Discovery of a new type of regulator of DNA replication : the glycolytic enzyme pyruvate kinase in Bacillus subtilis

Steff Horemans

université PARIS-SACLAY

NNT : 2019SACLE040

Discovery of a new type of regulator of DNA replication : the glycolytic enzyme pyruvate kinase in *Bacillus subtilis*

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'université Évry val-d'Essonne

École doctorale n°577 Structure et dynamiques
des systèmes vivants (SDSV)
Spécialité de doctorat: sciences de la vie et de la santé,

Thèse présentée et soutenue à Évry, 11/12/2019, par

Horemans Steff

Composition du Jury :

Bianca Sclavi DR, Université Pierre et Marie Curie (– Laboratory of Computational and Quantitative Biology)	Président
Nathalie Declerck DR2, CNRS (– Centre de Biochimie Structurale)	Rapporteur
Hannu Myllykallio Professeur chargé de cours, École polytechnique (– Laboratory for optics and Biosciences)	Rapporteur
Bianca Sclavi DR, Université Pierre et Marie Curie (– Laboratory of Computational and Quantitative Biology)	Examinateur
Dominique Le Coq Chercheur, Micalis, INRA (– Adaptation bacteriénne)	Examinateur
Jean-Luc Ferrat Maitre de conferences, Institut de Biologie Intégrative de la Cellule (– Biologie des Génomes)	Examinateur
Laurent Jannière CR1, Génoscope (– UMR 8030 Génomique Métabolique)	Directeur de thèse

Thèse de doctorat

Acknowledgements

The doctoral thesis you are about to read would not have been possible without the hard work and support of many people. Whether through financial, professional or other support, these people have left their mark on the work presented here.

I would like to thank the university of Évry Val d'Essonne, the university of Paris-Saclay and the doctoral school SDSV for the financial support that made this work possible as well as the administrative support they provided.

Additionally, I would like to thank my reporters, Hannu Myllykallio and Nathalie Declerck for taking the time to read this long document and providing feedback and I would like to express the hope that they will the read stimulating. I would also like to thank Jean-Luc Ferat, Dominique Le Coq and Bianca Sclavi for taking the time to be on the jury of the thesis defence and express the hope that they will consider this time well spent.

A doctoral thesis is not produced in a vacuum, but requires scientific input from many sources. I would like to thank François Képès for accepting me as his PhD student and the many stimulating discussions we had about science. I would also like to thank Bianca Sclavi and Olivier Espéli for being on my thesis committee and guiding me through these last three years. I would like to thank the many collaborators, who have significantly enriched this project. Specifically, I thank Alexandria Holland, Matthaios Pitoulias and Panos Soultanas, our collaborators from Nottingham who have contributed the biochemical work presented here. Next, I would like to thank our collaborators from Genoscope, namely Gabor Gyapay, Nathalie Vega-Czarny and Marcel Salanoubat who have contributed the flow cytometry results and Alain Perret, Christophe Lechaplais and Marcel Salanoubat who are going to contribute metabolomic studies. The work presented here would not be half as good without them and I would like to thank Marcel Salanoubat specifically for enabling these crucial collaborations.

But mostly, I am extremely grateful to current supervisor, Laurent Jannière without whom this work would not have been possible. He graciously accepted me as his student on his project after the retirement of François, taught me everything I needed to know and worked alongside me througout this final year. I feel blessed to have had the opportunity to push the boundaries of science just a little further with him.

I also would like to thank my current colleagues Nazim Sarica and Morgane Champleboux for their help on my project and the many discussions we had on life and other topics. I would also like to thank my former colleagues and especially Daniel Trejo Banos and François Bucchini for their support both inside and outside the lab. I would like to thank Sylvie Bobelet for guiding me through the many administrative messes I found myself in during my tenure here. All these people have been instrumental in

achieving this succes and helping me maintain my sanity.
In my personal life, I am extremely grateful to my girlfriend Hanne Verhaegen, my parents Frank Horemans and Carina Vermeulen and my brother Stijn Horemans, who have all supported me through the most difficult moments of this PhD.
To conclude, I would like to extend a big thank you to all the people whose vital contribution I am inevitably forgetting.

Table of contents

Table of figures

Résumé français de la thèse

La réplication de l'ADN est un processus biologique complexe et universel qui a pour fonction de transmettre entièrement et fidèlement le patrimoine génétique au cours des générations cellulaires. Pour assurer cette tâche, la réplication doit être étroitement coordonnée au cycle cellulaire. De la bactérie à l'homme, cette coordination est principalement assurée par l'action concertée de nombreux systèmes partiellement redondants qui, en contrôlant l'activité d'éléments spécifiques d'initiation de la réplication (protéines et origines de réplication), assurent une duplication du génome une et une seule par cycle cellulaire. Cependant, la réplication de l'ADN est aussi sous le contrôle du métabolisme. Chez les bactéries, ce niveau supérieur de régulation couple le taux global de réplication au taux de croissance. Il est assuré en modulant la fréquence d'initiation de la réplication et la vitesse des fourches de réplication en fonction de la richesse du milieu de culture. Le résultat de ce niveau supérieur de régulation est un fenêtrage temporel précis et reproductible de la synthèse d'ADN dans le cycle cellulaire. Chez la levure de boulanger (et peut-être tous les eucaryotes), le contrôle métabolique de la réplication confine la synthèse d'ADN dans une phase particulière d'un cycle métabolique intracellulaire qui est répété entièrement plusieurs fois par cycle cellulaire.

En dépit d'efforts importants depuis les années 60, le mécanisme et les éléments du contrôle métabolique de la réplication restent pour l'essentiel inconnus. Les hypothèses classiques stipulant que ce contrôle chez les bactéries dépend de la concentration de formes actives de facteurs d'initiation, d'une carence partielle des ADN polymérases en précurseurs ou de la concentration d'une alarmone (ppGpp) qui signale l'état métabolique des cellules et bloque la réplication à haute concentration, ont été tour à tour critiquées. Des hypothèses alternatives ont alors vu le jour dans lesquelles il est proposé que le contrôle de la réplication est un phénomène multifactoriel qui varie avec le milieu et qui dépendrait de systèmes capables de sentir le métabolisme et de communiquer cette information à la machinerie de réplication. Parmi les différentes processus cellulaires connus ou supposés reliés à la réplication se trouve le métabolisme central carboné (MCC), un groupe de réaction hautement conservé dans le monde vivant dont la fonction première est de brûler les sources de carbone pour générer l'énergie et les briques nécessaires à la production de biomasse. Puisque le MCC sent les ressources disponibles dans l'environnement et alimente de façon adaptée les principales voies de macrosynthèse, il est supposé que cette région métabolique pourrait également envoyer des signaux que les cellules utiliseraient pour ajuster de façon concertée le taux de réplication à la vitesse de croissance. Un tel processus de régulation a été récemment découvert pour la division cellulaire.

Plusieurs études suggèrent un rôle du MCC dans la réplication de l'ADN de la bactérie à l'homme. Chez la bactérie *B. subtilis*, qui est un organisme de référence dans ce domaine, il a été mis en évidence dès les

années 80 qu'une sous unité de la pyruvate déshydrogénase, une enzyme du MCC qui produit de l'acétyl-CoA à partir du pyruvate, inhibait l'initiation de la réplication dans un extrait membranaire en se fixant à une région riche en AT située près de l'origine de réplication. De plus, cette sous-unité et d'autres provenant d'enzymes métaboliques apparentées, ont été vues interagir avec des enzymes de réplication. Ces études suggèrent donc qu'une enzyme du MCC peut agir sur la réplication de l'ADN grâce à des interactions directes entre protéines métaboliques et enzymes de réplication.

Plus récemment, il a été montré chez *B. subtilis* que le contrôle métabolique de la réplication agit pour une bonne part sur l'étape d'initiation de la réplication et qu'il impliquerait des liens entre réplication et plusieurs fonctions cellulaires principales dont le MCC. Une étude assez large conduite précédemment au laboratoire a également montré que 11 gènes du MCC sont important pour la réplication. Ces gènes sont groupés dans la partie terminale de la glycolyse et dans un groupe de réactions situé juste en aval et appelé overflow. L'implication de ces gènes dans la réplication est dépendante du milieu de culture: si certains gènes sont nécessaires dans un seul milieu, d'autres sont requis dans deux ou trois milieux différents. De plus, un même gène peut être important pour l'initiation dans un milieu et pour l'élongation (la phase de synthèse d'ADN) dans un autre milieu. Dans un autre travail, l'équipe a démontré que les gènes de la partie terminale de la glycolyse sont génétiquement liés à trois enzymes de réplication qui jouent un rôle important dans l'initiation et l'élongation de la réplication. Ensemble, ces résultats suggèrent que le MCC est important pour un fenêtrage optimale de la réplication dans le cycle cellulaire et que ce processus dépend de nombreux liens entre MCC et réplication.

Dans ce travail de thèse, nous avons étudié chez *B. subtilis* le rôle de la pyruvate kinase PykA dans la réplication. Ce travail a été motivé par les diverses publications résumées ci-dessus qui suggèrent que PykA est un déterminant important du lien entre métabolisme et réplication. PykA est une enzyme du MCC qui catalyse la dernière réaction de la glycolyse et convertit le phosphoénolpyruvate et l'ADN en pyruvate et ATP. Chez *B. subtilis*, cette enzyme comprend deux domaines. Le premier, N-terminal, correspond au domaine catalytique classique. Le second, C-terminal, n'est pas essentiel à l'activité catalytique. De façon intéressante, il est fortement homologue au domaine « PEP utilizer » (nommé ici PEPut) présent dans d'autres enzymes métaboliques. Dans le système principal de transport des sucres (le système PTS pour phoshpoenolpyruvate:carbohydrate phosphotransferase), l'enzyme EI se lie au phosphoénolpyruvate qui est utilisé comme donneur pour phosphoryler de façon transitoire un résidu histidine dans le domaine PEPut de EI. Le groupement phosphoryl est ensuite successivement transféré aux protéines Hpr, EII pour être finalement transféré sur la molécule de sucre entrante. De façon remarquable, le résidu histidine phosphorylé dans EI est inclus dans un motif Thr-Ser-His (TSH) hautement conservé dans les différentes formes de PEPut et dans PykA. Puisque, dans une premier temps, PykA comme EI a pour substrat le phosphoénolpyruvate et porte un domaine PEPut potentiellement

fonctionnel et puisque, dans un second temps, PykA est importante pour la réplication, nous avons étudié la réplication dans différents mutants de PykA afin de rechercher des éléments dans PykA impliqués dans la communication avec la synthèse d'ADN.

Le travail a été réalisé dans un milieu néoglucogénique où l'activité catalytique de PykA ne joue pas de rôle important dans la croissance et le métabolisme cellulaire. Des expériences contrôles ont permis de confirmer ces points dans le milieu utilisé (milieu minimum complémenté en malate et hydrolysat de caséine) et pour l'ensemble des mutants testés. Dans un premier temps, nous avons étudié la réplication de l'ADN dans une souche délétée du gène *pykA*. Les différentes technologies utilisées ont permis de déterminer (i) le taux global de réplication (ratio *ori/ter* déterminé par PCR quantitative (qPCR)), (ii) le temps nécessaire pour répliquer le chromosome bactérien (période C ; par analyse de fréquence de marqueurs par qPCR), (iii) la vitesse de progression des fourches de réplication (en utilisant une formule qui prend en compte la taille du chromosome et la durée de la période C), (iv) le nombre de séquence origines par cellule (par cytométrie en flux) et (v) l'âge d'initiation de la réplication (en utilisant les données de cytométrie et le temps de génération). La mesure de ces paramètres de cycle cellulaire dans des cultures en phase exponentielle de la souche sauvage et du mutant ont révélé de nombreuses différences : dans le mutant, le ratio *ori/ter* est plus élevé, la période C est allongée, la vitesse des fourches de réplication et le nombre de séquences origines par cellules sont réduites et l'âge d'initiation est retardé. Tout cela montre que le système qui positionne dans le temps la réplication de l'ADN dans le cycle cellulaire est perturbé en absence de la pyruvate kinase. Ainsi, une enzyme du MCC est importante pour le contrôle temporel de la réplication.

Nous avons ensuite étudié la réplication dans des mutants du site catalytique de PykA. L'une des souches était délétée du domaine catalytique alors que d' autres portaient une mutation ponctuelle qui empêche la fixation du phosphoénolpyruvate, de l'ADP ou, simultanément, du phosphoénolpyruvate et du Mg^{2+} Une souche mutée également ponctuellement dans un élément de changement conformationnel a également été utilisée ainsi qu'une autre qui porte une petite délétion de 27 acides-aminés dans la région impliquée dans le changement conformationnel. Cette petite délétion a pour particularité d'être liée génétiquement à trois enzymes de réplication impliquées dans l'initiation et l'élongation de la réplication. Des expériences contrôles de croissance et de métabolomiques ont confirmé que toutes ces mutations affectent profondément l'activité catalytique de PykA. La mesure des paramètres de réplication ont montré que les mutations ponctuelles dans le site de fixation à l'ATP et dans l'élément de changement conformationnel n'ont pas d'effet sur le contrôle temporel de la réplication. Ce contrôle est par contre perturbé dans les autres mutants qui présentent une augmentation du ratio *ori/ter*, un allongement de la période C, une diminution de la vitesse de réplication ainsi que probablement, une avancée de l'âge d'initiation. Comme toutes les mutations inhibent l'activité catalytique de façon similaire, ces résultats montrent que c'est la

présence de la protéine PykA et non son activité métabolique ou son rôle dans le métabolisme cellulaire qui est importante pour le contrôle temporel de la réplication. De plus, ces travaux identifient trois régions dans PykA importantes pour la réplication. Deux d'entre elles affectent la fixation du phosphoénolpyruvate indiquant que le contrôle temporel de la réplication pourrait dépendre de la concentration de ce métabolite du MCC.

Nous avons ensuite déterminé les paramètres de réplication dans des mutants de PEPut. Ces mutants étaient soit délétées de PEPut, soit mutées ponctuellement dans le motif TSH de PEPut. Dans ces dernières souches, les résidus Thr, Ser et His étaient individuellement mutés en Ala ou Asp, un acide-aminé utilisé pour mimer la phosphorylation dans les protéines. L'analyse des paramètres de réplication montre que la mutation Thr> Asp a une réplication fortement perturbée : le ratio *ori/ter* et la période C sont fortement réduits, la vitesse de réplication est fortement accélérée, le nombre de séquences origines par cellule est légèrement réduit et l'âge d'initiation est extrêmement retardé. Tout ceci montre une forte perturbation du contrôle temporel de la réplication dans ce mutant. Puisque le mutant Thr>Ala montre une augmentation à peine détectable du ratio *ori/ter* et puisque le résidu Asp mime la phosphorylation, ces résultats suggèrent (i) qu'une phosphorylation permanente du résidu Thr dans le motif TSH interfère dramatiquement avec le contrôle temporel de la réplication alors que (ii) une phosphorylation transitoire de ce résidu est requise pour permettre un contrôle temporel sauvage de la réplication. Par ailleurs, les résultats suggèrent que les résidus Ser et His du motif TSH ne sont pas impliqués dans la réplication dans les conditions métaboliques utilisées puisque les mutants Ser>Ala et His>Ala ont un ratio *ori/ter* sauvage. Cette déduction est contredite par la légère réduction du ratio observée dans les mutants Ser>Asp et His>Asp à moins de supposer que la présence d'un signal de phosphorylation au niveau des résidus Ser et His est perçue de façon erronée par la cellule comme une phosphorylation partielle du résidu Thr, provoquant une faible réduction du ratio. Enfin, dans le mutant délété de PEPut, une réduction modéré du ratio *ori/ter* est détectée suggérant que l'absence de PEPut est moins délétère pour la réplication qu'une phosphorylation permanente du résidu Thr dans le motif TSH.

Dans leur ensemble, les résultats montrent que les domaines Cat et PEPut fusionnés dans la protéine PykA ont des influences opposées sur le contrôle temporel de la réplication : le domaine Cat prévient une initiation précoce et une vitesse lente des fourches de réplication alors que la limitation du niveau de phosphorylation de la Thr dans le motif TSH de PEPut prévient une initiation tardive et une vitesse très élevée des fourches de réplication. De plus, puisqu'aucun rôle métabolique n'a été assigné à PEPut, ces résultats confirment que l'engagement de PykA dans le contrôle temporel de la réplication dépend de la protéine elle-même et non de son activité métabolique ni de son rôle dans le métabolisme cellulaire.

Pour mieux comprendre le fonctionnement de PEPut et son interface avec Cat, nous avons ensuite

construit des souches codant pour des molécules de PEPut non fusionnées à Cat (molécules PEPut libres). Pour cela, la phase ouverte de lecture de PEPut et son site de fixation au ribosome ont été mis sous le contrôle d'un promoteur inductible à l'IPTG et la construction appelée HyPEPut a été insérée dans un locus du chromosome éloigné de *pykA*. Cette insertion a été effectuée dans la souche sauvage et dans différents mutants de *pykA*. Des souches contenant des formes mutées de HyPEPut ont également été construites. Nous avons dans un premier observé que la synthèse de molécules de PEPut libres est associée à une nette réduction du ratio *ori/ter* dans le contexte sauvage, délété de *pykA* et délété de PEPut (c'est à dire codant pour Cat seulement). Cet effet inhibiteur, qui augmente avec le niveau d'expression de HyPEPut, n'est par contre pas observé dans un contexte produisant PEPut à partir du promoteur naturel de *pykA*. Comme le promoteur inductible de HyPEPut permet la production de beaucoup moins de protéines que le promoteur naturel de *pykA* (environ 3.000 contre 30.000 molécules), ce résultat montre que la molécule PEPut non fusionnée à Cat a un effet négatif sur le ratio *ori/ter* à des concentrations modérées mais pas à des concentrations élevées et que cet effet est dominant sur de fortes concentrations de PEPut quand ce peptide est fusionné à Cat (comme dans le contexte sauvage). Ils montrent également que les formes libres de PEPut ne complémentent pas l'absence de PEPut dans PykA puisque le ratio *ori/ter* est faible dans la souche produisant séparément Cat et PEPut. Ceci suggère que PEPut et Cat doivent être fusionnés dans une molécule unique pour un contrôle temporel sauvage de la réplication. Les effets opposés de l'expression de HyPEPut sur le ratio *ori/ter* dans des contextes où PEPut est massivement produite à partir du promoteur naturel de *pykA* dans un état fusionné à Cat (inhibition du ratio) ou libre (pas d'inhibition du ratio) suggèrent de plus que les molécules libres de PEPut sentent un facteur limitant inhibiteur du ratio mais pas les molécules fusionnées à Cat. On peut alors proposer qu'à des fortes concentrations, seulement une proportion des molécules PEPut libres reçoivent un signal du facteur limitant inhibiteur rendant le système de signalisant inefficace alors que la majorité de ces molécules reçoivent le signal quand elles sont à faible concentration, activant efficacement le système de signalisation pour *in fine* réduire le ratio *ori/ter*.

Nous avons ensuite analysé l'effet sur le ratio *ori/ter* de mutations dans le motif TSH de HyPEPut en contexte sauvage. Les résultats montrent que le ratio est aussi faible dans les cellules exprimant HyPEPut sauvage qu'avec la mutation Thr>Ala, alors qu'il est encore plus réduit dans en contexte HyPEPut avec la mutation Thr>Asp. Par contre, le ratio n'est plus inhibé quand le résidu His de HyPEPut est remplacé par Ala ou Asp. Ceci montre que le résidu His des molécules libres de PEPut est essentiel pour recevoir et envoyer le signal du facteur limitant inhibiteur. Ceci attribue une fonction au résidu His du motif TSH de PEPut qui n'avait pas été détectée quand ce domaine est fusionné à Cat. Par ailleurs, comme dans le contexte où PEPut et Cat sont fusionnés, la mutation Thr>Asp augmente l'effet inhibiteur de PEPut libre. Ceci renforce l'hypothèse selon laquelle la phosphorylation du résidu Thr est importante pour le bon

fonctionnement de PEPut et suggère que les résidus Thr et His du motif TSH de PEPut assurent des fonctions importantes distinctes dans le contrôle temporel de la réplication.

Enfin, nous avons étudié l'effet de l'expression de HyPEPut dans des mutants catalytiques de PykA. Nous avons observé que, comme dans une souche sauvage, l'expression de HyPEPut est associée à une nette réduction du ratio *ori/ter* dans tous les mutants catalytiques de PykA sauf dans celui contenant une petite délétion de 27 acides-aminés. Dans ce cas, l'effet inhibiteur de HyPEPut est abrogé. Ceci établit un lien fonctionnel étroit entre 27 acides aminés du domaine Cat de PykA (coordonnées 208 à 234), le résidu His de PEPut, le facteur limitant inhibiteur et les trois enzymes de réplications liées génétiquement à ces 27 résidus de PykA.

Dans une collaboration avec Panos Soutanas (Nottingham, UK), nous avons ensuite recherché d'éventuels effets directes de PykA sur l'activité d'enzymes de réplication *in vitro*. Ce travail a été centré sur les trois enzymes de réplication liées génétiquement à PykA : l'hélicase DnaC, la primase DnaG et l'ADN polymérase DnaE. Ces trois enzymes forment un complexe qui joue des rôles importants dans l'initiation et l'élongation de la réplication. L'hélicase DnaC ouvre l'ADN dans la région origine au moment de l'initiation de la réplication et sépare les brins d'ADN dans la fourche de réplication pendant la phase d'élongation. La primase DnaG synthétise des amorces ARN au moment de l'initiation et dans les fourches de réplication. Ces amorces sont utilisées pour démarrer la synthèse d'ADN à l'origine de réplication et pour synthétiser le brins discontinu dans la fourche de réplication. L'ADN polymérase DnaE allonge les amorces ARN produites par DnaG à l'origine ainsi que dans les fourches de réplication pour assurer une synthèse partielle ou totale du brin discontinu. Les enzymes de réplication ont été purifiées et leurs activités testées en présence ou en absence de PykA comme dans nos études précédentes. En utilisant un système d'expression et de purification de *E. coli* classiquement utilisé en biologie moléculaire, nous avons purifié la forme native de PykA de *B. subtilis*. A l'aide de tests biochimiques (test de Michalis Menten, test de Hill et spectroscopie de masse), nous avons montré dans un premier temps que la forme purifiée de PykA est métaboliquement active et forme, comme attendu, un tétramère très stable.

Des tests de réplication ont ensuite été conduits avec l'ADN polymérase DnaE. Dans des mélanges réactionnelles contenant différentes matrices de réplication, nous avons ajouté des concentrations faibles de DnaE en combinaison ou non avec des concentrations équimolaires de PykA. Les tests montrent une stimulation de la synthèse d'ADN en présence de PykA dans toutes les conditions utilisées. Des expériences contrôles ont montré que cette stimulation était spécifique de PykA car elle n'était pas observée quand l'enzyme métabolique était remplacée par la BSA dans les mélanges réactionnels. De plus, des expériences de titration associées à des tests de décalage sur gel (gel shift) ont montré que la

stimulation de la synthèse d'ADN par PykA n'était pas due à une stimulation de la fixation de DnaE sur la matrice d'ADN. Il est conclu de ces expériences que PykA stimule l'activité ADN polymérase de DnaE probablement via une interaction directe entre ces deux protéines. Puisque le domaine PEPut purifié ne provoque pas cette stimulation, celle-ci dépend probablement d'un contact direct entre DnaE et le domaine Cat de PykA.

