activated ligation). We derive thermodynamic costs for the dissipative generation of random copolymer sequences. This allows to place energetic bounds on scenarios that rely on the appearance of rare structures. In Ch.8, we set up a statistical framework to study transient compartmentalization. This new form of multilevel selection has no lineages: surviving compartments vanish after growth and selection, which means contents may multiply by more than a factor 2 (experimentally: $> 10^6$). This mechanism is shown capable of overcoming parasite invasions, induce cooperation and lower error thresholds. Compositional noise is derived from growth kinetics. Polymerization can drastically decrease such noise, which can improve selection. On this level of description, a new parasite catastrophe emerges: a complexation catastrophe. In Ch.9, we formulate a mechanism-based scenario. The scenario is based on the structural features of chemistry, multicompartment autocatalysis and multi-level evolution, but does not specify any molecules: they can be introduced a posteriori. We provide a foundation on which rigorous future scenarios can be built.

KEYWORDS

Thermodynamics, Information, Chemical Networks, Origin of Life, Compartments, Chemical Evolution.