Des tests permettant de mesurer l'activité hélicase de DnaC en présence ou en absence de PykA ont ensuite été conduits. Ils montrent que PykA a un effet inhibiteur marginal sur DnaC. Cependant, la stimulation de l'activité hélicase de DnaC par la primase DnaG n'est pas observée en présence de PykA. Ainsi, dans certains contextes, PykA peut inhiber significativement DnaC. Ensemble, ces résultats suggèrent que PykA peut modifier positivement ou négativement les activités d'enzymes de réplication de façon spécifique grâce à des interactions directes avec les enzymes de réplication.

En conclusion, nos résultats chez *B. subtilis* mettent en évidence des phénotypes et des activités non-métaboliques à une enzyme du MCC qui catalyse la dernière étape de la glycolyse : la pyruvate kinase. Cette enzyme semble donc dotée de fonctions cryptiques. Les phénotypes ont été observés dans des conditions de culture néoglucogénique où l'enzyme est produite en abondance (environ 30.000 molécules par cellule) mais où sa fonction métabolique n'est pas requise pour la croissance cellulaire. Leur analyse met systématiquement en évidence des défauts de réplication. Suivant la position des mutations dans la protéine, des défauts opposés sont observés. Dans les mutants du site catalytique, l'initiation de la réplication est avancée dans le cycle cellulaire et la vitesse des fourches de réplication est ralentie provoquant un allongement du temps nécessaire à la réplication du chromosome. A l'inverse, dans les mutants du domaine PEPut, l'initiation est retardée et la vitesse des fourches de réplication accélérée. Ces défauts de réplication qui changent le positionnement dans le temps de la réplication, indiquent donc que les activités cryptiques de la pyruvate kinase chez *B. subtilis* sont importantes pour un bon contrôle temporel de la réplication dans le cycle cellulaire. Le mécanisme sous-jacent reste inconnu mais il semble dépendre de plusieurs facteurs. L'un de ces facteurs pourrait être le phosphoenolpyruvate puisqu'au moins deux des trois mutations catalytiques affectant la réplication inhibent la fixation de ce métabolite à PykA. Un autre facteur pourrait être le niveau de phosphorylation du résidu Thr situé dans le motif TSH de PEPut. Il semble que plus ce résidu est phosphorylé, plus l'initiation est retardée et la synthèse d'ADN rapide. Un troisième facteur serait un inhibiteur limitant non identifié. Enfin, l'étude des phénotypes associés à une production modérée de PEPut non fusionné au domaine Cat indique des liens fonctionnels étroits entre 27 acides aminés du domaine Cat, le résidu His du motif TSH de PEPut et le facteur inhibiteur non identifié.

Des études effectuées dans des conditions métaboliques différentes (un shift métabolique) indiquent que

la réponse de la réplication aux variations nutritionnelles peut faire appel à des déterminants et donc des activités cryptiques différentes de PykA. En effet, dans ces nouvelles conditions de culture, le changement de réplication dépend du domaine PEPut, et plus particulièrement du résidu His de son motif TSH, mais pas du domaine Cat. Ceci montre que le comportement des activités cryptiques de PykA change avec le métabolisme cellulaire et donc que PykA agit comme une plateforme qui évalue le métabolisme pour envoyer des signaux adaptés à la réplication afin de positionner dans le temps la réplication de l'ADN dans le cycle cellulaire.

Plusieurs données indiquent que PykA agit sur le contrôle de la réplication en ayant une action directe ou indirecte sur trois enzymes clé de réplication : l'hélicase DnaC, la primase DnaG et l'ADN polymérase DnaE. Premièrement, nous avons montré précédemment que, dans un milieu glycolytique riche (LB), (i) le gène *pykA* est important pour l'initiation de la réplication, (ii) que 27 acides aminés du site catalytique de PykA sont génétiquement liés à DnaC, DnaG et DnaE et que (iii) ce lien génétique pouvait être accompagné de changements conformationnels dans ces enzymes de réplication. Ici, nous montrons *in vitro* que la protéine PykA purifiée change les propriétés de réplication de DnaE et DnaC probablement via des interactions directes et spécifiques entre PykA et ces enzymes de réplication.

Dans son ensemble, ce travail apporte toute une série d'évidences conduisant à la conclusion que, chez la bactérie *B. subtilis*, la pyruvate kinase est dotée de plusieurs activités cryptiques qui jouent un rôle direct ou indirect dans le contrôle temporel de la réplication. Ce système de régulation peut impliquer des effets opposés orchestrés par les domaines Cat et PEPut de l'enzyme et ces effets seraient médiés par des voies de signalisations pouvant impliquer la concentration intracellulaire du phosphoénolpyruvate et des événements de phosphorylation. Ces observations attribuent donc pour la première fois une fonction de régulation temporelle de la réplication de l'ADN à une enzyme métabolique. Etant donné que le contrôle temporel de la réplication est universel, que les enzymes du MCC sont fortement conservées et que de nombreuses études mettent en évidence des liens entre le MCC et la réplication, nous proposons ici que des enzymes du MCC présentent des activités cryptiques de réplication nécessaires au contrôle temporel de la réplication dans la plupart des organismes. Par ailleurs, il est bien établi que l'altération du contrôle temporel de la réplication augmente l'instabilité génétique et que le cancer présente deux caractéristiques précoces et essentielles : une altération du MCC (effet Warburg) et une instabilité génétique. Nous proposons donc ici que l'émergence de cancers pourrait résulter de changements dans le MCC qui, en altérant le contrôle temporel de la réplication, augmenterait l'instabilité génétique pour engager les cellules dans le processus de cancérisation.

Chapter 1: Introduction to the thesis

1. Introduction to the cell cycle in eukaryotes, prokaryotes and archaea

1.1 The cell cycle in eukaryotes

The purpose of the cell cycle is to ensure that key periodic biological processes for reproduction are performed every cycle in a timely fashion. These key processes include (i) build-up of the necessary energy and building blocks (ii) faithful and efficient replication of DNA (iii) segregation of DNA to the cell poles and (iv) division.

In eukaryotic cells, the cell cycle is divided into temporally separated phases, each responsible for a specific task. The cell cycle in eukaryotes consists of two main phases: interphase and mitosis and is reviewed in [1,2]. These different phases of the eukaryotic cell cycle are summarized in figure 1A.

The interphase itself is divided into three phases: Gap 1 (G_1), Synthesis (S) and Gap 2 (G_2). The G_1 phase is used to build up energy and building blocks such as nucleotides, proteins and many more required for the next phase of the cycle. During S phase, eukaryotic cells replicate their DNA. DNA is replicated from specific sequences called origins of replication and each eukaryotic chromosome has multiple origins and is linear. Specific regulatory mechanisms ensure that DNA replication occurs once and only once each cell cycle. Likewise, G_2 is used to prepare for mitosis and to make sure that DNA replication has been completed before going into the M-phase.

Mitosis covers the process of nuclear division. Mitosis itself is divided into (i) prophase, (ii) metaphase, (iii) anaphase and finally (iv) telophase. During the prophase, the nuclear membrane and nucleolus are disassembled into vesicles, the mitotic spindle (consisting of two centromeres and a network of microtubules) is assembled and the DNA is condensed into chromatids. During metaphase, the condensed chromatids are aligned along the cell equator; equidistant from the centromeres. During the anaphase, the spindle complex pulls the sister chromatids to opposite sides of the cell and segregates them. Finally, the effects of prophase are reversed during telophase. DNA decondenses, the nuclear membrane and the nucleolus are reformed and the spindle complex is disassembled. Additionally, cells that are not committed to the cell cycle are said to be in the phase of G_0. In short, the eukaryotic cell cycle is a multiphase ordered process culminating in the duplication of the cell.

Correct progression through these different phases is regulated mainly by a network of

threonine-serine kinases called cyclin dependent kinases (CDK)[1,2]. Three classes of CDK's exist depending on their time of activity: G_1, S phase and Mitotic CDK's.[2] CDK's themselves are continuously expressed throughout the cell cycle, but require the binding of cyclin proteins to be activated. The cyclin proteins however are periodically expressed in distinct phases of the cell cycle, hence rendering CDK's active at specific timepoints during the cell cycle. Additionally, the timing of activity of CDK's is further regulated by inhibitors of CDK activity that either prevent CDK-cyclin binding or inactivate the CDK-cyclin complexes[1]. Active CDK's then exert their control over the cell cycle by phosphorylating key target proteins that are important during each specific phase of the cell cycle[1]. For instance, phosphorylation of the protein Rb during early G_1 ultimately leads to the activation of two transcription factors that promote gene transcription of targets required for the S-phase, including cyclin A and E[3]. Cyclin A dependent CDK's stimulate replication initiation by phosphorylating DNA polymerases[4] and Cyclin E-CDK2 phosphorylates histone H1, which may be important for controlling chromosome condensation during replication[1]. Using this phosphorylation system, the CDK network assures timely progression through the cell cycle, as shown in figure 2.
Additionally, the eukaryotic cell cycle contains 3 major checkpoints that make sure that the previous steps have been executed correctly[1,2]. The restriction point is the point where a cell transitions from G_0 to G_1. After this point, the cell is committed to completing the entire cell cycle, regardless of any environmental signals it receives. The checkpoint at the transition between G_1/S ascertains that the cell has enough building blocks to finish replication and to check for DNA damage.

Likewise, the checkpoint at the transition between G_2/M makes sure that mitosis can proceed without delay and that no DNA damage was accumulated after replication and G_2. Finally, spindle checkpoint makes sure that chromatids are properly aligned with the spindle complex during the metaphase-anaphase transition. During these checkpoints, specific protein kinases capable of detecting the presence of DNA damage induce the expression of CDK inhibitors to arrest the cell cycle[1]. For instance, DNA damage during the G_1/S checkpoint is detected by the ATM protein, whose kinase activity depends on the presence of DNA damage. Next, ATM phosphorylates p53 which in turn activates transcription of p21, an inhibitor of CDK activity[5]. Together, the CDK network and checkpoint controls enable the orderly progression of the eukaryotic cell cycle through its different phases and ensure correct reproduction of the cell.

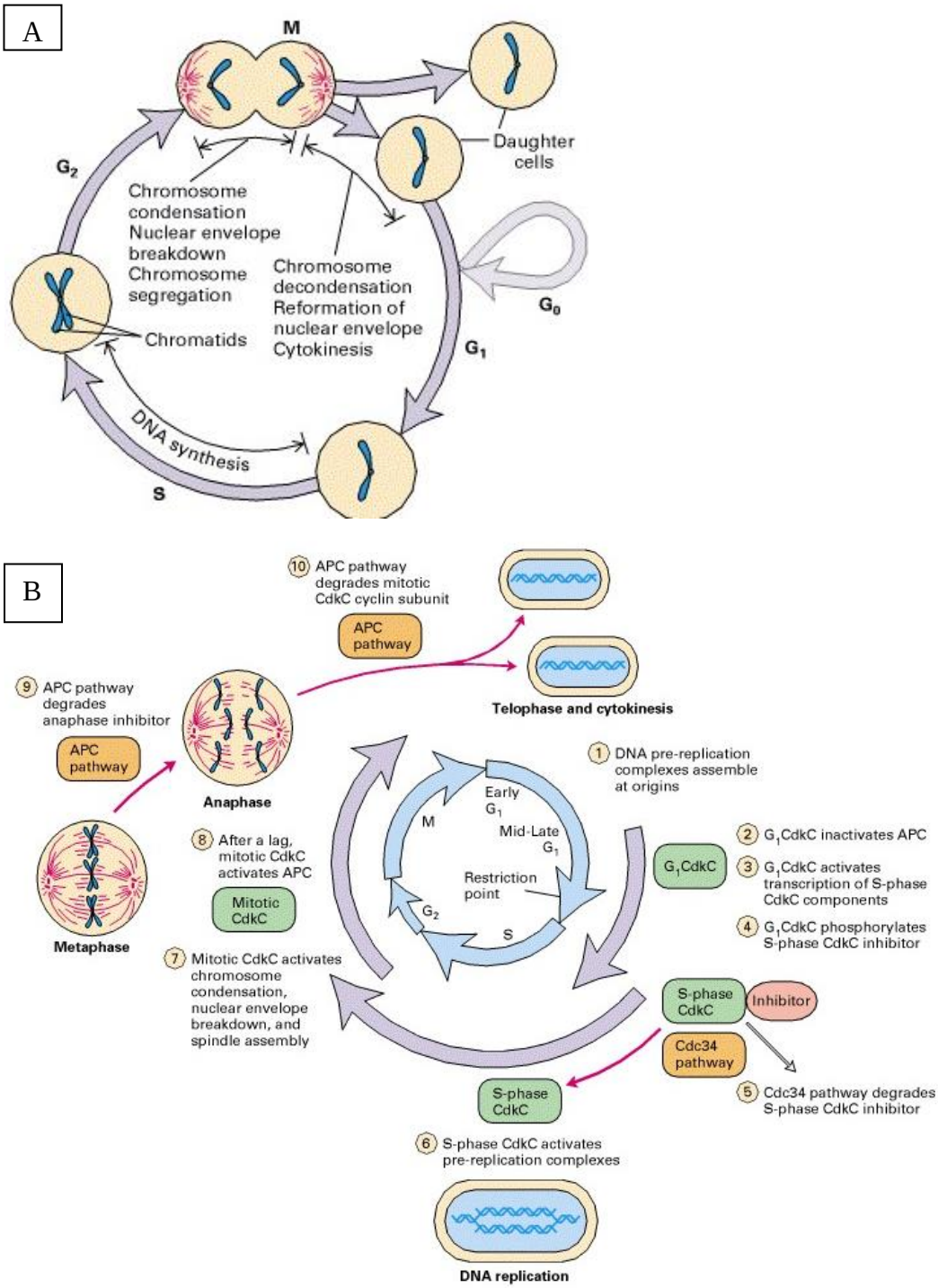


Figure 1 The cell cycle in eukaryotes

(A) The eukaryotic cell cycle: The fate of a single chromosome is depicted during the cell cycle. The eukaryotic cell cycle is an ordered sequence of temporally separated events. First, the cell enters the G1 phase, the phase between the birth of the cell and the initiation of DNA replication. Next, the cell duplicates its chromosomes during the S phase and enters G2 with twice as many chromosomes. During mitosis, the nucleoid is disassembled into vesicles, DNA is condensed and the spindle structure is formed (prophase) chromatids are aligned at the cell equator (metaphase) chromatids are segregated by the spindle network (anaphase) the events of prophase are reversed (telophase) and the cell divides. Taken from 2 Regulation of the

eukaryotic cell cycle using cyclin dependent kinases (CDK's): The eukaryotic cell cycle is regulated using a network of CDK's, which fall in three classes: G1 CDK's, S-phase CDK's and Mitotic CDK's. CDK's require the activity of cyclins. Taken from 2

1.2 The cell cycle in bacteria

In contrast to the eukaryotic cell cycle, the key biological processes of the bacterial cell cycle (preparation, replication, segregation and division) are not completely temporally separated into different phases. The bacterial cell cycle is reviewed in [6–8] and shown in figure 2A. Historically, the cell cycle in bacteria has been divided into three phases. The first phase, called the B-period, is the time between cell birth and the initiation of DNA replication and is akin to the G_1 phase in eukaryotes. The second phase, called the C-period, encompasses the time needed to fully replicate the chromosome. The DNA replication process in bacteria differs from the one in eukaryotes in a number of important ways. First, bacterial chromosomes are usually circular as opposed to linear chromosomes in eukaryotes. This allows DNA replication in bacteria to start from a single origin of replication (*oriC*) where it proceeds bidirectionally until a termination region in the opposite half of the chromosome is reache (*terC*). Additionally, bacterial chromosomes are segregated as they are replicated, with *oriC* moving to different sides of the cell and the bulk of the chromosome following. The ParAB system (Soj) and the SMC complex are responsible for this segregation. Two models of segregation have been observed. In the model bacterium *Escherichia coli (E. coli*), replichores separate along the transversal axis of the cell and occupy two different halves of the cell with *oriC* and *terC* located at the poles or mid-cell (depicted in figure 2B). Alternatively, replichores separate along the longitudinal axis and *oriC* and *terC* end up at different cell poles in the model bacterium *Bacillus subtilis* (*B. subtilis*). The replicated chromosomes are concatenated at the end of the C-period[7]. Chromosome separation is achieved using the DNA translocase FtsK, which colocalizes at the division site and recruits topoisomerase IV to decatenate the DNA and XerCD to promote homologous recombination between diff sites and chromosome separation. Finally the third phase, called the D-period, is the time between finalizing DNA replication and the end of cell division[8]. During this phase, bacteria start forming a Z ring made up of a tubulin homologue FtsZ when the middle of the cell is free of chromosomal DNA. The Z-ring defines the division site and recruits other cell division proteins, membrane invagination proteins and cell wall biosynthetic enzymes to complete division[8].

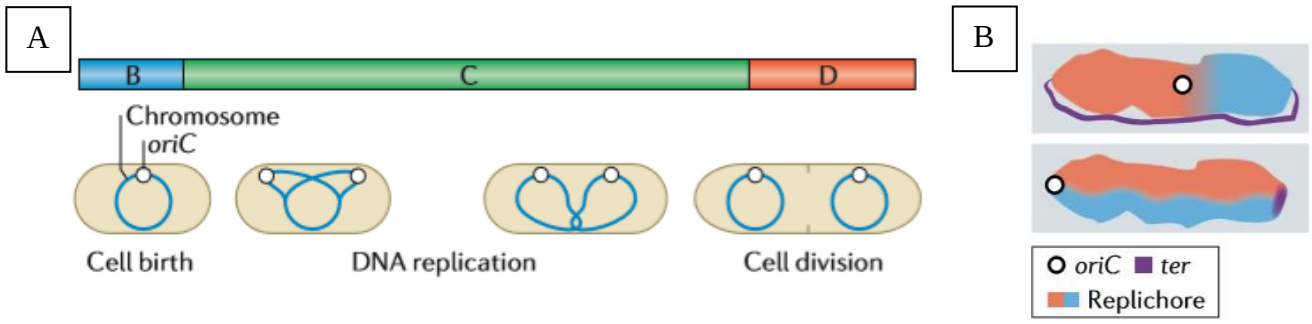


Figure 2 : The cell cycle in bacteria

(A) Events of the cell cycle in bacteria : The bacterial cell cycle (slow growing) consists of three periods. During the B period, bacteria grow and accumulate nutrients to initiate the cell cycle. During the C period, the bacterial chromosome is fully replicated bidirectionally from a single origin of replication (oriC) and simultaneously segregated to different halves of the cell (see figure 2B) The replicated chromosomes are concatenated which is resolved by homologous recombination at the end of the C period. During the D period, a Z-ring starts to form at mid-cell when it is free of genomic DNA. This Z-ring defines the division site and becomes the anchor point for proteins that will ultimately lead to division. Taken from [8]. (B) Orientation of chromosome segregation Top: Segregation orientation in the model organism E. coli Bottom: Segregation in the model organisms B. subtilis and C. crescentus. Taken from [8].

Another feature of the bacterial cell cycle that differs from the eukaryotic one is that no overarching regulatory network has been discovered that ensures orderly progression through the cell cycle[9]. CDK analogues do not seem to exist in bacteria [9]. Instead, a multitude of regulatory mechanisms that control the timing of DNA replication initiation, DNA replication elongation, cell division have been described in detail without a unifying principle that ensures the orderly progression of the cell cycle[9]. Hence, the cell cycle in bacteria is thought to consist of a series of independent steps, that are coordinated by a multitude of mostly unknown signaling systems[10]. Nevertheless, bacteria also have DNA damage response mechanisms like the SOS response (reviewed in [11]) that can arrest progression of the cell cycle and nucleoid occlusion factors like SlmA in *E.coli* and Noc in *B. subtilis* prevent the formation of the Z-ring before the chromosomes are segregated[8], akin to checkpoint control in eukaryotes.

1.3 The cell cycle in archaea

The cell cycle in archaea contains features of both eukaryotic and bacterial systems as well as unique features to the phylum[12]. The archaeal cell cycle is summarized in figure 1D. Like in eukaryotes, the cell cycle in archaea consists of a temporally ordered sequence of events. The archaeal cell cycle generally consists of a G_1 phase, an S phase, a G_2 phase and an M phase. The G_1 phase in archaea is generally short, whereas the S (30-35%) and the G_2 phase (>50%) take up the bulk of the cell cycle. Most archaea have a circular chromosome that is replicated from 1 or more origins of replication. During S-phase, all the origins fire synchronously. The replication machinery of all known archaea consists of homologues to eukaryotic proteins. After the completion of replication and during the G_2 phase, the chromosomes are resolved using the XerA homologous recombination system similar to the one in bacteria. This is remarkable since the multi origin replication of most archaea makes the decatenation of the replicated chromosomes more complex. This might be the reason why G_2 takes up the bulk of the cell cycle. During the M phase, the duplicated genomes are segregated and the cell divides. In the *Sulfolobales* order of archaea, chromosome segregation is achieved using the SegAB system, akin to ParAB in bacteria. Other segregation systems are currently unknown, but some species show rapid segregation of chromosomes, akin to the anaphase in eukaryotes, whereas others show a more gradual separation akin to the one in bacteria. Cell division systems are diverse in archaea. Some archaea divide using FtsZ, akin to bacteria. Others divide using the recently discovered Cdv system, which forms a spiral inside the membrane that constricts the cell similar to the activity of VP4, a vesicle forming protein in eukaryotes. Still other species have neither system, suggesting that other division systems remain to be discovered in archaea. The mechanisms that enable correct progression through the phases of the cell cycle remain poorly understood in archaea. However, checkpoint control systems have been reported in archaea[13,14].

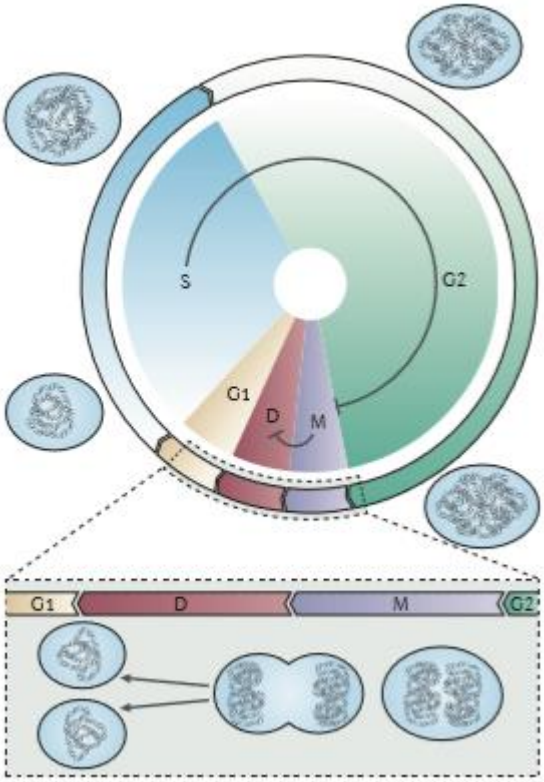


Figure 3: The cell cycle in archaea

Like the eukaryotic cell cycle, the key biological cell cycle processes of the archaeal cell cycle are separated into a G_1 phase, S phase, G_2 phase and M phase. During S phase, the circular archaeal chromosomes are replicated from multiple origins of replication using eukaryotic protein homologues. During G_2 phase, the multi-origin replicated chromosomes are decatenated using recombination systems homologous to bacterial ones and the cell is prepared for mitosis. During M phase, the archaeal chromosomes are segregated using segregation systems akin to bacterial ones. Finally, the cell is divided using a division machinery that is akin to either bacterial or eukaryotic systems or is entirely unique to archaea depending on the species studied. Taken from [12]

1.4 Proper timing of cell cycle events

The proper timing and execution of events throughout the cell cycle is vital for proper survival in any organism. Temporal defects in the cell cycle could lead to the formation of anucleate, inviable mini cells or cause division through the chromosomal DNA. Each of these defects reduces cellular viability. Furthermore, in multicellular organisms such as humans, temporal disruption of the cell cycle is associated with unrestrained cell proliferation and therefore cancer. Mutations in cyclins, CDK activators or inhibitors and CDK substrates are frequently observed in tumors[1] and several anti-tumor drugs have been developed to control CDK activity[115]. In all organisms, proper progression of the cell cycle starts by restricting the process of DNA replication to a specific time window in the cell cycle, also called DNA replication gating. The focus of this work is on the mechanisms that regulate DNA replication

gating in the model bacterium *B. subtilis*. In the next section, the process of DNA replication in *B. subtilis* is described and contrasted with the one in the other model bacterium *Escherichia coli*.

2. DNA replication in *B. subtilis* and *E. coli*

2.1 Basic mechanisms of DNA replication in *B. subtilis* and *E.coli*

The replication process is reviewed here[16] and consists of three phases: initiation, elongation and termination, depicted in fig. 4A. What follows is a model for the replication process in the bacterium *B. subtilis*.

DNA replication initiation itself consists of three major steps: (i) the assembly of the DnaA initiator oligomer at the origin of replication, (ii) the opening of the DNA and (iii) the recruitment of the other parts of the replication machinery at the unwound start site. First, the initiator protein DnaA binds to specific DNA sequences in the origin of replication called the DnaA boxes[17]. When bound to ATP, DnaA assembles into a right handed oligomer, thus forming the nucleoprotein complex[18,19]. The assembly of DnaA-ATP into the nucleoprotein complex triggers the unwinding of a nearby AT rich region called the DNA Unwinding Element (DUE)[20,21]. Next, the initiator proteins DnaB and DnaD bind to the origin[22] and remodel the DNA [23] prior to replisome loading[24]. DnaA-ATP then recruits the processive helicase DnaC in *B. subtilis*[24], which is loaded on the DNA by the helicase loader protein DnaI[24]. Afterwards, DnaC recruits the primase DnaG and the polymerase β-clamp DnaN, which then recruits the other elements required for replication such as the lagging strand polymerase DnaE and the leading strand polymerase PolC[25]. The process of DNA replication initiation is summarized in figure 4B.

During elongation, new DNA is synthesized by a molecular machine called the replisome. Single molecule studies suggest that the replisome consists of three DNA polymerase complexes PolC, DnaE and the primase DnaG, the hexamer DNA helicase DnaC, the β-clamp DnaN, three processivity clamps and a pentameric clamp loader[26]. DNA replication in any organism only occurs in one direction (5'end to 3'end) which results in one strand being continuously replicated (leading strand) and one strand being discontinuously replicated (lagging strand). The helicase DnaC forms a homohexameric ring that sits at the head of the replication fork on the lagging strand. Using the energy obtained from ATP hydrolysis, the DnaC ring mechanistically unwinds dsDNA. Separated DNA strands are then coated by the single stranded DNA binding protein (SSB) . This prevents both reassembly of the DNA duplex

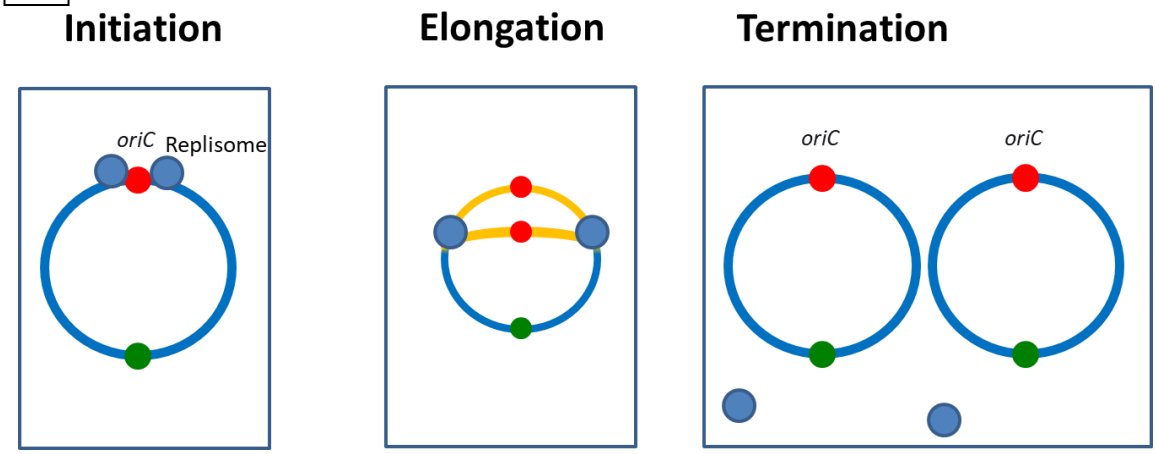
and protects the single stranded DNA against degradation by nucleases[27–29]. DnaG is bound to the helicase ring and catches single stranded DNA from the lagging strand for RNA primer synthesis right after duplex unwinding[30,31]. The single stranded DNA with its RNA primer is then handed off to the DNA polymerases for further extension[30]. The DNA polymerases DnaE and PolC use deoxyribonucleotides (dNTP's) to replicate the chromosome. DnaE was found to be required for lagging strand synthesis, whereas PolC is required for both[32]. Furthermore, it was found that both PolC and DnaE can extend DNA primers, but only DnaE can extend RNA primers[32]. For this reason, it is assumed that PolC is responsible for continuous synthesis of the leading strand and DnaE is responsible for extending the RNA primers on the lagging strand, before handing it over to PolC[32].

Replication termination is controlled by a polar mechanism, where there is a permissive and non-permissive direction in which the terminus, the region directly opposed to the origin, can be approached. In *B. subtilis*, the binding of two homodimers of the replication termination protein (RTP) at A and B sites is necessary to stop replication[33,34]. Crystal studies have shown that the binding of RTP is different at the A site compared to the B site[35], which may give rise to the establishment of a permissive and non-permissive direction. However, RTP binding to the A site is also required [36].

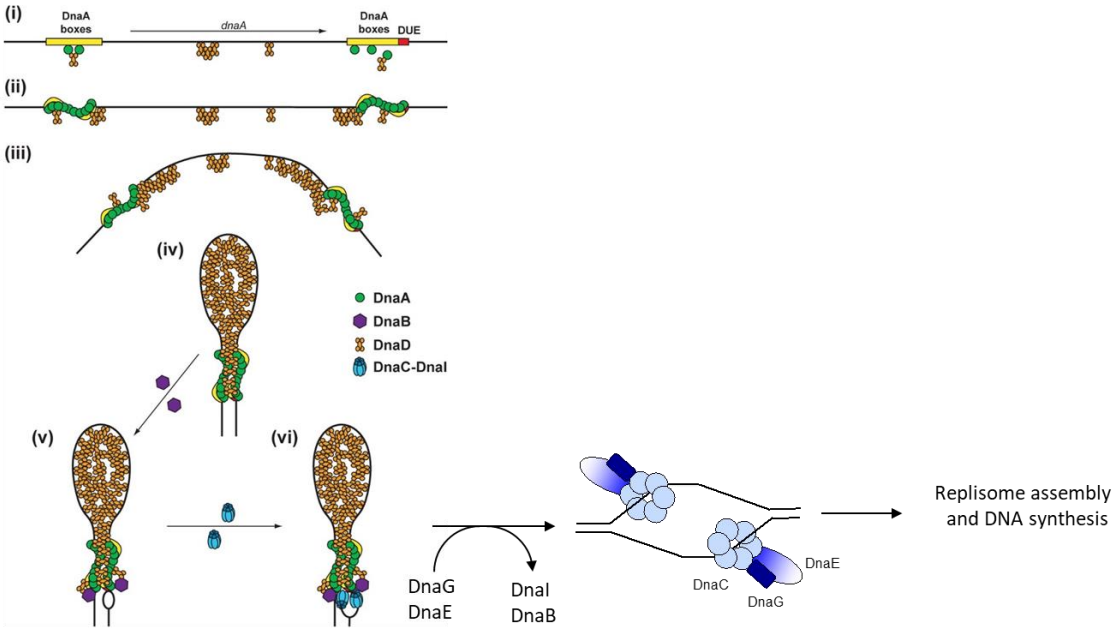
This general mechanism of DNA replication is highly conserved in *E. coli*, but there are a few differences[16]. First, the origin of replication in *B. subtilis* is bipartite i.e. it consists of two intergenic regions, separated by *dnaA*[37]. In contrast, the *E. coli* origin of replication is continuous i.e. consist of only one intergenic region[16]. Second, *E. coli* does not have the *B. subtilis* DNA remodellers DnaB and DnaD[16]. Third, in *E. coli*, the DNA helicase loader is called DnaC and the DNA helicase is called DnaB[16]. Fourth, elongation in *E.coli* is done solely by DNA polymerase III (PolIII) [27–29]as opposed to *B. subtilis* leading strand polymerase PolC and lagging strand polymerase DnaE[32,38]. Fifth, while the principle of termination is maintained in *E. coli*, the termination protein is monomeric and called Tus in *E.coli*[39–41] as opposed to the homodimeric RTP from *B. subtilis*[16]. These small differences highlight the bacterial diversity in the specifics of the replication process as well as the conservation of the general process.

In the next section, the classical mechanisms by which the timing of DNA replication is regulated in the model bacteria *E. coli* and *B. subtilis* are reviewed.

A


Initiation
Elongation
Termination
oriC
Replisome
oriC
oriC

B



C

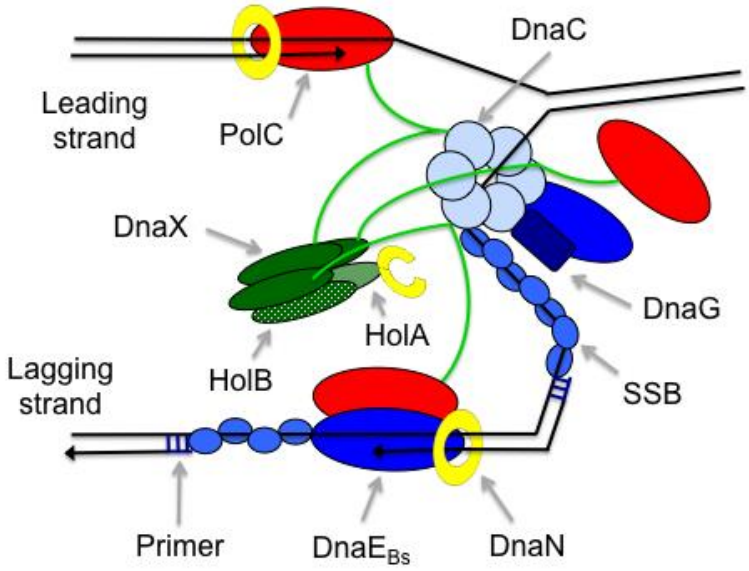


Figure 4 : DNA replication initiation and elongation in the model bacterium B. subtilis:

(A) The three phases of DNA replication (B) DNA replication initiation: DNA replication initiation: When bound to ATP, the initiator protein DnaA binds to DnaA boxes in the origin of replication and oligomerizes. This oligomerization triggers the binding of DnaD, that loops the origin and recruits DnaB. DnaB binding leads to the unwinding of the DNA unwinding element (DUE), which triggers the loading of the DNA helicase loader, the DNA helicase and all the other proteins that collectively form the replication machine (replisome). Taken from [42] (C) DNA replication elongation in B. subtilis. First, the helicase DnaC unwinds dsDNA. ssDNA is immediately coated with single stranded binding proteins (SSB) to prevent reassembly and degradation. Next, the leading strand DNA is handed off to the leading strand polymerase (PolIII vs PolC), which is bound to the DNA by the β-clamp DnaN. The lagging strand is handed off to the primase DnaG, which synthesizes an RNA primer. This RNA

primer is used to extend the lagging strand by DnaE and PolC. The RNA primer is then replaced by DNA (PolI vs DnaE) and ligated to the newly extended strand using DNA ligase. The replisome is held together by DnaX, HolA and HolB.

2.2 Classical mechanisms of DNA replication timing control in *B. subtilis* and *E. coli*

The goal of gating DNA replication to a specific time window is to ensure that replication initiates once and only once per cell cycle. Classical mechanisms for controlling DNA replication timing primarily focus on controlling the initiation step and act on DnaA and *oriC*. In *B. subtilis*, only a handful of proteins are known to regulate timing of DNA replication and all act by regulating the assembly of the DnaA oligomer[16].

YabA is the major regulator of DNA replication timing during vegetative growth[43]. Deletion of *yabA* leads to over initiation phenotypes, growth defects and asynchronous replication[43]. YabA consists of an N-terminal tetramerization domain and a C-terminal domain that can interact with both DnaA and DnaN[44]. Despite a lot of research however, the mechanistic details of how YabA controls DNA replication timing remain elusive. To date, two models have been proposed. In the first model, YabA tethers DnaA to DnaN at the replisome, thereby preventing DnaA to initiate new rounds of replication when the first one is finished[45], consistent with an observed change in DnaA localization inside *ΔyabA* cells[45]. In the second model, YabA binds DnaA and prevents assembly of the DnaA oligomer[46]. This inhibition is relieved by competitive binding with DnaN[46], consistent with the observation that DnaN concentrations are correlated with initiation frequency[46].

Similarly, the segregation protein Soj (ParA) is another regulator that influences timing of DNA replication by modulating the assembly of DnaA[47,48]. Soj is an ATPase and its effect on replication depends on whether it is bound to ADP or ATP[49]. Soj-ADP is monomeric[50] and its interaction with DnaA prevents the assembly of DnaA at the origin. Conversely, Soj-ATP is dimeric and stimulates assembly of the DnaA helix[48].

Additionally, six DnaA box clusters have been found outside the origin region of *B. subtilis* and their deletion causes over initiation[51]. This could mean that these regions titrate DnaA away from the origin. However, the effect was only strong when all six were deleted, casting doubt on how important this mechanism is for controlling the timing of DNA replication[51].

Comparatively, more diverse mechanisms of controlling DNA replication timing have been discovered in *E.coli*[16].

Like YabA and Soj in *B. subtilis*, the replication control protein DiaA influences DNA

replication timing in *E. coli* [52] by affecting the oligomerization of DnaA at the origin of replication[53]. Deletion of *diaA* is associated with replication timing defects[52] and the DiaA tetramer is thought to act as a molecular scaffold that supports DnaA oligomerization at the origin of replication[53].

Furthermore, binding of the nucleoid associated proteins IHF and Fis in the origin region also replication timing[54,55]. IHF binding in the origin region is thought to bring two DnaA boxes closer together and hence stimulate the assembly of the DnaA oligomer[56]. Conversely, Fis binding in the origin region inhibits DNA replication initiation by preventing the binding of both IHF and DnaA[54].

Like Fis, the SeqA protein in *E. coli* affects DNA replication timing by preventing the binding of DnaA at the origin of replication[57–59]. Concretely, SeqA binds hemimethylated GATC sequences in the origin region (and beyond)[60], which prevents both binding of DnaA there as well as the transcription of *dnaA*[57–59].. The purpose of SeqA is to prevent immediate reinitiation after a new round of replication was launched. SeqA inhibition is relieved by methylation of the newly replicated strand using the Dam methyl transferase[61].

In contrast to *B. subtilis* however, oligomerization and binding of DnaA in *E. coli* is actively regulated by controlling the conversion between DnaA-ATP and DnaA-ADP[16]. DnaA in *S. aureus* (a close relative of *B. subtilis*) was found to exchange ATP for ADP 10 times more frequently, making the conversion between DnaA-ATP and DnaA-ADP less viable as a regulatory mechanism in these organisms[62].

The *E.coli* replication control protein Hda, the main *E.coli* regulator of DNA replication timing during vegetative growth, functions in this way. Hda is an ATPase whose effect on replication depends on whether it is bound by a nucleotide. Hda-ADP binds to DnaA and stimulates its ATPase activity, converting it from DnaA-ATP to DnaA-ADP[63]. Conversely, apo-Hda forms oligomers and multimers, which may suggest that Hda oligomerization is important for its activity[63]. Interestingly, Hda also binds DnaN and the replisome and this binding is required for Hda-DnaA binding[63]. This has led to the model that the role of Hda is to prevent immediate reinitiation after a new round of replication by disassembling the DnaA oligomer after the replisome is loaded[64].

DnaA box clusters in *E. coli* outside the origin region not only titrate DnaA, but also control the conversion between DnaA-ATP and DnaA-ADP[16]. To date, three major DnaA box clusters have been discovered in *E.coli*. The first locus *datA* is bound right after initiation by IHF, which is essential for *datA* function[65]. When bound by IHF, *datA* binds DnaA-ATP and stimulates assembly and hydrolysis of ATP there, thereby preventing immediate reinitiation[66].

Conversely, binding of DnaA to the DnaA reactivating sequence (DARS) 1 or 2, stimulates nucleotide exchange and recharges DnaA-ADP with ATP[67]. This process of recharging of DnaA has been shown to be influenced by IHF and Fis binding in the DARS region as well[68]. Recently, DnaA in *E. coli* was also found to be acetylated on a lysine residue that is crucial for nucleotide binding in a growth phase dependent way[69].

A summary of the different strategies to regulate DnaA activity is given in figure 5.

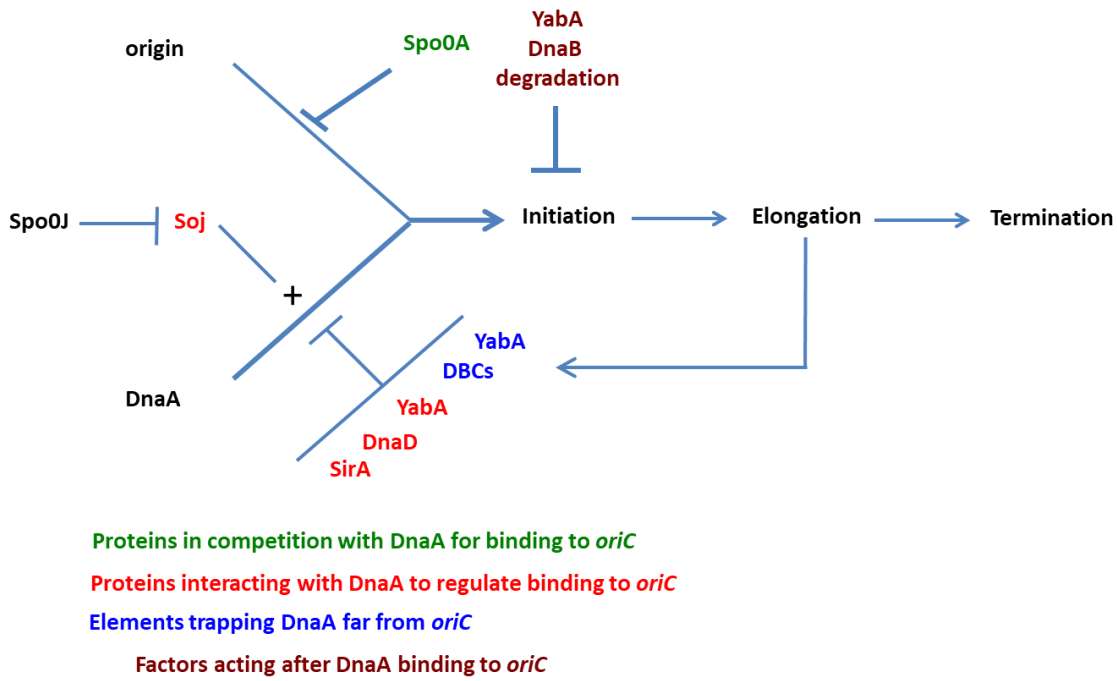


Figure 5: Regulation of DnaA mediated initiation in E.coli and B. subtilis

Regulation of DnaA acts on (i) the nucleotide bound state of DnaA (ATP or ADP) done by the E.coli regulators Hda and the datA/DARS loci and (ii) the binding of DnaA and its assembly into the initiation competent DnaA oligomer, done by the other mechanisms.

The classical mechanisms described in this section assure that replication happens once and only once during each cell cycle during steady state conditions.

However, most bacteria and lower eukaryotes must also be able to adapt their growth rate to varying environmental conditions[70]. These changes in growth rate require coordinated changes in the timing of the cell cycle[70]. Despite decades of research, the mechanisms that enable organisms to adjust their cell cycle dynamics, including DNA replication timing, to changing

nutritional conditions remain poorly understood[70]. In the next section, the coupling of growth rate and DNA replication timing in bacteria is discussed. Next, the possible molecular mechanisms that might underlie this coupling between DNA replication timing and metabolism are critically evaluated in the model organisms *B. subtilis* and *E. coli*.

2.3 Metabolic control of DNA replication: coupling DNA replication timing to nutrient richness and growth rate

2.3.1 Evidence for a metabolic control of DNA replication in bacteria and lower eukaryotes

Early studies in the model bacteria *E. coli* and *B. subtilis* have revealed that the growth rate increases with increased nutrient availability and that this has consequences for DNA replication timing[71]. At low growth rates, the B, C and D period lengths increase with decreasing growth rates[72]. At high growth rates, the C and D periods remain virtually constant with increasing growth rates[71]. Paradoxically, both *E. coli* and *B. subtilis* achieve growth rates that are faster than their C and D periods at a certain threshold of nutrient availability. In order to resolve this issue, these bacteria start multiple replication cycles within one cell, a phenomenon called multifork replication[71,73].

Following these observations, Cooper and Helmstetter elaborated a model in which the timing of replication initiation is under growth rate control and that the timing of all other events in the cell cycle followed at a fixed time interval[71]. This widely influential model successfully explained how bacteria balance largely constant C and D periods at high growth rates with nutrient dependent changes in growth rate[71].

Not much later, the hypothesis that the accumulation of cell mass regulates the timing of DNA replication became popular. Donachie[74] observed that the ratio of the average cell initiation mass to number of origins per cell was constant throughout many different media that enable fast growth in *E. coli*[71] and *Salmonella*[75]. More recently, single cell studies have demonstrated that the accumulation of cell mass since birth, not cell mass itself, before the onset of DNA replication is constant[76,77].

This model of cell mass accumulation controlling DNA replication timing has been criticized[10,78–82]. Briefly, the studies mentioned above correctly noted a correlation between cell mass and DNA replication timing, but incorrectly concluded a causal relationship[10]. Moreover, recent studies in *B. subtilis* and *E. coli* have shown that cell size and DNA replication timing can be uncoupled, suggesting that there is no causal relationship[83,84,85]. For instance, mutants in

B. subtilis that are 35% shorter than wild type still initiate with normal timing[83]. This means that replication can be uncoupled from cell size in *B. subtilis*[83]. Additionally, initiation mass in *E.coli* can vary 1,6 to 2 fold in slow growing cells, despite being constant in cells growing at a normal rate[78,84,85]. Finally, this view of the cell cycle as a single process controlled through DNA replication timing by the accumulation of cell mass doesn't explain why changes in nutrient availability can also affect the subsequent steps in the cell cycle[70] (See 1.4.2).

For these reasons, several groups have started to argue that growth rate dependent control of replication timing is a multifactorial process, which may involve sensing cellular metabolism and communicating this information to the replication machinery[10,70,86–88]. Currently, this view is favored, since accumulating evidence suggests that the necessary molecular factors that govern replication timing may depend on the growth rate[78,80,85].

Additionally, some evidence points to a direct role for metabolic signals to control DNA replication timing. It has been known for many years that yeast cultured in continuous culture condition creates synchronized oscillations of oxygen consumption[89] . These oscillations are called the yeast metabolic cycle (YMC) and are accompanied with oscillations in gene expression of over half of the yeast genome[90]. Interestingly, the proteins necessary for DNA replication (polymerases, primases etc) reach peak expression at the same time during the reductive phase of the YMC and DNA replication was also restricted to this phase[90]. This temporal compartmentalization of DNA replication to the reductive phase within the YMC is thought to protect DNA from damage caused by oxidative agents produced during the oxidative phase[90–92]. This example highlights both the importance of proper timing of DNA replication within the cell cycle as well as the fine-tuned way this timing responds to the metabolic state of the cell.

In the next section, the known mechanisms coupling DNA replication timing to growth rate are discussed.

2.3.2 Mechanisms of metabolic control of replication

2.3.2.1 Metabolic control of initiation: classical explanations

The concentration of DnaA has been proposed to couple DNA replication initiation timing to growth rate. First, the concentration of DnaA correlates well with the growth rate in bacteria during rapid growth[93]. Conversely, cells starved from amino acids stop expressing DnaA and don't initiate new rounds of replication[94]. Second, the concentration of DnaA has been shown to directly impact replication initiation timing. Artificial overexpression of DnaA leads to overinitiation[95–97]. By contrast, artificial inhibition of DnaA leads to underinitiation

phenotypes[98]. Third, DnaA expression levels *in vivo* have been demonstrated to be under metabolic control. Specifically, DnaA expression is under the control of ppGpp[88]. ppGpp is a small nucleotide that can inhibit the transcription of many genes by interfering with the transcription complex [99,100]. DnaA is among the genes whose expression is affected by ppGpp, although it is currently unknown if this effect is direct or indirect[101]. Interestingly, ppGpp synthesis is stimulated upon amino acid or carbon starvation and therefore couples DnaA expression levels to nutrient availability[100].

Finally, DNA gyrase has been shown to influence DnaA dependent replication initiation in *B. subtilis* by inhibiting DnaA binding[102]. The sensitivity to DNA topology was partly mediated by YabA[102]. Interestingly, the activity of DNA gyrase has long been known to depend on the metabolic state of the cell[103] allowing it to potentially couple DNA replication timing to metabolism. These classical mechanisms coupling DNA replication to metabolism are shown in figure 6.

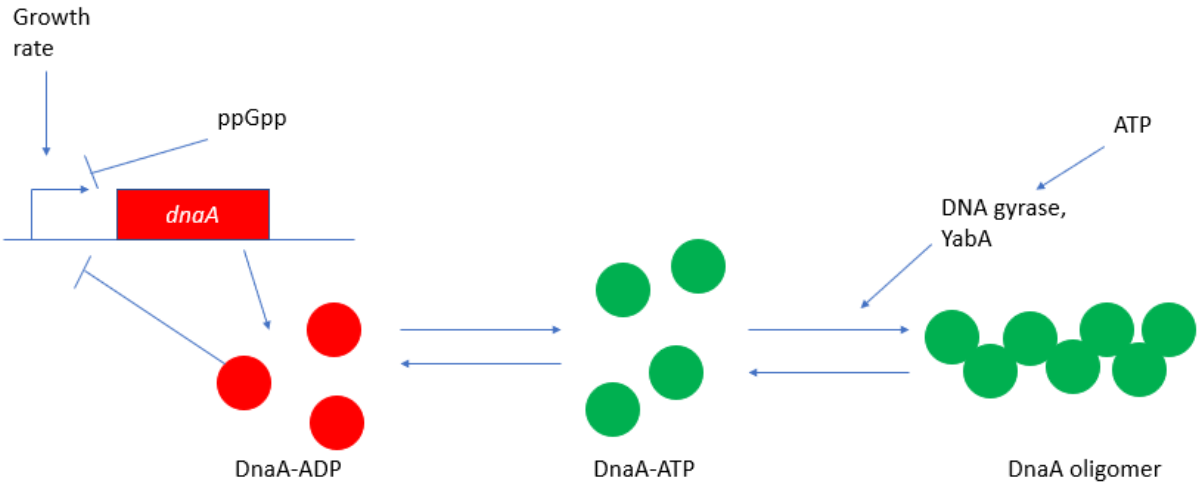


Figure 6: Classical mechanisms of metabolic regulation of DNA replication initiation

Metabolic regulation of initiation acts primarily on DnaA. DnaA expression increases with increasing growth rates, yielding higher initiation frequencies. Under AA starvation conditions, ppGpp prevents expression of DnaA and inhibits new rounds of replication. Additionally, DNA gyrase activity has recently been shown to impact DNA initiation frequency in a DnaA dependent manner and DNA gyrase activity has long been known to be influenced by nutrition.

2.3.2.2 Metabolic control of elongation: classical mechanisms

DNA replication elongation is known to be subjected to metabolic control as well.

The most obvious mechanism by which elongation rate is controlled by nutrition is through

the availability of deoxynucleotide (dNTP) precursors[70]. The replication rate in bacteria that require exogeneous addition of the nucleotide thymine varies with the concentration of thymine [104]. Additionally, the expression of ribonucleotide reductase, an enzyme required for the production of nucleotides, is regulated by DnaA to adapt the production of nucleotides to the requirements for elongation[105]. This suggest that the levels of dNTP influence replication rates under limiting conditions. Interestingly, the expression of *nrdAB* (main gene for RNR) is also under the control of the transcription factor NrdR[106], whose activity is itself controlled by the binding of ATP[107].

In addition to controlling DNA replication initiation rate, ppGpp has been shown to control DNA replication elongation rate under starvation conditions. ppGpp inhibits primase activity in both *E. coli* and *B. subtilis* [108]. Curiously enough though, halting DNA replication through inhibition of primase activity does not induce the SOS response[109]. The classical mechanisms involved in coupling DNA elongation rates to metabolism are summarized in figure 7.

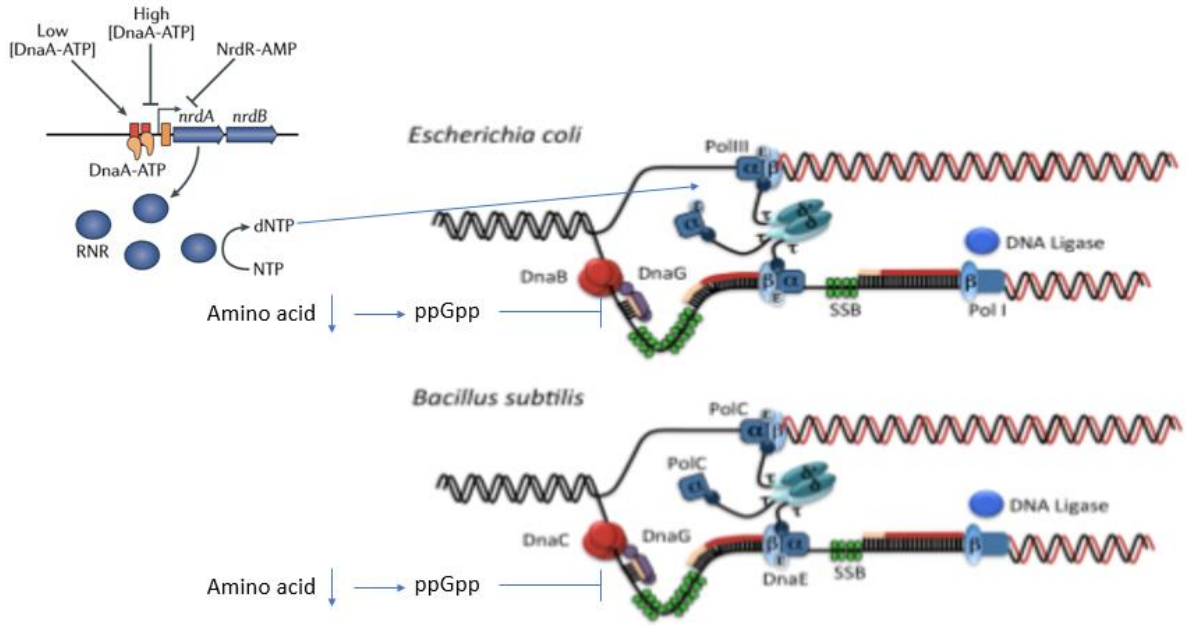


Figure 7 : Classical mechanisms of metabolic regulation of DNA replication elongation: Ribonucleotide reductase, an enzyme that is crucial for the production of nucleotides in E. coli, is under metabolic control of both DnaA-ATP and NdrR-ATP. This ensures the timely production of nucleotides when there is enough energy and when there is need for them. Nucleotides are rate limiting for the replication machinery. Additionally, ppGpp can block primase activity and stall replication fork on top of its effects on initiation Taken from [8,16]

2.3.2.3 Beyond classical mechanisms of metabolic control of replication: a role for direct metabolic signaling

However neither classical regulators of replication timing nor the known metabolic regulatory pathways for control of DNA replication suffice seem to fully explain coupling. In both *E.*

coli and *B. subtilis*, the notion that the concentration of DnaA-ATP is a limiting factor for proper timing of DNA replication during fast growth conditions has been challenged[110,111]. While the effect of ppGpp on replication at high concentrations during nutritional stress is well established, its role during exponential growth remains up for debate[110,112]. Deletion of the key classical regulators of DNA replication timing *yabA* and *soj* in *B. subtilis* impacted replication frequency, but not the coupling between DNA replication and metabolism[110]. In short, few mechanisms have been proposed that satisfactorily couple DNA replication timing to metabolism, especially under fast growth conditions in bacteria.

Recent studies suggest that key cellular metabolites and metabolic enzymes might directly send signals that regulate the cell cycle, even under fast growth conditions.

In *E. coli*, acidic phospholipids like phosphotidylglycerol (PG) and cardiolipin (CL) have been shown to regulate DNA replication initiation reviewed in [113]. *In vitro* studies have shown that CL and PG can stimulate the exchange of ADP for ATP in the *E. coli* DnaA[114–116]. The speed of this release depends strongly on the fluidity of the membrane[115,116]. Acidic phospholipids have also been known to inhibit binding of DnaA to *oriC*[117]. Interestingly, the presence of phospholipids in the membrane has been directly connected to cellular growth[118,119], making this particular signaling system well suited for coupling DNA replication timing to growth rate.

During the yeast metabolic cycle (YMC), acetyl-CoA levels, which is a crucial intermediate in cellular metabolism (see below), were found to respond to the addition of rich carbon sources. Increased acetyl-CoA levels caused increased acetylation levels of histones that were specifically associated with genes important for growth through the activity of an acetyltransferase Gcn5[120]. This also led to increased entry in the cell cycle[120]. The authors also pointed out that under rich growth conditions, acetyl-CoA levels are constitutively high, which may be the reason why this type of metabolic control of the cell cycle has not been observed before.

Other recent studies have suggested that the metabolic enzyme GADPH is involved in regulating viral replication in yeast and plants[121,122]. Specifically, GADPH was found to stimulate plus strand synthesis of the RNA virus *Tombusvirus* both using *in vitro* assays and cell extracts[121]. Moreover, a negative mutant of GADPH inhibited viral replication *in vitro* and viral build-up *in vivo*[121]. GADPH was also found to bind the RNA virus *Bamboo Mosaic Virus* and negatively regulate its replication *in vivo*[122].

These examples illustrate that key metabolites and key metabolic enzymes might signal metabolic cues that regulate the cell cycle and DNA replication timing, even in fast growth

conditions and in a diverse array of organisms.Incidentally, both acetyl-CoA and GADPH belong to a highly conserved set of reactions called central carbon metabolism (CCM), which lies at the hearth of both energy and building block production.

The main hypothesis pursued in this work is that key metabolites and/or enzymes in CCM couple DNA replication timing to the metabolic state of the cell in bacteria. In the next chapter, the metabolic reactions that constitute CCM are reviewed with an emphasis on their importance for cellular energy and building block production. Next, the regulation of CCM activity in *B. subtilis* is discussed. Finally, an overview is presented of the accumulating evidence that CCM is committed to replication control in a wide range of organisms.

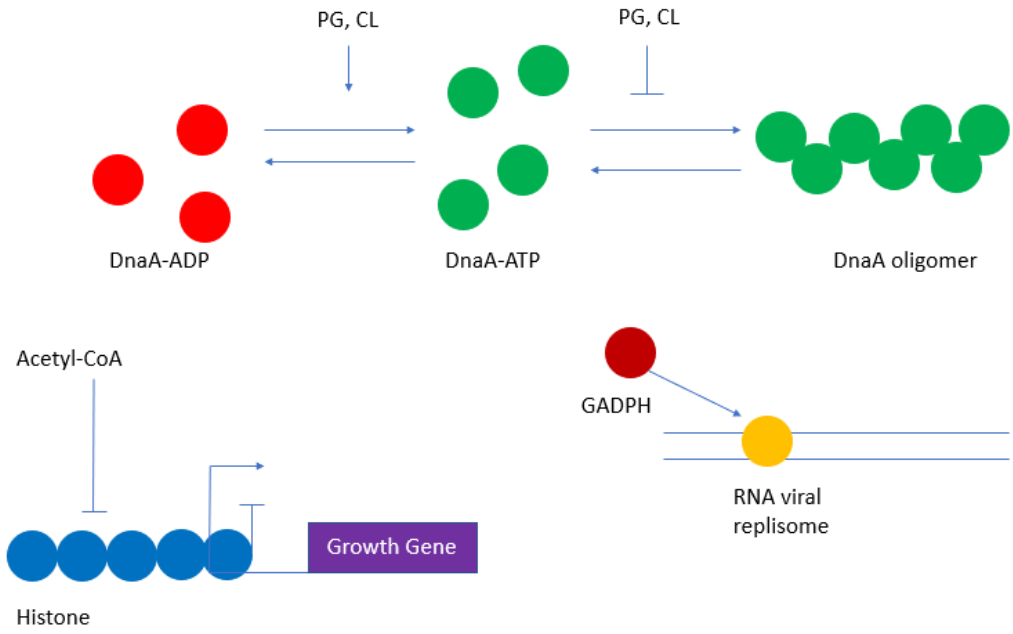


Figure 8 : Recent examples of direct metabolic signalling that impacts replication and growth.

(Top) Acidic phospholipids like phosphatidylglycerol (PG) or cardiolipin (CL) are increasingly recognized to have a role in DNA replication initiation in E. coli. PG and CL regulate replication by (i) stimulating nucleotide exchange and recharging DnaA and (ii) sequestering DnaA away from oriC. The presence of these lipids in the membrane is associated with growth. (Bottom Left) During the Yeast Metabolic Cycle, acetyl-CoA levels have been associated with increases acetylation of histones associated with genes important for growth. Histone acetylation causes the histones to dissociate and expression of growth genes is no longer repressed. (Bottom right) The metabolic enzyme GADPH has been found to influence DNA replication speed of several plant viruses through direct interactions.

3 Central carbon metabolism (CCM) and its commitment in DNA replication

3.1 Central carbon metabolism (CCM): brief introduction

3.1.1 CCM in eukaryotes

Central carbon metabolism (CCM) is a collection of reactions aimed at extracting energy and precursors for macromolecular synthesis from nutrients[86]. It consists of about 30 key reactions that function by oxidizing glucose[123] and are highly conserved throughout all phyla[86] .

Central carbon metabolism in higher eukaryotes consists of four major pathways: glycolysis, the Krebs cycle, the pentose phosphate pathway and overflow metabolism and is reviewed in [124].

Glycolysis comprises the set of reactions that breakdown glucose into pyruvate. In human cells, glycolysis is performed by the Embden-Meyerhof-Parnas pathway (EMP) pathway[124]. The EMP pathway is shown in figure 6. The EMP pathway net produces 2 ATP and 2 NADH in terms of energy and reductive power[124]. Additionally, the EMP pathway also produces the intermediates required for the Krebs cycle (pyruvate) and the pentose phosphate pathway (glucose-6-phosphate, fructose-6-phosphate and glycerldehyde-3-phosphate) (See below). On top of that, the EMP pathway fulfills important anabolic needs by providing key intermediates for non-essential amino acid production in humans[124]. Specifically, key intermediates of the EMP pathway include 3-phosphoglycerate (serine, glycine, cysteine) and pyruvate (alanine)[124]. The EMP pathway is shown in figure 9.

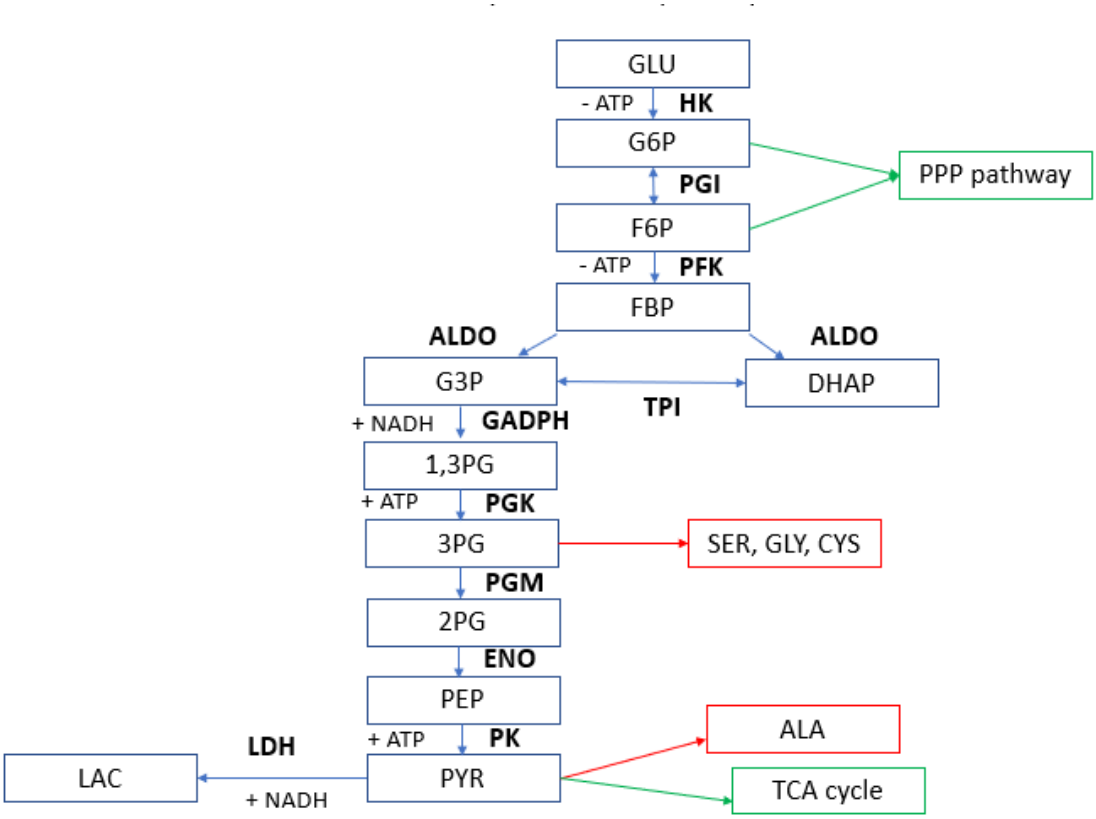


Figure 9 : The Embden-Mayerhof-Parnas pathway for glycolysis + the overflow pathway

Glossary: (Molecules) GLU: glucose, G6P: Glucose-6-phosphate, F6P: Fructose-6-phosphate, FBP: Fructose-1,6-bisphosphate, DHAP: Dihidroxyacetonephosphate, G3P: Glyceraldehyde-3-phosphate, 1,3PG: 1,3 bisphosphoglycerate, 3PG: 3-phosphoglycerate, 2PG: 2-phosphoglycerate, PEP: phosphoenolpyruvate, PYR: pyruvate, LAC: lactate, ALA: alanine, SER: Serine, GLY: glycine, CYS: cysteine (Enzymes): HK:hexokinase, PGI: phosphoglucose isomerase, PFK: phosphofructokinase, ALDO: aldolase, TPI: triosephosphate isomerase, GADPH: glyceraldehyde phosphate dehydrogenase, PGK: phosphoglycerate kinase, PGM: phosphoglycerate mutase, ENO: enolase, PK: pyruvate kinase, LDH: lactate dehydrogenase

The Krebs cycle (or Tricarboxic Acid or TCA) comprises the set of reactions that are responsible for the oxidative decarboxylation of pyruvate, resulting in the formation of 38 ATP, mostly from NADH and $FADH_2$ and 6 CO_2 starting from glucose. [124]. The Krebs cycle reactions are executed in the mitochondrion in humans[124]. Importantly, the Krebs cycle considerably extends the anabolic capabilities of the EMP pathway[124]. Specifically, it produces the key intermediates oxaloacetate (aspartate and asparagine) and α-ketoglutarate (glutamate, proline, glutamine and arginine) for non-essential amino acid production. The Krebs cycle also produces acetyl-CoA, the fundamental building block for lipid biosynthesis[124]. The Krebs cycle is shown in figure 10.

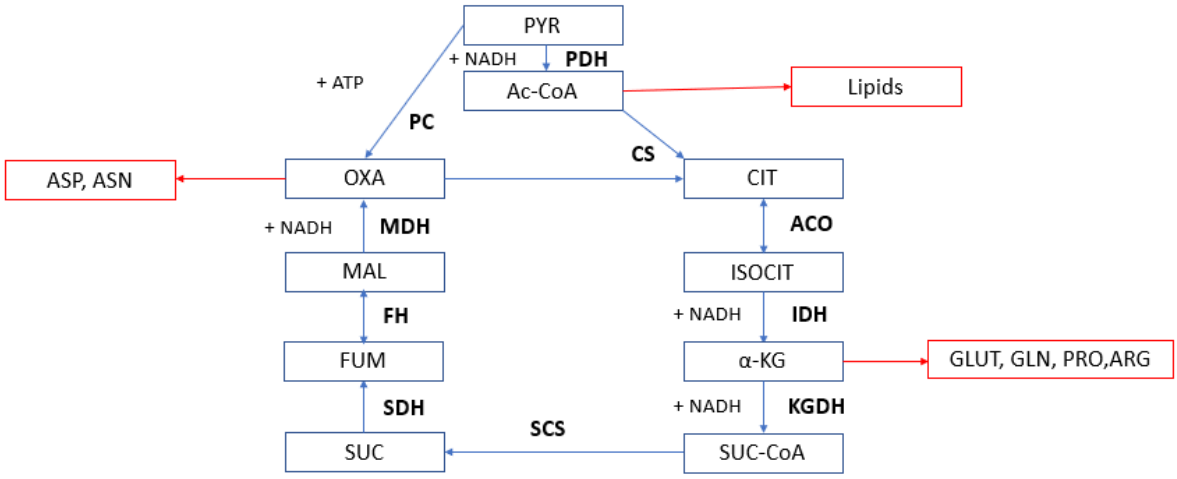


Figure 10 : The Tricarboxyic acid (TCA) or Krebs cycle

Glossary : (Molecules) PYR: pyruvate, LAC: lactate, ALA: alanine, SER: Serine, GLY: glycine, CYS: cysteine, Ac-CoA: Acetyl-CoA, CIT: citrate, ISOCIT: isocitrate, α-KG: α-ketoglutarate, SUC-CoA: succinyl-CoA, SUC: succinate, FUM: fumarate, MAL: malate, OX: oxaloacetate, (Enzymes) PDH: pyruvate dehydrogenase, CS: citrate synthase, ACO: aconitase, IDH: isocitrate dehydrogenase, KGDH: ketoglutarate dehydrogenase, SCS: succinyl-CoA synthetase, SDH: succinate dehydrogenase, FH: fumarase, MDH: malate dehydrogenase, PC: pyruvate carboxylase

The pentose phosphate pathway is a metabolic shunt that starts with glucose-6-phosphate and ends with the EMP intermediate 3-glyceraldehyde-phosphate and fructose-6-phosphate [124]. Its most important role is in providing the cell with the key intermediate ribose-5-phosphate, which is the precursor for nucleotide biosynthesis and NADPH, which provides reductive power to drive biochemical reactions[124]. The pentose phosphate pathway is shown in figure 11.

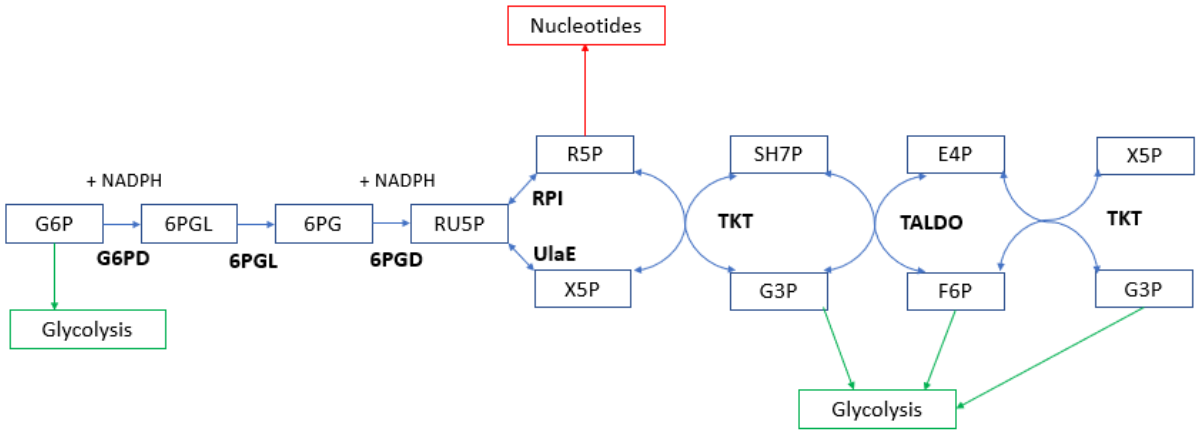


Figure 11: The pentose phosphate pathway

Glossary : (Molecules) 6PGL: 6-phosphogluconolactone, 6PG: 6-phosphogluconate, RU5P:

ribulose-5-phosphate, R5P: ribose-5-phosphate, X5P: xylose-5-phosphate, SH7P: sedoheptulose-7-phosphate, E4P: erythrose-4-phosphate (Enzymes) GP6D: glucose-6-phosphate dehydrogenase, 6PGL: 6-phosphogluconolactonase, 6PGD: 6-phosphogluconate dehydrogenase, RPI: ribose-5-phosphate isomerase, UlaE: Ribose-5-phosphate-3-epimerase, TKT: transketolase, TALDO: transaldolase

The overflow pathway is an alternative route for pyruvate oxidation[124]. Instead of being completely oxidized, pyruvate is converted into lactate, thereby regenerating 1 NADH per pyruvate molecule[124]. Despite being less efficient in terms of energy production and production of intermediates, overflow pathways regenerate NAD^+ much faster, allowing for a faster flux through glycolysis[124]. For this reason, overflow pathways are often active in highly proliferative cells such as tumors[124].

Despite this top-down description, CCM can be fed at multiple points and most cells can use several organic molecules as carbon source for growth. Sugar molecules like glucose or fructose enter the EMP pathway at the top and are broken down throughout the EMP and TCA pathways. This top-down flux regimen is called a glycolytic growth condition. In human cells, glutamine, which enters the TCA cycle via oxaloacetate, can support growth as well[124]. In this case, the flux through the CCM is reversed

3.1.2 CCM in bacteria and archaea

Compared to eukaryotes, central carbon metabolism in bacteria is more diverse. In bacteria, breakdown of glucose into pyruvate is predominantly, but not exclusively achieved using the EMP pathway. The model bacteria *B. subtilis*, *E.coli* and many others use the EMP pathway as their major pathway for glycolysis[123] and most bacteria can make all amino acids using CCM intermediates[123]. However, a widespread alternative to the EMP pathway is the Entner-Doudoroff pathway (ED), which is shown in figure 7. The ED pathway converts glucose-6-phosphate into pyruvate and glyceraldehyde-3-phosphate, which is turned into pyruvate using the same final reactions as EMP[123]. The ED pathway is the main glycolytic pathway in *Pseudomonas*, *Zymomonas* and *Azotobacter* and is a minor pathway in *E. coli*, that enables it to use gluconate as a carbon source[123]. The ED pathway has a net yield of 1 ATP and 2 NADH[123]. The ED pathway is shown in figure 12.

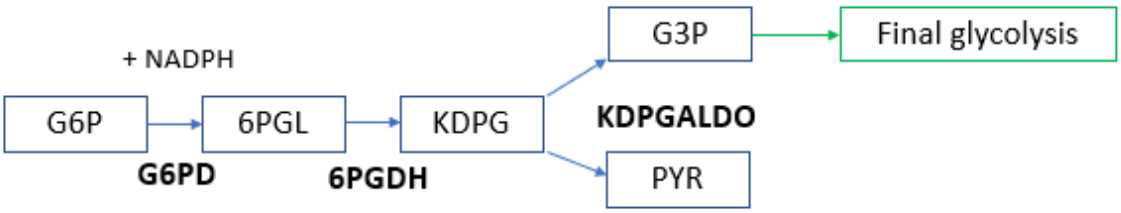


Figure 12: The Entner-Douderoff pathway

Glossary : KDPG: 2-keto-3-deoxy-phosphogluconate, 6PGDH: 6-phosphogluconate dehydrogenase, KDPGALDO: 2-keto-3-deoxy-6-phosphogluconate aldolase

The TCA cycle and PPP pathway are universally present in some form in all bacteria due to their importance for producing building blocks[123]. Full complements of both are present in the model bacteria *B. subtilis* and *E. coli*, but this isn't the case for all bacteria. The PPP pathway becomes the major pathway of energy production in bacteria that utilize pentoses and don't have the alternative EMP or ED pathways[123]. In some bacteria lacking a full TCA cycle, PPP pathway also helps fully oxidize pyruvate[123].

Archaea have similar metabolic complexity to their bacterial counterparts, but demonstrate even greater variability in their CCM[125]. Only modified versions of EMP and ED pathways have been identified so far in *Archeae*[125]. Pentose sugars are broken down differently from their bacterial counterparts and PPP pathway is not fully present if at all[125]. Some key CCM enzymes/reactions in bacteria and eukaryotes have been replaced in *Archeae*[125].

Regardless of the form it takes in any particular organism, CCM as a set of reactions that serve both catabolic and anabolic cellular needs, is universally conserved. Given its importance for both energy production and biosynthesis of important precursors of macromolecular synthesis, CCM is uniquely well positioned to sense both supply and demand of cellular energy and building blocks[86]. All organisms, including the model organism *B. subtilis* expend a lot of energy to produce CCM enzymes in high copy numbers[126] and some studies suggest the trade-off between enzyme cost and energy efficiency is the reason for the prevalence of both EMP and ED[127]. It is therefore important to tightly regulate the flux through these pathways. The mechanisms responsible for controlling CCM activity in *B. subtilis*, the object of this study, are discussed in the next section.

3.2 Regulation of CCM activity: the PTS system and carbon catabolite repression

3.2.1 The phosphoenolpyruvate:carbohydrate phosphotransferase (PTS) system

In a mixture of carbon sources, *B. subtilis* preferentially uses the carbon source that most readily supports growth before moving onto others, creating a hierarchy of carbon sources. To make this sophisticated phenotype possible, *B. subtilis* has developed a very complex regulatory network capable of controlling CCM activity. For good reviews on this topic, the reader is referred to [128,129].

In bacteria, the phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS) is a key player in regulating CCM activity. First and foremost, the PTS system is responsible for the uptake of a large variety of carbohydrates, including sugars and sugar derivatives such as sugar alcohols, amino sugars, disaccharides and many others[129]. Equally important though are the many regulatory functions the components of the PTS system have in the regulation of CCM activity[129].

The PTS system consists of four soluble and one membrane spanning protein that end up transporting PTS carbohydrates and phosphorylating them upon entry. The PTS system uses phosphoenolpyruvate (PEP) as both an energy source and a phosphoryl donor[129]. At the first step, PEP is bound by the protein EI, the first component of the PTS system[130]. Next, EI is phosphorylated on a conserved histidine by PEP[130], which will subsequently transfer this phosphorylation to histidine 15 of the protein HPr[131]. This protein subsequently transfers the phosphate to the sugar specific PTS components EIIA on a conserved histidine residue[132,133]. EIIA then transfers the phosphate onto EIIB on a conserved cysteine residue[134]. EIIC is the membrane spanning part of the PTS system that holds the carbohydrate, which will ultimately receive the phosphate from PEP and this phosphorylation event lowers the affinity of EIIC for the carbohydrate[135].

The different parts of the PTS system form a phosphorylation cascade that transfers a phosphate from PEP to the incoming carbohydrate[129]. This cascade is shown in figure 13. The phosphorylation status of the proteins forming this cascade depends on two things. First, the uptake of more efficiently metabolized sugars leads to increased dephosphorylation of the PTS components[129]. Second, the PEP to pyruvate ratio also affects the phosphorylation state[136,137]. A high PEP to pyruvate ratio is indicative of starving conditions or growth on non-PTS sugars[137,138], whereas a low PEP to pyruvate ratio is indicative of metabolically active cells[138].

These phosphorylation signals can then be used to regulate CCM activity. It controls catabolic gene expression through (i) carbon catabolite repression (CCR), (ii) operon-specific inducer exclusion and (iii) induction prevention. These are discussed in the next section.

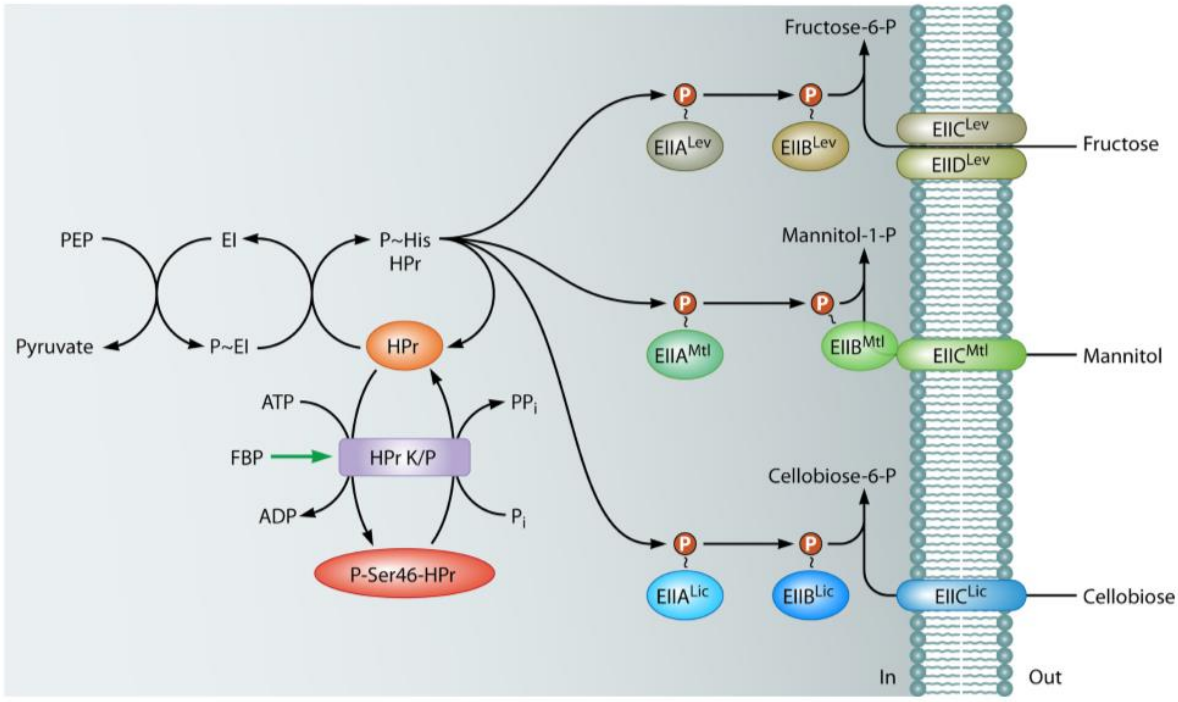


Figure 13: The PTS system in B. subtilis.

Shown here are the PTS systems for fructose ($EIIX^{Lev}$), mannitol ($EIIX^{Mtl}$)and cellobiose ($EIIX^{Lic}$). Taken from [129]

3.2.2 Regulation of central carbon metabolism: Carbon catabolite control

The most important mechanism for regulating CCM activity is carbon catabolite repression (CCR)[128]. This regulatory program represses gene expression of any gene involved in the catabolism of secondary carbon sources, when glucose is present. In *B. subtilis*, the key players of CCR are a global transcription factor CcpA, the HPr protein from the PTS system, the bifunctional HPr kinase/phosphorylase (HPrK) and the glycolytic intermediates glucose-6-phosphate and fructose-1,6- bisphosphate [139,140]. Carbon catabolite repression in *B. subtilis* critically depends on the phosphorylation state of HPr. HPr can be both phosphorylated and dephosphorylated at Ser46 by HPrK[141]. ATP and fructose-1,6-bisphosphate are both indicators of high glycolytic activity and stimulate HPr phosphorylation by HPrK [142–144]. By contrast, inorganic phosphate, an indicator of poor glycolytic activity, stimulates HPr dephosphorylation by HPrK[143,145]. When phosphorylated at Ser46, HPr-P binds to CcpA[146,147]. This binding is stimulated by the glycolytic indicators glucose-6-phosphate and fructose-1,6-bisphosphate[148,149]. The HPr-P-CcpA complex then binds to *cre* sites and represses the

transcription of catabolic genes of carbon sources secondary to glucose[146,147]. PTS mediated carbon catabolite repression is summarized in figure 14.

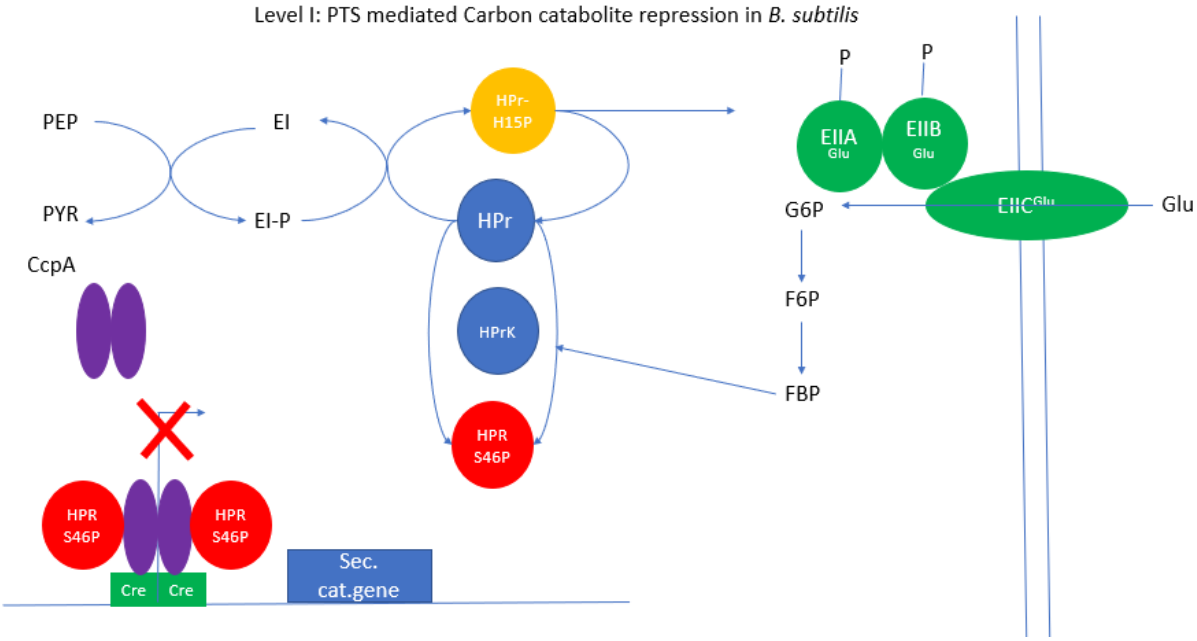


Figure 14: Carbon catabolite repression

In the presence of glucose, a phosphorylation cascade from PEP through the PTS system results in the import of glucose-6-phoaphate, which is transformed into fructose-1,6-bisphosphate (FBP). FBP stimulates the kinase activity of HprK, resulting in the phosphorylation of HPr on serine 46. HPr-S46-P then interacts with the global regulator CcpA, that binds to cre sites and prevents transcription of any gene involved in the catabolism of secondary carbon sources. Based on [129].

3.2.3 Regulation of central carbon metabolism: operon specific control

In addition to CCR regulation, the expression of catabolic genes for many secondary carbon sources is also controlled by an operon-specific transcription factor that is induced when it detects the secondary carbon source[128]. In order to prevent activation of these genes in a mixture of carbon sources, CCR prevents the induction of the transcription factor, either by excluding the secondary carbon source (inducer exclusion) or by preventing induction of the transcription factor (induction prevention)[128]. Both mechanisms have been described for Gram positive bacteria like *B. subtilis*.

Inducer exclusion in Gram positive bacteria is achieved by inactivating the transporter proteins of secondary carbon sources in the presence of glucose using phosphorylation dependent interactions with PTS components. In *Lactobacillus brevis*, HPr-Ser-P can bind and inactivate the galactose permease[150]. In *S. thermophilus*, HPr-His-P can activate the

lactose permease[151]. Phosphorylation on the Ser of HPr prevents this activation[152]. Similarly, in *Enterobacteriae*, the PTS glucose specific component EIIA^{Gluc} is known to interact with and inactivate lactose permease, glycerol kinase, the maltose, melibiose, galactose and arabinose transporters when it is phosphorylated[129].

Induction prevention can be achieved in a variety of ways.

Many secondary carbon source specific transcription factors have duplicated PTS regulatory domains (PRD's) that can be phosphorylated by at least two PTS components[129]. These domains enable hierarchical usage of secondary carbon sources, as the activity of the transcription factors depends both on the availability of the carbon source and on the absence of glucose[153]. In *B. subtilis*, the LevR transcription factor requires phosphorylation by HPr to achieve full activity[154]. Another well-studied example is the LicT antiterminator[155–158].

Alternatively, specific transcription factors can be inactivated by sequestration near the membrane. This is achieved through the fusion of EIIB to its cognate sugar specific transporter and interaction of the cognate transcription factor with the unphosphorylated EIIB. An example of this regulatory mechanism includes the MtlR from *B. subtilis*[159]. Conversely, CCM activity can also be regulated by interactions with phosphorylated PTS components[129]. The HPr-CcpA system described above is one example. Other examples include the stimulation of adenylyl cyclase through interaction with phosphorylated EIIA^{Gluc} [160]and the antiterminator LicT in *B. subtilis* through interaction with phosphorylated EIIB for β-glucoside[161]. .

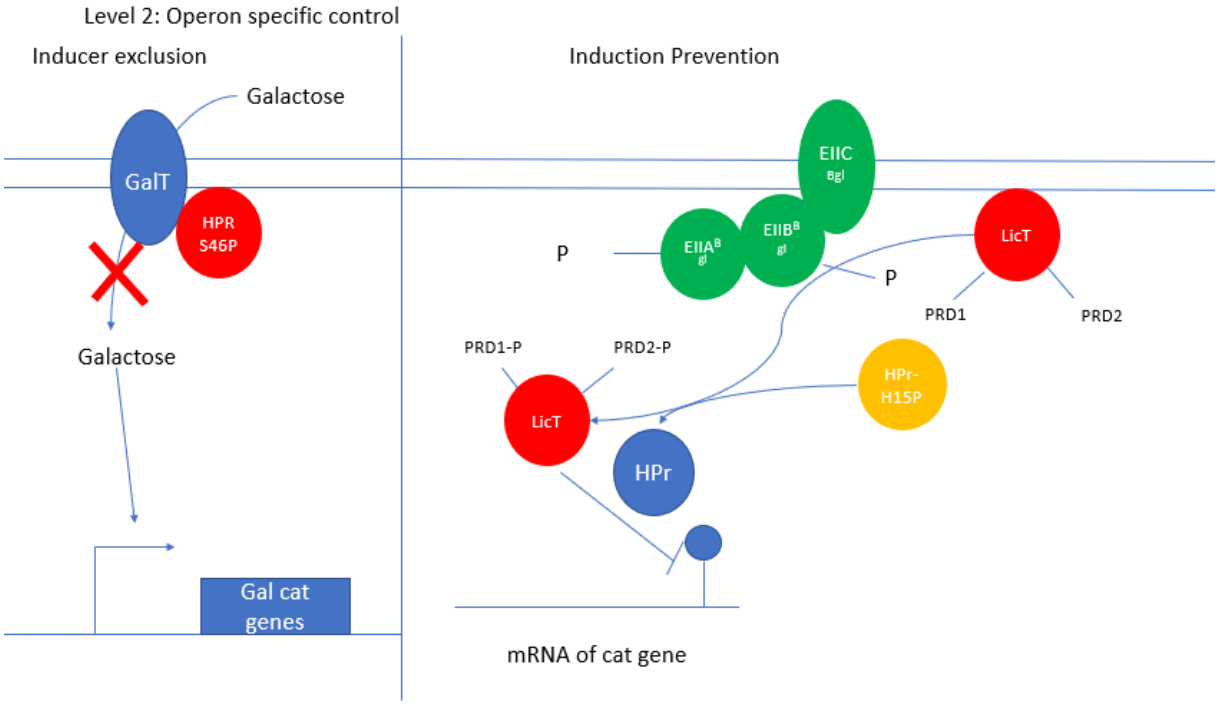


Figure 15: Local control of gene expression by the PTS system

PTS mediated signalling also influences gene expression more locally, enabling hierarchical usage of carbon sources. (Left) Inducer exclusion: Interaction of HPr-S46-P (a sign of the presence of glucose) with the galactose transporter (GalT) prevents the import of galactose in the presence of glucose. This prevents gene expression of galactose specific genes. (Induction prevention) Induction prevention can be achieved through (i) direct phosphorylation dependent interactions with PTS components (ii) phosphorylation of secondary carbon source specific regulators and (iii) sequestration of the regulators at the membrane. Shown is the example of LicT. LicT has two PTS regulatory domain (PRD), which are phosphorylated using EIIB and HPr. When phosphorylated, LicT has full antiterminator activity. When not phosphorylated, it is sequestered near the membrane.

Collectively, these different mechanisms are responsible for optimizing CCM activity under different growth conditions at all times based on the measurement of a few key metabolites. This principle is summarized in figure 16. Thanks to this and CCM's unique position in cellular metabolism (see section 2.1), the activity of CCM perfectly encapsulates the metabolic state of the cell. This makes sensing CCM activity theoretically ideal to couple DNA replication timing to the metabolic state of the cell. In the next section, the evidence will be reviewed that CCM is indeed involved in gating DNA replication timing

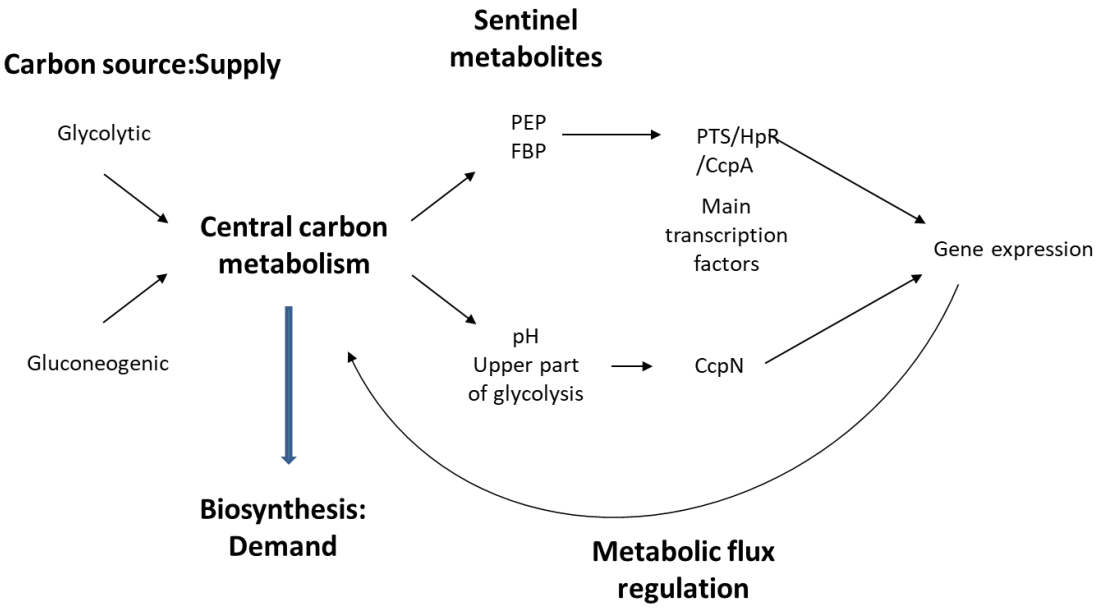


Figure 16: CCM perfectly reflects the metabolic state of the cell : CCM senses both supply and demand. Furthermore, the activity of CCM is dynamically modulated by a complex regulatory system that involves sensing key CCM metabolites such as PEP and regulating metabolic gene expression, thus ensuring optimal flux through CCM in every metabolic condition.

3.3 Commitment of central carbon metabolism in replication

3.3.1 Functional connections between CCM enzymes and DNA replication

In recent years, evidence has been building that CCM is involved in the control of replication timing in prokaryotes and eukaryotes alike. This evidence is summarized in figure 17.

In 2007, *Janniere et al* found that mutations in the four terminal genes of CCM (*pgk, pgm, eno* and *pykA*) could suppress the effects of temperature sensitive (Ts) mutations in the replication proteins DnaG, DnaE and DnaC on survival at elevated temperatures in *B. subtilis.* [162] This restoration of viability to Ts mutants in replication enzymes occurred over a narrow range of concentrations of the CCM enzymes and did not appear to be due to defects in growth rate or an activation of major stress responses[162]. Rather, the restoration phenotype depended at least in part on maintenance of high activity of temperature sensitive mutants at elevated temperatures[162]. The authors therefore proposed that this suppression is evidence of a genetic system connecting CCM to the gating of DNA replication[162]. In follow-up studies, it was shown that multiple CCM deletions in *B. subtilis* can also affect replication initiation (*pykA* and *pgk*) and elongation (*gapA* and *ackA*) directly [86,110]. Both papers suggested the existence of multiple medium dependent links between CCM and replication [86,110].

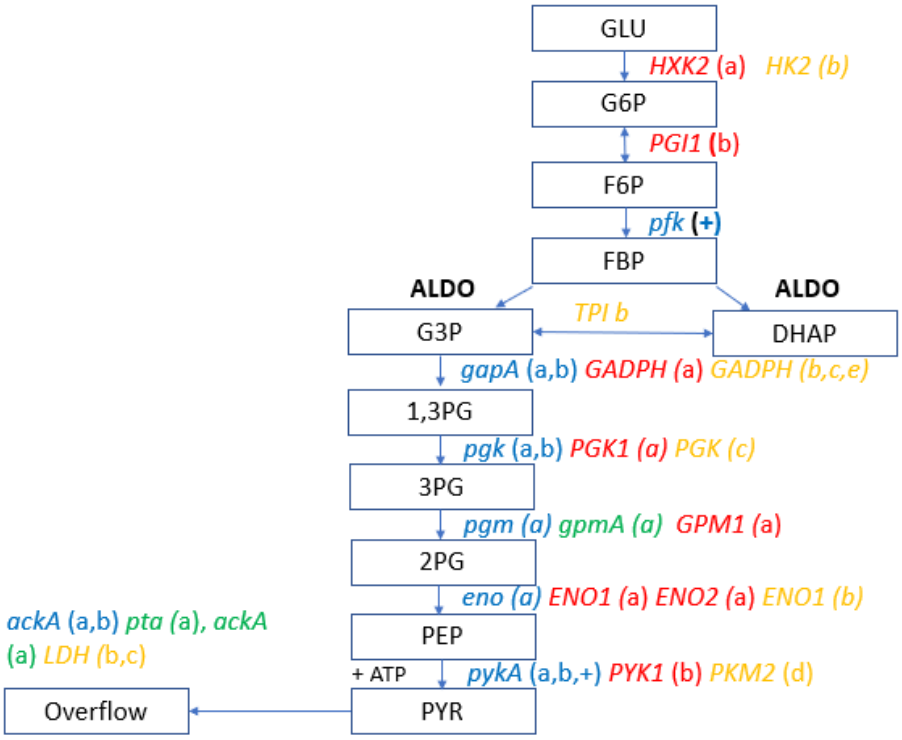
Similar experiments in *E.coli* also revealed that several CCM (*pta, ackA, tkt* and *gpmA*) genes are functionally connected to DNA replication proteins as CCM deletions could both restore

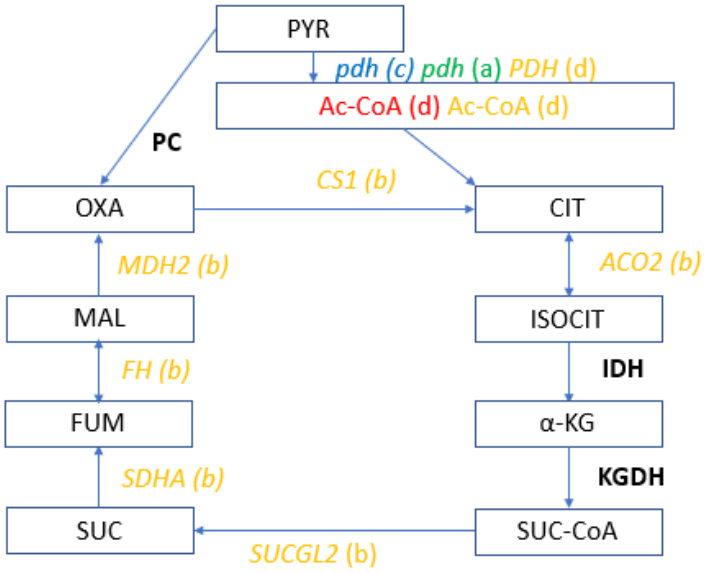
viability at elevated temperatures[163] and could reverse cell filamentation and improper nucleoid positioning in Ts sensitive backgrounds at elevated temperatures[164].

Bioinformatic analysis in prokaryotes supports a role for CCM in control of replication timing as well. *In silico* analysis of fully sequenced bacterial genomes found a highly conserved physical linkage between two glycolytic enzymes (pyruvate kinase and phosphofructokinase) and the DNA polymerase DnaE[165]. This conserved proximity on the genome sequence is thought to indicate related functions.

The connection between CCM and replication doesn't seem to be limited to prokaryotes. In budding yeast, the following results were obtained: (i) three CCM enzymes (*HXK2*, *ENO1* and *GPM1*) were found to be functionally linked to MCM1[166]. MCM1 is a multifunctional protein that can bind DNA to stimulate replication initiation and also controls the expression of many genes, including some involved in DNA replication[167,168]. (ii) Inactivation of glycolytic genes in yeast arrested them at the G_1 phase (*PYK1*) or at the S-phase (*PGI1*)[169–171]. (iii) Mutations in Cdc40p, a splicing factor required for both G1/S and G2/S transition are suppressed by overexpressing the glycolytic enzymes *GAPDH*, *PGK1*, *GPM1* and *ENO2*[172]. (iv) The origin recognition complex (ORC) interacted with CCM genes near replication origins[173].

In human fibroblasts, recent studies found that silencing of several glycolytic (*HK2*, *TPI*, *GADPH*, *ENO1* and *LDHA*) ,TCA cycle (*CS1*, *ACO2*, *SUCLG2*, *SDHA*, *FH* and *MDH2*) and PPP genes (*H6PD*, *PRPS1*, *G6PD*, *RBKS* and *TALDO*) impacted either the delay for entry into S-phase or its efficiency or both.[174,175].


GLU
HXK2 (a) HK2 (b)
G6P
PGI1 (b)
F6P
pfk (+)
FBP
ALDO
ALDO
TPI b
G3P
DHAP
gapA (a,b) GADPH (a) GADPH (b,c,e)
1,3PG
pgk (a,b) PGK1 (a) PGK (c)
3PG
pgm (a) gpmA (a) GPM1 (a)
2PG
eno (a) ENO1 (a) ENO2 (a) ENO1 (b)
ackA (a,b) pta (a), ackA (a) LDH (b,c)
PEP
+ ATP
pykA (a,b,+) PYK1 (b) PKM2 (d)
Overflow
PYR


PYR
pdh (c) pdh (a) PDH (d)
Ac-CoA (d) Ac-CoA (d)
PC
CS1 (b)
OXA
CIT
MDH2 (b)
ACO2 (b)
MAL
ISOCIT
FH (b)
IDH
FUM
α-KG
SDHA (b)
KGDH
SUC
SUC-CoA
SUCGL2 (b)

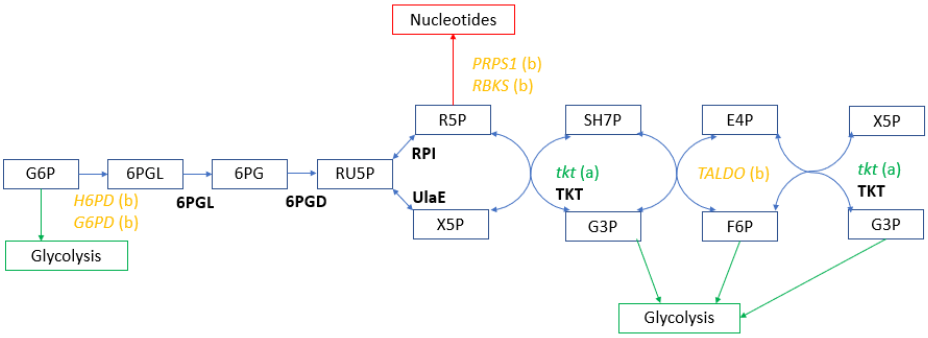


Figure 17: Connections between CCM and replication in the model organisms B. subtilis, E. coli, S. cerevisiae and higher eukaryotes (human/plant)

Different colors are used for observed connections in different organisms (Blue) B. subtilis (Green) E. coli (Red) S. cerevisiae (Orange) Higher eukaryotes (Human/Plant) Different kinds of experimental support are depicted using a letter code: (a) Suppression of a mutation in a replication gene or a regulator of replication (b) Deletion or knockdown affects replication parameters or the transition into S-phase (+) Metabolic genes are located close to a replication protein in many bacteria, indicating a functional connection. (c) Direct protein-protein interactions with a replication protein or regulator of replication (d) Stimulates post-translational modification of histones relevant for S-phase progression (e) Stimulates S-phase expression of histones relevant for S-phase progression through protein-protein interactions. Depicted are connections for glycolysis (top), TCA cycle (middle) and PPP (bottom)

3.3.2 Mechanistic evidence for a connection between CCM enzymes and DNA replication

Theoretically, the mechanisms connecting CCM and replication can be divided into two broad categories: indirect and direct mechanisms.

Given its strong interconnection with all cellular processes, it is highly likely that CCM influences replication indirectly. First, CCM is responsible for the production of nucleotides through PPP, which are necessary for successful completion of replication as discussed earlier.[123,176]. Second, deletion of CCM activities can sometimes lead to the activation of stress responses that can influence replication. For instance, a study in *E.coli* found that the suppression of temperature sensitivity in DnaA by the previously found *ackA-pta* pathway[163,164] were due to an activation of the stress response pathway[177]. Furthermore, the deletion of the *ackA-pta* pathways were associated with an accumulation of pyruvate and resulted in an altered activity of pyruvate dehydrogenase (PDH)[177]. Deletion of parts of *pdh* also resulted in suppression of Ts DnaA[177]. These results underscore the difficulty of interpreting these deletion studies of CCM enzymes due to pleiotropic effects.

Nevertheless, evidence is accumulating that CCM may play a more direct role in regulation of

replication timing as well.

Direct protein-protein interactions (PPI) have been shown to modulate the activity of replication proteins. DnaE, DnaC and DnaG were shown to interact in *B.subtilis* in a ternary complex and this interaction was shown to improve the fidelity of DnaE[178]. Later, it was found that PPI between DnaE and other replication proteins also improved its fidelity considerably and that DnaE was essential for both initiation and elongation[179]. Similar PPI interactions between replisome components through the Ctf4 hub protein are important in yeast and human replisomes for proper DNA replication[180].This suggests that the replisome is a potential target for controlling the timing of replication through direct PPI.

In bacteria, some CCM enzymes directly interact with replication proteins and modulate their activity. A yeast-two-hybrid screen in *B.subtilis* revealed that DnaC and DnaG interact with subunits of the pyruvate dehydrogenase, branched ketoacid dehydrogenase and acetoin dehydrogenase[181]. The binding of the primase DnaG to pyruvate dehydrogenase was thought to inhibit replication[182].

Direct interactions between CCM enzymes and replication proteins have been observed in eukaryotes as well.

In humans, nearly all CCM enzymes enter the nucleus, making direct interactions *in vivo* possible[183,184].Additionally, some human CCM enzymes (*LDH, PGK, GADPH*) bind DNA and stimulate replication *in vitro*[185–188] and *in vivo (PGK)*[189,190].

In plants, both GADPH and PGK influence viral replication through direct interaction with replisomes of several viruses [121,122,191,192].

Post-translational modifications of replication proteins constitute another method to control DNA replication timing within the cell cycle. It is a well-known system for regulating cellular activities in eukaryotes, including replication (See the CDK system for regulating eukaryotic cell cycle: section 1.1) More recent lines of evidence suggest that post-translational modifications may play a key role in regulating replication activity in prokaryotes as well. First, proteome technologies have revealed post-translational modifications in many bacterial proteins, including replication proteins[193–196]. Second, independent studies have shown that modifications of replication proteins in prokaryotes happen and alter their activity. For instance, DnaA is acetylated in *E.coli* and this acetylation is important for its activity[69]. Additionally, phosphorylation of the bacterial single stranded DNA binding protein (SSB) is also important for replication[197,198]. However, direct post-translational modifications of replication proteins by CCM effectors have not yet been described to the best of our knowledge. By contrast, histone modification and cell cycle dependent histone expression regulate gene

expression, control the timing of DNA replication and therefore constitute another level of control on DNA replication timing (for review, see [199]).

Indeed, CCM effectors directly influence the (modification) status of histones to promote growth and DNA replication.

In yeast, the CCM metabolite acetyl-CoA stimulates growth by increasing histone acetylation near genes involved in growth[120]. One of these genes is *CLN3*, an early G_1 cyclin that permits the subsequent expression of hundreds of cell cycle genes[200].

In humans, the following results were obtained: (i) GAPDH proved to be an essential part of the OCA-S complex. This complex binds tightly to the Oct-1 transcription activator, which is responsible for the expression of H2B, an essential histone for S-phase progression[201]. (ii) PDH translocates in a cell cycle dependent manner into the nucleus and generates the acetyl-CoA needed for correct acetylation of histones involved in progression through the G_1/S transition and S-phase[202]. (iii) PKM2 phosphorylates histone H3 and changes expression of cyclin D1 among others, which is important for cell cycle progression in cancer cells[203].

In summary, a lot of evidence has emerged that CCM activity is indeed coupled to replication control. In this work, the role of the glycolytic enzyme pyruvate kinase in DNA replication timing control is evaluated.

4 Pyruvate kinase in *B. subtilis*: structure, function and role in replication

4.1 The metabolic function of pyruvate kinase in *B. subtilis*

Pyruvate kinase is a strongly conserved CCM enzyme that catalyzes the final reaction of glycolysis. It converts PEP and ADP into pyruvate and ATP. This reaction is part of a group of reactions called the PEP-pyruvate-oxalate node, reviewed in [204] .

The PEP-pyruvate-oxalate node is an important metabolic switch point that directs carbon fluxes to appropriate directions and helps control the switch between gluconeogenesis and glycolysis[204]. Under glycolytic conditions, the final products of glycolysis (PEP and pyruvate) enter the TCA cycle to replenish molecules withdrawn for anabolic purposes using pyruvate kinase (*pykA)* and pyruvate carboxylase (*pycA*)[204]. Under gluconeogenesis, the TCA cycle intermediates oxaloacetate or malate are converted into the intermediates of glycolysis that were used for anabolic conditions using the enzyme PEP carboxykinase (*pckA*)[204]. Thus, the enzymes of the PEP-pyruvate-oxalate node in general are crucial for the switch between

glycolysis and gluconeogenesis.

The activities of these enzymes during the switch between gluconeogenesis/glycolysis is partially transcriptionally and partially post-transcriptionally regulated. *pckA* expression is repressed under glycolytic conditions[205–207], but *pykA* and *pycA* expression is more or less constitutive[207–209]. This constitutive investment in PykA production under gluconeogenic conditions is remarkable, since *pykA* is one of the most highly expressed genes in the cell**.** By contrast, PycA is strongly allosterically regulated using acetyl-CoA[209], but PykA activity is only known to respond to the concentrations of its substrate PEP[208].

4.2 A role for PykA in cell cycle control

Recent evidence in the model bacterium *B. subtilis* suggests that pyruvate kinase (*pykA)* may be involved in control of DNA replication timing.

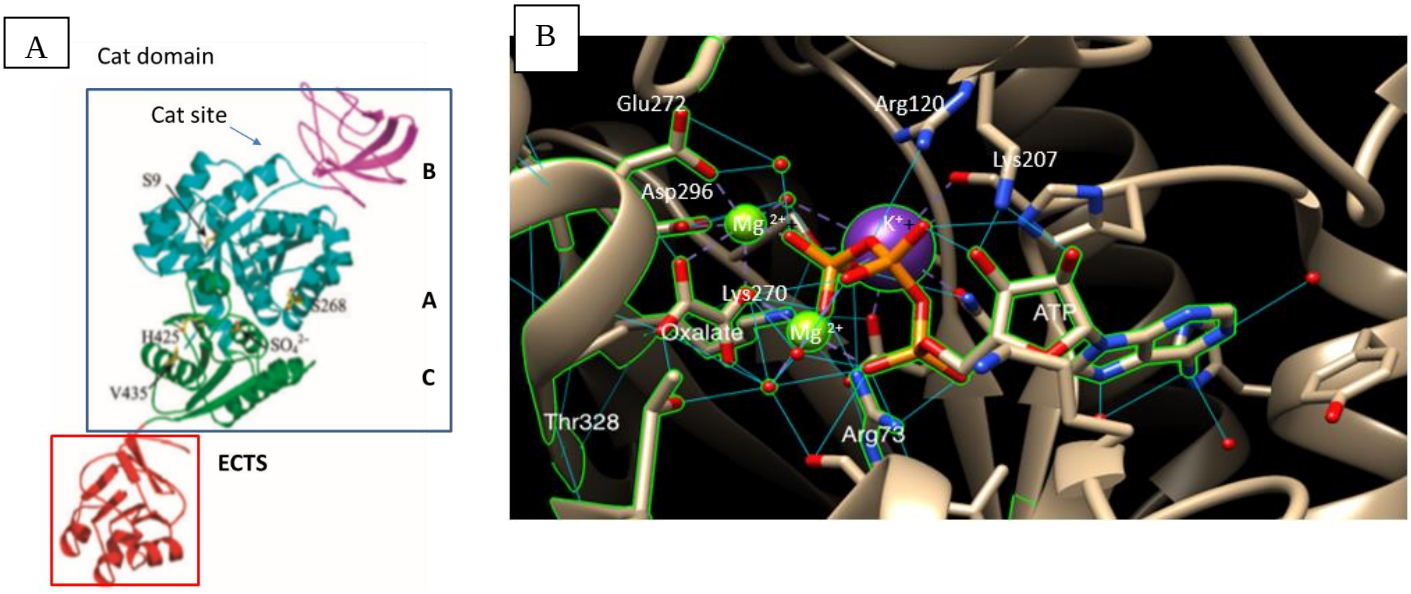
First, a mutation in pyruvate kinase (*pykA*) in *B. subtilis* was able to suppress Ts mutations in the *dnaE* gene in LB and this suppression was proposed to occur through a post-translational mechanism[162]. In a follow-up study, *pykA* deletion was also found to affect DNA replication initiation in LB and LB + malate [86]. In the same organism, *pykA* deletion rescues an assembly defect in a Ts sensitive FtsZ ring assembly required for cell division and causes a defect in cell division in wild type backgrounds, which can be rescued by exogenous addition of pyruvate[210]. Taken together, these results indicate that PykA is required for proper cell cycle and replication control in *B.subtilis*. However, given their reliance on deletion of *pykA* , it is not clear from these studies if it is required for its catalytic activity or if the PykA protein itself plays a role in replication control. In the next section, the structure of the PykA protein is reviewed.

4.3 Structure of pyruvate kinase (PykA) in *B. subtilis*

The structure of pyruvate kinase has been solved for many organisms[211–218], but unfortunately not for *B. subtilis*. However, the structure of a near identical pyruvate kinase from the closely related species *Geobacillus staerothermophilus* has been solved[214] and is shown in figure 18A. The overall structure of pyruvate kinase is strongly conserved throughout the tree of life and reviewed in[219]. Most pyruvate kinases exist in the cell as homotetramers[214].

Pyruvate kinase monomers consist of three domains (A,B and C) that are highly conserved throughout the tree of life and together constitute the catalytic domain[211–218]. Eukaryotic pyruvate kinases have an additional N-terminal domain[211,216,217], which bacterial pyruvate kinases lack[212–214,218]. However, the pyruvate kinases of *B. subtilis* and *G. staerothermophilus* have an extra C-terminal sequence ECTS[214,220], which is conserved among the *Bacilli*[220].

In general, pyruvate kinases are formed by monomer-monomer and dimer-dimer interactions[219]. Monomer-monomer interactions are mediated through interactions of the A-domain along a 2-fold axis, also called the r-axis or large interface[219]. The dimer-dimer interface of PykA is mediated by the C domains along a 2-fold axis called the q-axis or small interface[219]. Additionally, the extra ECTS domain is predicted to interact with a neighboring A domain and is thought to stabilize the tetramer, even though its deletion does not affect catalysis[214].



C

	PykA (*B. subtilis*)	PKM2
Binding site (PEP)	R32	R73
Metal binding (K)	N34	N75
Metal binding (K)	S36	S77
Metal binding (K)	D66	D113
Metal binding (K)	T67	T114
Binding site (ATP)	R73	R120
Binding site (ATP)	K156	K207
Transition state stabilizer	K220	K270
Metal binding (Mg)	E222	E272
Binding site (PEP)	G245	G295
Metal binding (Mg)	D246	D296
Binding site (PEP)	D246	D296
Binding site (PEP)	T278	T328

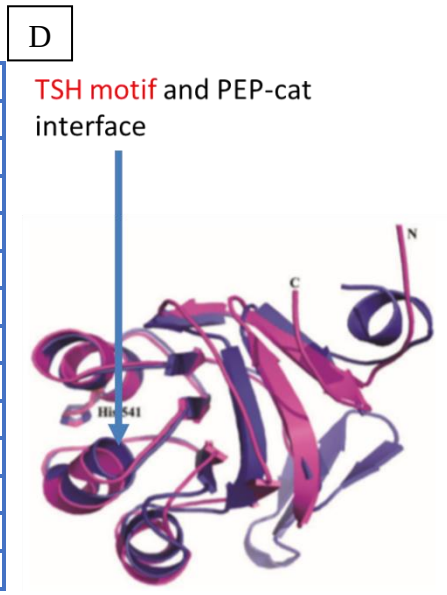


Figure 18: The structure of pyruvate kinase

(A) Domain structure of the monomer pyruvate kinase in G. staerothermophilus .Shown is

domain A (cyan), domain B, (magenta) domain C (green) and the ECTS Picture taken from [214] (Red) (B) Catalytic site of human PKM2 complexed with substrates Picture was made using the visualization software Chimera. Crystal structure obtained from [214] (C) Important and corresponding catalytic sites and their function in B. subtilis and PKM2. List obtained from UniProt. (D) Structure of the PEPut domain

The catalytic site of pyruvate kinase has been studied in detail and conserved residues important for catalysis are summarized in fig. 18B and C. The pyruvate kinase catalytic site requires one divalent cation (Mg^{2+}) and one monovalent kation (K^{+}). Other key residues are 3 positively charged amino acids (2 R and 1 K) and 4 negatively charged residues (2 D and 2 E). In *B. subtilis*, K^{+} is coordinated by the amino acids N34, S36, D66 and T67. Mg^{2+} is coordinated by the amino acids E222 and G245D246. The substrates PEP and ADP are partly bound by these cations, with Mg^{2+} being essential for PEP binding. PEP is additionally bound by R32, G245D246 and T278. ADP, on the other hand, is bound by R73 and K156. Incidentally, mutation of the equivalent amino acid of R73 in rabbit pyruvate kinase has been shown to decrease ADP binding[215]. Finally, phosphate transfer from PEP to ADP is stabilized by K220. Mutation of this amino acid in pyruvate kinase of *G. staerothermophilus* or yeast has been shown to completely abolish catalytic activity, though its effect on substrate binding is less clear[221,222].

In contrast to the highly conserved catalytic mechanism, regulation of Pyk in different organisms is much more variable[219]. Most Pyk are either regulated by FBP or by AMP and sugar monophosphates[219]. PykA is regulated by ribose-5-phosphate and by AMP[214]. The effector binding site in most Pyk is situated in the C-domain and usually consists of a phosphate binding loop, a mobile loop and an effector loop[219]. In general, the effector loop remains buried in a pocket formed by the phosphate binding loop[219]. Upon binding of the phosphate binding loop to the effector, the effector loop becomes ordered. The effector binding site in Pyk of *G. staerothermophilus* is thought to interact with R5P and AMP through the amino acids T381, S383, T386 and H425[214], but no structural data is available of Pyk of *G. staerothermophilus* complexed with an effector[219]. Mutation of H425 to alanine led to a decrease in AMP and to a lesser extent R5P binding[214]

Currently, the function of the ECTS of pyruvate kinase in *Bacilli* remains unknown, but its structure is shown in figure 18D. The sequence is homologous to the one of PEP utilizer domains found in the PTS component EI and the metabolic enzymes pyruvate phosphate dikinase (PPDK) and PEP synthase[220]. In the PTS component EI, the PEP utilizer domain hosts

the conserved histidyl that transfers phosphate abstracted from PEP to HPr[223]. This similarity might suggest that it is responsible for PTS regulation of its activity. However, deletion of the PEP utilizer domain in *G. staerotermophilus* pyruvate kinase had no effect on its metabolic activity[220]. By contrast, in the metabolic enzyme PPDK, the PEP utilizer domain is a key part of the catalytic activity. It is responsible for turning ATP and pyruvate into AMP, pyrophosphate and PEP and the conserved histidyl transfers the phosphate groups between the nucleotide binding domain and the pyruvate binding domain[224]. Interestingly, a TSH motif around the catalytic H is conserved in plant PPDK and in *Bacilli* PEP utilizers. The T in this motif regulates the catalytic activity of H during the light/dark transition in plant PPDK[225], whereas the middle S has no effect on catalytic activity or substrate specificity[226].

5. Objectives

The major goal of this work is to demonstrate for the first time that (i) PykA moonlights in the control of DNA replication gating and to generate insights into the moonlighting mechanism and (ii) that it can couple DNA replication gating to the metabolic state of the cell.

The first goal of this work was to demonstrate that PykA plays a role in control of DNA replication gating beyond its catalytic activity and to identify the regions by which it may exert control over replication timing. This was achieved by (i) analysing the effect of *pykA* mutations in the PEP utilizer and catalytic domains on replication and cell cycle control in a medium in which PykA metabolic activity is dispensable and (ii) by integrating *in vivo* analysis methods (qPCR, flow cytometry and microscopy) with *in vitro* replication assays to get a holistic view of the effects of PykA on replication. These results are described in Chapter 2 and will be published in Spring 2020.

The second goal of this work was to demonstrate that PykA is not just a replication protein, but also couples DNA replication timing to the metabolic state of the cell. This was achieved by analysing the effects of *pykA* mutations during a shift from a rich gluconeogenic to a rich glycolytic medium using the analysis methods from the first part. These results are described in Chapter 3.

Strain construction was performed as indicated in Supplementary table I. Growth rate and qPCR experiments (*ori-ter*, marker frequency) were performed by Steff Horemans (except for mutated *hyPEPut* constructs). Flow cytometry was performed by Laurent Jannière, Gabor Gyapay, Nathalie Vega-Czarny and Marcel Salanoubat . *In vitro* experiments were performed by Alexandria Holland, Matthaios Pitoulias and Panos Soultanas.

6. Key methodology used in this thesis

6.1 Marker frequency analysis and (*ori-ter* ratio) for determining replication speed

Ori-ter ratio analysis is a method that enables the detection of the average number of ongoing replication events(Fig. 19A)[82]. It is based on the quantification of the number of DNA copies close to the origin versus the number of DNA copies close to the terminus[82]. For instance, if a cell is in the process of replicating its chromosome (single replication), this cell will contain twice as many origin sequences compared to terminus sequences. Conversely, the ratio becomes one again when the chromosome is fully replicated. Hence, the *ori-ter* ratio reflects the time it takes to complete chromosome replication and is used in literature to calculate the C-period[82]. In this work, the *ori-ter* ratio is used as a general replication parameter to see if replication is impaired in a particular strain. The C-period and replication speed are determined using the quantification of about 10 DNA sequences between the origin and terminus, which allows for a more accurate determination.

Quantification of DNA sequences is done using qPCR in this work, but could also be done by deep sequencing (Fig 19B)[227]. Like conventional PCR, DNA is denatured, primers anneal to the region of interest and the primers are extended using a DNA polymerase, doubling the amount of DNA during each cycle[227].. Unlike conventional PCR, an extra dye (SyBr Green) is present that becomes fluorescent upon binding to DNA. This fluorescence is proportional to the amount of DNA that is present, is captured by a real-time PCR machine and enables accurate quantification of the amount of DNA present in the sample[227].

A

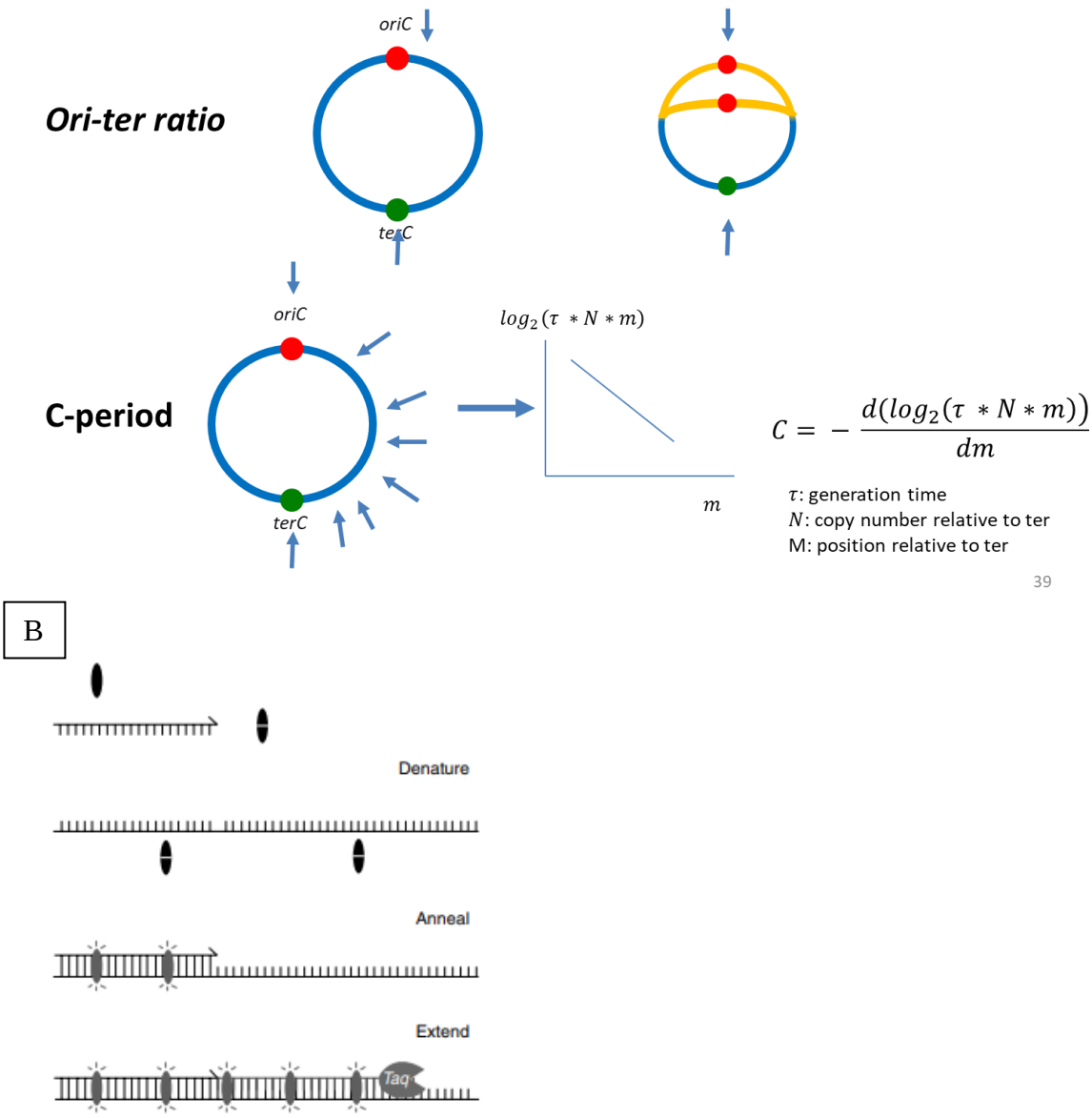


Figure 19: Principles of marker frequency analysis for determining replication fork speed (A) Basic idea behind ori-ter ratio and marker frequency analysis (B) Quantification of gene copy number by qPCR using SybR Green Taken from [227]

6.2 Run-out flow cytometry for determination of initiation parameters

Initiation parameters were determined using run-out flow cytometry in this work[228]. If a cell is replicating, it has two or more origins per cell (Fig 20A). During run-out flow-cytometry, chloramphenicol is added to the culture. This prevents future initiation events, but allows the current active cycles to be finished. After 4h-6h of waiting, the chromosomes are fully replicated and the DNA content in each cell reflects the original number of active origins per cell. Next, the DNA content of each cell is stained with a fluorescent dye (DAPI or Hoechst) and quantified using flow cytometry (Fig 20B). From this experiment, the number of active

origins per cell can therefore be calculated, a measure of initiation frequency. Furthermore, the proportion of cells that haven't started initiation can be used to determine when initiation started during the cell cycle i.e. the so-called age of initiation.

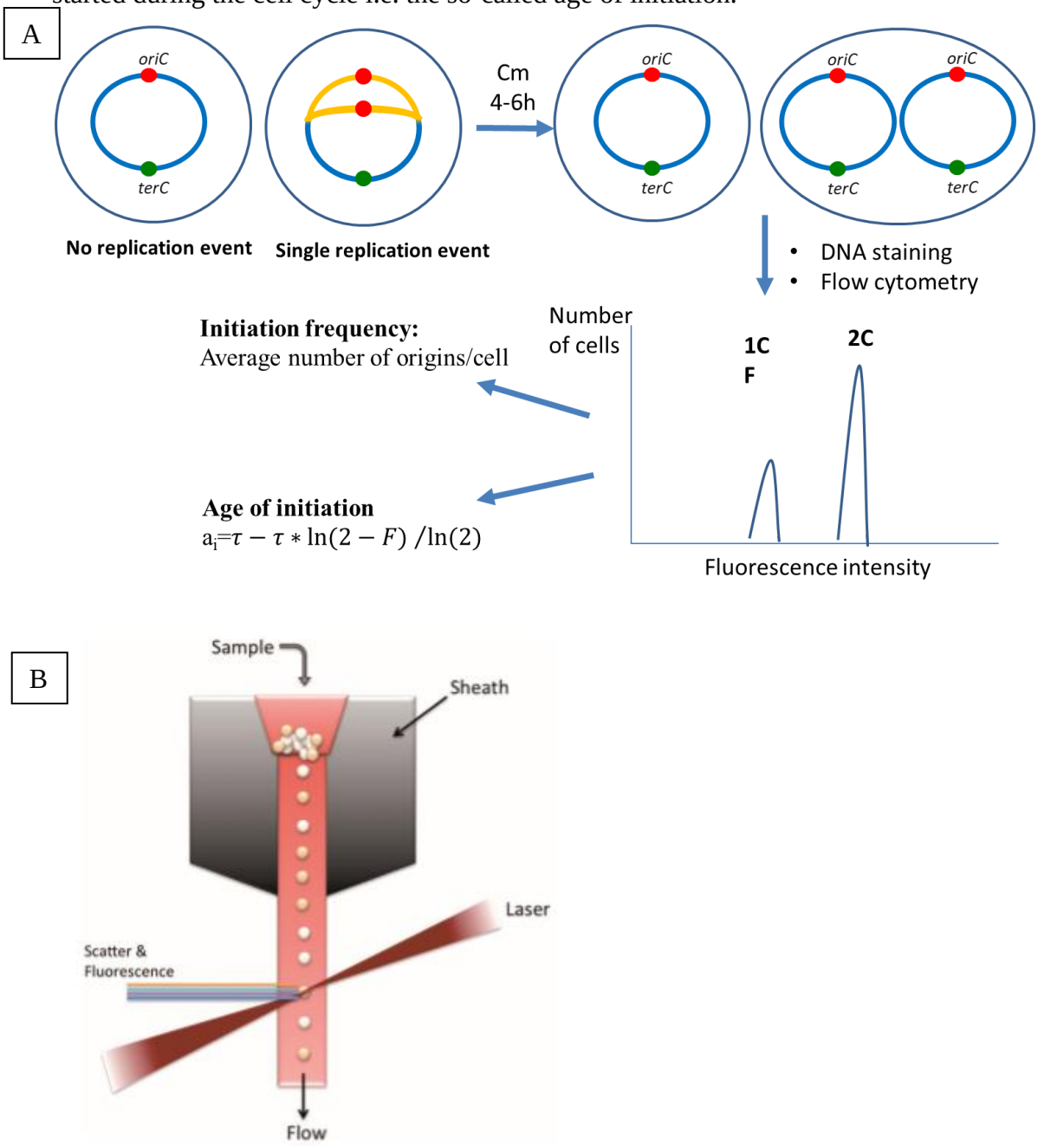


Figure 20: Principles of run-out flow cytometry for determining initiation parameters

(A) Basic principle of using flow cytometry for determining initiation parameters (B) Basic principle of quantification through flow cytometry Taken from [229]

6.3 *DnaE polymerase and DnaC helicase assays*

The basic principles behind the DnaE polymerase and DnaC helicase activity assays by the team of Panos Soultanas are depicted in fig. 21[178]. Briefly, DnaE polymerase activity is measured by mixing the DnaE polymerase with MS13 DNA and a single stranded radioactive

primer that binds to MS13 DNA (Fig 20A). DnaE extends this primer as part of its activity and the more active it is, the more DNA product is produced. The DNA product produced by DnaE is radioactive due to the radioactive primer, which is used to quantify the DNA product using gel electrophoresis and radiography.

Helicase activity is measured in a similar way (Fig 20B) [178]. First, a radioactive primer is annealed to MS13 DNA. Helicase is then incubated with that. Helicase then displaces the radioactive primer from MS13 as part of its activity, which leads to its dissociation from the column. The amount of dissociated primer versus bound primer is determined from using gel electrophoresis and radiography.

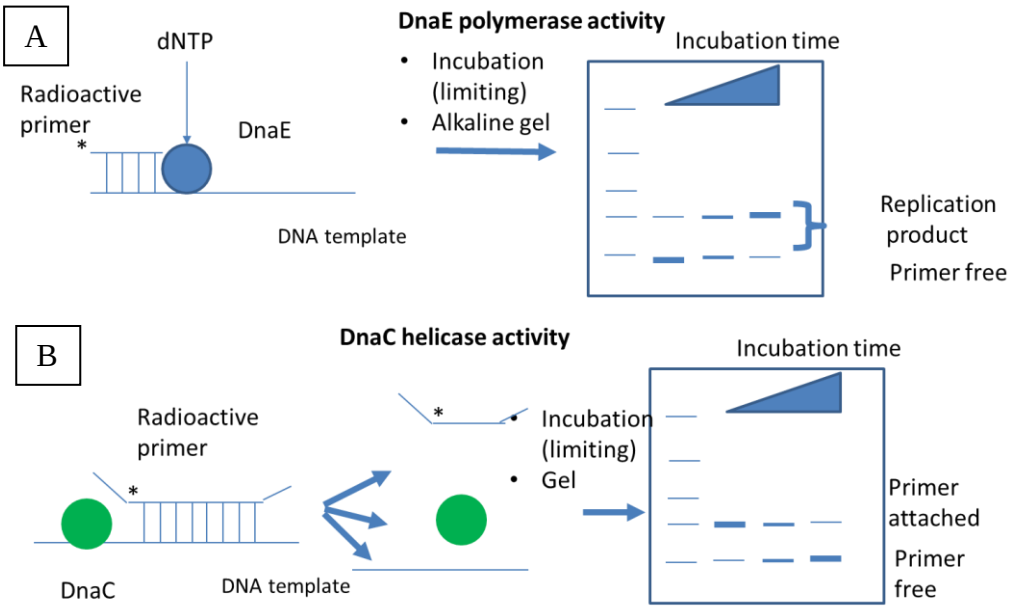


Figure 21: DnaE polymerase activity and DnaC helicase activity assays

(A) Principle of DnaE polymerase activity measurement (B) Principle of DnaC helicase activity measurement

Chapter 2: PykA controls DNA replication timing

The glycolytic pyruvate kinase of *Bacillus subtilis* moonlights in replication via several pathways including direct interactions with replication enzymes for helping cells to temporalize DNA replication in the cell cycle

Steff Horemans[1], Alexandria Holland[2], Matthaios Pitoulias[2], Panos Soultanas[2¶] and Laurent Janniere[1¶]

[1] : Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France

[2] : Centre for Biomolecular Sciences, School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK

[¶] : Corresponding authors

Laurent Janniere: laurent.janniere@univ-evry.fr

Panos Soultanas : panos.soultanas@nottingham.ac.uk

SUMMARY

DNA replication is a complex biological process required for ensuring faithful reproduction of genetic material in progeny. To cope with nutritional conditions, cells evolved sophisticated systems driving a metabolic control of DNA replication. These controls are thought to depend, at least in part, on central carbon metabolism. As the glycolytic pyruvate kinase was proved tightly connected to replication, its role in DNA synthesis is investigated here in the model bacterium *Bacillus subtilis*. In this organism, the pyruvate kinase (PykA) consists of the catalytic domain (Cat) and of an extra C-terminal domain homologous to PEP utilizers (PEPut). In other metabolic enzymes, the PEP utilizer is phosphorylated at a conserved TSH motif to drive sugar import and catalytic or regulatory activities. By measuring various replication parameters in cells encoding the wild-type or mutated forms of PykA, we identified Cat and PEPut mutations associated with replication defects. Evidence showing that these defects are due to changes in the PykA protein amino-acid chain rather than on an alteration of its metabolic activity were then provided. Interestingly, the replication defects found in Cat and PEPut mutations are opposite: while in Cat mutants the velocity of replication forks is reduced and the age of initiation advanced, forks are very fast and initiation is delayed in PEPut mutants. Further investigations showed that Cat residues involved in phosphoenolpyruvate binding and in a genetic link with the replication enzymes DnaC, DnaG and DnaE are important for replication as well as the H amino-acid and phosphorylation of the T residue of the TSH motif of PEPut. Finally, *in vitro* studies showed that the Cat domain of PykA affect replication activities of DnaE and DnaC probably via direct protein-protein interactions. Overall, these results show that PykA moonlights in replication independently of its metabolic activity. This non-metabolic activity is needed to help cells to temporalized DNA replication in the cell cycle and operates through different pathways driven by the Cat and PEPut domains and involving direct and indirect effect on replication enzymes involved in initiation and elongation of replication and possibly phosphorylation events. By assigning a moonlighting activity in replication control to CCM enzymes, this work suggests that CCM mutations can increase genetic instability and pave the way to tumorigenesis by altering replication control.

INTRODUCTION

DNA replication is gated to a specific timeslot within the cell cycle to ensure that chromosome replication occurs once and only once every cell cycle. This gating is maintained by two distinct systems. First, overlapping mechanisms control the activity of replication initiators and replication origins. These are responsible for initiating DNA strand opening and recruiting the replication machinery at origins every cell cycle. (Jameson & Wilkinson, 2017; Siddiqui et al., 2013) and prevent re-initiation before the end of the cell cycle. Second, the timing of DNA replication is also connected to the availability of nutrients.(Buchakjian & Kornbluth, 2010; Ewald, 2018; Helmstetter, 1996). In bacteria, this higher level of regulation couples the rate of replication to the growth rate by modulating the initiation frequency and/or the speed of replication (Bipatnath et al., 1998; Schaechter et al., 1958; Sharpe et al., 1998; Skarstad et al., 1986). In *Saccharomyces cerevisiae* (and possibly higher eukaryotes), metabolic control confines DNA synthesis to a specific phase of a metabolic cycle (Burnetti et al., 2016; Klevecz et al., 2004; Papagiannakis et al., 2017; Slavov & Botstein, 2011; Tu, 2005). In either case, the net result of these two systems is a precise and reproducible timing of DNA synthesis within the cell cycle across a wide range of nutritional conditions.

In contrast to the classical regulatory mechanisms, the exact nature of determinants involved in the metabolic control of replication remains elusive. In bacteria, long-standing hypotheses postulating that this control depends on (i) the concentration of the active form of the replication initiator DnaA (DnaA-ATP), (ii) on restricting DNA polymerases activity by limiting precursor concentrations or (iii) on the concentration of the alarmon (p)ppGpp have been challenged (Flåtten et al., 2015; Hernandez & Bremer, 1993; Mathews, 2015; Murray & Koh, 2014). For this reason, several groups have started to argue that metabolic control of replication is a multifactorial process where the effectors vary with nutrient richness and which may involve directly sensing cellular metabolism and communicating this information to the replication machinery (see for instance (Barańska et al., 2013; Boye & Nordström, 2003; Wang & Levin, 2009).

Central carbon metabolism (CCM) is a collection of metabolic reactions thought to be particularly involved in communicating information about the metabolic state of the cell(Nouri et al., 2018). In most organisms, it consists of glycolysis/gluconeogenesis, tricarboxylic acid (TCA) cycle, the overflow pathway and penthose phosphate pathway (PPP) (Kim & Gadd, 2008). Collectively, these reactions convert nutrients into the energy and the nucleotides, sugars, lipids and amino acids the cell needs for survival (Kim & Gadd, 2008). Moreover, the activity of CCM is dynamically

modulated by a complex regulatory system that ensures optimal CCM activity across different nutritional conditions(Deutscher et al., 2014; Görke & Stülke, 2008). This regulatory system involves sensing the concentration of key CCM metabolites such as PEP and adjusting the expression of metabolic enzymes accordingly(Deutscher et al., 2014). This group of highly conserved reactions is therefore in a strategic position for producing signals that cells may use to couple key cellular decisions to the metabolic state of the cell, as was recently demonstrated for cell division in *Escherichia coli* and *Bacillus subtilis* (Chien et al., 2012; Monahan et al., 2014; Weart et al., 2007).

In addition, several lines of evidence have suggested that CCM might be involved in the gating of DNA replication in different organisms. In *B. subtilis*, deletion of the metabolic enzymes responsible for the terminal four reactions of glycolysis rescued temperature sensitive mutations in the primase DnaG, the helicase DnaC and the lagging strand polymerase DnaE in the LB medium (Jannière et al., 2007). Furthermore, follow-up studies showed that these mutations also impacted the gating of DNA replication within the cell cycle (Murray & Koh, 2014; Nouri et al., 2018). Moreover, similar connections between CCM and DNA replication gating have been found in *E. coli* (Maciąg et al., 2011, 2012), yeast (V. K. Chang et al., 2004; Victoria K. Chang et al., 2003; Chen & Tye, 1995; Dickinson & Williams, 1987; Hartwell, 1973; Kaplan & Kupiec, 2007; Shor et al., 2009; Sprague, 1977) and human cells (Fornalewicz et al., 2017; Konieczna et al., 2015) albeit involving different metabolic reactions, suggesting a conserved functional link between CCM and the gating of DNA replication.

However, the main limitation of the studies connecting CCM to gating is the drastic physiological effects of depletion of metabolic enzymes, complicating their interpretation. In *E. coli*, deletion of the overflow pathway enzymes *ackA-pta* has been consistently functionally connected to DNA replication and repair(Maciąg et al., 2011, 2012, 2012). However, a follow-up study found that deletion of *ackA-pta* lead to an activation of stress responses, an accumulation of pyruvate and an altered activity of another metabolic enzyme (*pdh*) in LB (Tymecka-Mulik et al., 2017). At present, it is unclear whether CCM indirectly influences DNA replication through its metabolic importance for the cell, or whether some CCM enzymes/metabolites have an additional moonlighting function in gating DNA replication. In spite of this limitation, it is likely that at least some CCM enzymes moonlight in DNA replication gating. First, CCM enzymes are well-documented to have several moonlighting functions in bacteria to higher eukaryotes (see for review (Commichau & Stülke, 2007; Snaebjornsson & Schulze, 2018)). Second, several CCM enzymes from different organisms directly interfere with DNA replication *in vitro*(Baxi & Vishwanatha, 1995; Grosse et al., 1986; Huang & Nagy, 2011; Jindal & Vishwanatha, 1990; Kumble et al., 1992; Li et al., 2018, p. 1; Popanda et al., 1998; K.

R. Prasanth et al., 2011; K. Reddisiva Prasanth et al., 2017; Tian et al., 2015).

One of the CCM enzymes most frequently associated with both DNA replication gating and moonlighting is pyruvate kinase (Pyk). Deletion of *pyk* suppressed temperature sensitive mutations in replication enzymes in *B. subtilis* (Nouri et al., 2018) and *S. cerevisae* (Dickinson & Williams, 1987; Hartwell, 1973; Sprague, 1977). Additionally, pyruvate kinase isoform 2 (PKM2) in human cells is known to moonlight in a wide variety of cellular processes, including the regulation of S-phase histone expression, DNA repair, DNA segregation, transcription and cell-cell communication (Snaebjornsson & Schulze, 2018).

Here, we report on a possible moonlighting role of pyruvate kinase (PykA) in DNA replication gating in *B. subtilis*. PykA catalyzes the final reaction of glycolysis by converting phosphoenolpyruvate and ADP into pyruvate and ATP and is an important regulator of glycolytic flux (Schormann et al., 2019). The 3D structure of pyruvate kinase has been solved for many organisms(Cheng et al., 1996; Larsen et al., 1998; Mattevi et al., 1995; Muirhead et al., 1986; Rigden et al., 1999; Suzuki et al., 2008; Zhong et al., 2017), but not for PykA . However, a structure has been obtained for the one of *G. staerothermophilus*, which only differs by two amino acids (Suzuki et al., 2008).

In general, pyruvate kinase is a fairly well conserved protein across all different organisms (Schormann et al., 2019). Pyruvate kinase exists as a homotetramer in most known organisms, although monomers and dimers have been observed (Schormann et al., 2019). Additionally, the overall structure and active site configuration are also strongly conserved in all known organisms. On the other hand, the activity of pyruvate kinases is regulated by a variety of effectors, most notable fructose-1,6-bisphosphate (FBP) or AMP and sugar monophosphates like ribose-5-phosphate(Schormann et al., 2019). The mechanisms behind this regulation shows remarkable versatility, making pyruvate kinase an interesting drug target, despite the large degree of conservation of the active site(Schormann et al., 2019).

Pyruvate kinase in *B. subtilis* consists of 4 domains: the A, B, C and Extra C-terminal sequence (ECTS)(Suzuki et al., 2008). The A, B and C domains collectively make up the catalytic site, are strongly conserved and are sufficient for catalysis(Suzuki et al., 2008). The monomer-monomer interface of pyruvate kinases is mediated by the A domains along a 2-fold axis called the r-axis or large interface(Schormann et al., 2019). The dimer-dimer interface of PykA is mediated by the C domains along a 2-fold axis called the q-axis or small interface(Schormann et al., 2019). Additionally, the extra ECTS domain is predicted to interact with the A domain of a neighboring monomer through the amino acids E209 and L508(Suzuki et al., 2008). This interaction is thought to stabilize the

tetramer(Suzuki et al., 2008). The catalytic domain also contains the catalytic site and the effector binding site that regulates PykA catalytic activity.

The catalytic site is situated at the interface of the A and B domain in most Pyk, including PykA. The residues important for catalysis in the active site are strongly conserved. (Schormann et al., 2019). The corresponding amino-acids in the *B. subtilis* protein are listed Fig. S1A. Amino acids N34, S36, D66 and T67 coordinate K^+ while amino acids E222 and G245D246 coordinate Mg^{2+} (Schormann et al., 2019). The substrates PEP and ADP are partly bound by these cations, with Mg^{2+} being essential for PEP binding(Schormann et al., 2019). PEP is additionally bound by R32, G245D246 and T278. ADP, on the other hand, is bound by R73 and K156. Finally, phosphate transfer from PEP to ADP is stabilized by K220. Mutation of this amino acid completely abolishes catalytic activity, though its effect on substrate binding is less clear(Schormann et al., 2019).

The effector binding site in most Pyk is situated in the C-domain and usually consists of a phosphate binding loop, a mobile loop and an effector loop(Schormann et al., 2019). In general, the effector loop remains buried in a pocket formed by the phosphate binding loop (Schormann et al., 2019). Upon binding of the phosphate binding loop to the effector, the effector loop becomes ordered(Schormann et al., 2019). The effector binding site in Pyk of *G. staerothermophilus* is thought to interact with R5P and AMP through the amino acids T381, S383, T386 and H425(Suzuki et al., 2008), but no structural data is available of Pyk of *G. staerothermophilus* complexed with an effector(Schormann et al., 2019). However, mutation of H425 to alanine did lead to a decrease in AMP and to a lesser extent R5P binding (Suzuki et al., 2008).

In addition to the catalytic domain, the pyruvate kinases of *B. subtilis* and related species have a conserved extra C-terminal sequence (Nguyen & Saier, 1995; Sakai, 2004). The function of this C-terminal peptide is not known and its deletion in the related *G. staerotermophilus* pyruvate kinase has no effect on its catalytic activity (Sakai, 2004). However, the sequence of the extra peptide is homologous to the one of PEP utilizer (PEPut) domains found in the phosphoenolpyruvate:carbohydrate phosphotransferase (PTS) component EI and the metabolic enzymes pyruvate phosphate dikinase (PPDK) and PEP synthase (PEPS) (Fig. S1B) (Sakai, 2004) In these proteins, the PEPut domain transfers phosphates from one side of a protein to another. In the component EI of PTS, the mobile PEPut hosts a conserved histidyl that transfers phosphate abstracted from PEP to HPr (Teplyakov et al., 2006). By contrast, in the metabolic enzymes PPDK and PEPS, PEPut is a key part of the catalytic activity. It is responsible for turning ATP and pyruvate into

AMP, pyrophosphate and PEP and the conserved histidyl transfers the phosphate groups between the nucleotide binding domain and the substrate binding domain (Goss et al., 1980; Herzberg et al., 1996).

Interestingly, a TSH motif around the catalytic H is conserved in PEPut of EI, PEPS, PPDK and PykA (Fig. S1B) (Sakai, 2004). In PEPS and PPDK, the T residue of the motif is phosphorylated at the expense of ADP to inhibit the catalytic activity of H-P. The phosphorylation/dephosphorylation of this residue is catalyzed by serine/threonine kinases of the DUF299 family (Jim N Burnell, 2010; Jim N. Burnell & Chastain, 2006; J.N. Burnell & Hatch, 1984; Tolentino et al., 2013). In contrast, the middle S has no effect on catalytic activity or substrate specificity (Tolentino et al., 2013). However, this residue may be phosphorylated in PykA as well (Eymann et al., 2004; Mader et al., 2012; Reiland et al., 2009).

As mentioned above, PykA has been functionally connected to DNA replication gating, but it is not clear whether these results indicate a moonlighting activity in the regulation of DNA replication gating or an indirect consequence of its metabolic activity (Jannière et al., 2007; Murray & Koh, 2014; Nouri et al., 2018). The goal of this work is therefore to demonstrate that PykA indeed moonlights in the gating of DNA replication and to identify regions of PykA involved in this moonlighting activity. To achieve this, (i) the effect of mutations in the catalytic and PEPut domains of *pykA* on DNA replication timing and cell physiology were assessed in a medium where the metabolic activity of PykA is dispensable and (ii) the effect of the protein PykA on DnaE, DnaG and DnaC replication activities was analysed *in vitro*. This work (i) demonstrates that PykA moonlights in the gating of DNA replication via direct and indirect pathways and (ii) identifies regions in both the cat and PEPut domain important for this activity, assigning a concrete role to PEPut domain for the first time.

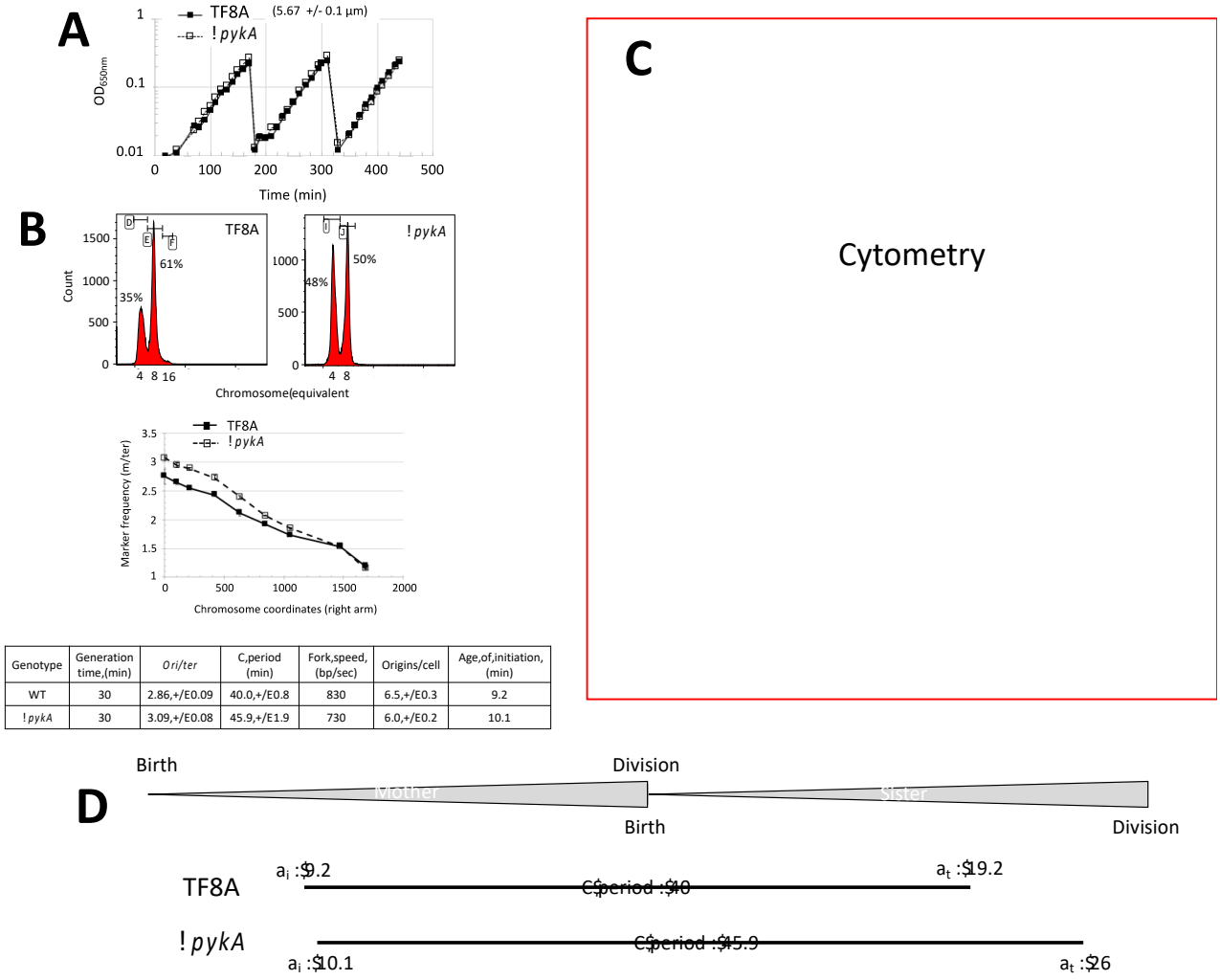
RESULTS

1. The temporal control of DNA replication is compromised in *pykA* null cells

PykA is a glycolytic enzyme that has no major function in neoglucogenic media (ref). To gain insights into its commitment in DNA replication, we thus analyzed replication in *pykA* mutants cultivated in the rich neoglucogenic medium MC (malate + casein hydrolysate). Control experiments confirm that PykA is dispensable in this medium as the size and growth rate of cells deleted for *pykA* (DGRM25, *DpykA*) are indistinguishable to that of the wild-type strain growing in the same medium (Fig. 1A). Moreover, metabolomic studies showed similar metabolite profiles in both strains in a minimal medium supplemented with malate (Fig. S2). Despite this global similarity, detailed qPCR and flow cytometry analyses showed that the mutant exhibits significant replication phenotypes in comparison to the wild-type strain: the *ori/ter* ratio is higher, the C period is extended, the speed of the replication forks is reduced, the concentration of origin

sequences is lower and the age of initiation is pushed back (Fig. 1B). As microscope studies showed no major DNA segregation defect (Fig. 1C), we compare the cell cycle of the mutant and the wild-type strain (Fig. 1D). Results clearly show that the temporal gating of replication in the cell cycle is compromised in *pykA* null cells growing in the MC medium and that this phenotype may not be due to the loss of PykA metabolic activity.

Note: the published (Nouri et al) ori/ter and C period of DpykA in mal+casa are nicely confirmed here if the latter value is corrected for the growth rate difference (previously 2.36 db/h and here 2 db/h).



Genotype	Generation time,(min)	*Ori/ter*	C,period (min)	Fork,speed, (bp/sec)	Origins/cell	Age,of,initiation, (min)
WT	30	2.86,+/E0.09	40.0,+/E0.8	830	6.5,+/E0.3	9.2
! *pykA*	30	3.09,+/E0.08	45.9,+/E1.9	730	6.0,+/E0.2	10.1

Fig. 1: The temporal gating of DNA replication is disturbed in *pykA* null cells. (A) Growth rate and size of wild-type and *DpykA* cells in MC. Cells were grown exponentially in the neoglucogenic MC medium for more than 20 generations using successive dilutions. All along the experiment, growth was assessed by spectrophotometry (OD_{650nm}, a typical experiment is shown) and cell size was analyzed at a representative timepoint by microscopy upon FM4-64 membrane staining to ease the measurement of pole-to-pole distances. (B) Cell cycle parameters analysis. The *ori/ter* ratio, C period and fork speed were determined by qPCR from cells collected at OD_{650nm} = 0.08 to 0.25 after exponential growth in MC for more than 10 generations. For determining the number of origins/cell and the age of initiation, chloramphenicol was

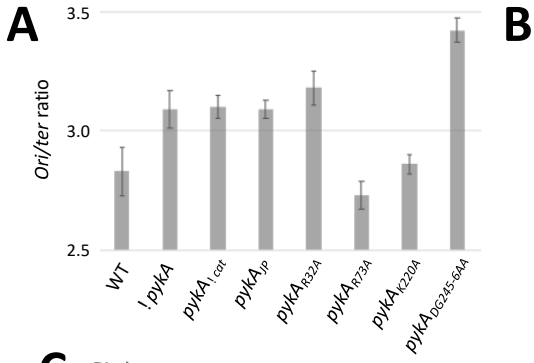
added to exponentially growing cells at OD_{650nm} = 0.15 (runout experiment). The drug inhibits replication initiation and cell division but allows completion of ongoing rounds of replication. After 4 hours of drug treatment, cells were analyzed by flow cytometry after DNA staining with Hoechst 33258. Top panel: typical runout histograms with the % of cells containing 4 or 8 chromosomes. Mid panel: typical marker frequency slot. The nearly monotonous decrease of marker frequencies from the origin to the terminus showed that there is no pause site along the chromosome and that the velocity of replication forks in the two strains are different. Bottom panel: cell cycle parameters (mean and SD are from 3 to 8 independent biological repeats in most instances). The cell cycle parameters were determined as described in Material and Method. (C) Microscope study of chromosome segregation. Cells growing exponentially in MC were incubated with FM4-64 and DAPI to stain the membrane (colored in red) and the nucleoid (colored in blue), respectively. Shown are the phase contrast and the FM4-64 and DAPI merge of a typical analysis. (D) Schematic representation of cell cycles. The triangles represent two successive cell cycles. Replication of sister cells initiated in mother cells. Numbers are in min; a_i: age of initiation; a_t: age of termination.

2. Different regions in the catalytic domain of PykA but not its catalytic activity prevent reduction in replication fork velocity and early initiation

To gain insights into how *pykA* deletion alters replication gating, we first analyzed cell cycle parameters in mutants of the catalytic domain (Cat). To that end, we used a strain deleted for Cat ($pykA_{\Delta cat}$) and mutants that affect the binding of phosphoenolpyruvate ($pykA_{R32A}$), ADP ($pykA_{R73A}$), phosphoenolpyruvate and Mg^{2+} ($pykA_{DG245\text{-}6AA}$) or the transition state stabilizer ($pykA_{K220A}$) (Fig. S1) (Schormann et al., 2019). We also used a 27 amino-acid deletion (208-234) ($pykA_{JP}$) that overlaps the $pykA_{K220A}$ mutation and that suppresses a thermosensitive mutation in the lagging strand polymerase DnaE that affects DNA elongation ($dnaE_{2.4}$) (Jannière et al., 2007; Paschalis et al., 2017). Control experiments confirmed that the Cat mutations inactivate or strongly inhibit PykA metabolic activity as the corresponding strains grow on a glycolytic medium as poorly as the strain lacking PykA (Fig. S3A). They also confirmed that disruption of the catalytic activity of PykA has no major outcome on cellular metabolism in MC as Cat mutants have a growth rate in MC and a metabolomic profile in a minimal media supplemented with malate similar to the wild-type strain (Fig. S3B and S4). Moreover, Cat mutations have no detectable effect on segregation/compaction of DNA (Fig. S3C).

Determination of cell cycle parameters showed that two Cat mutations cause no significant cell cycle alteration ($pykA_{R73A}$ and $pykA_{K220A}$) while the remaining four ($pykA_{\Delta cat}$, $pykA_{JP}$, $pykA_{R32A}$ and $pykA_{DG245\text{-}6AA}$) are associated with a high *ori/ter* ratio, an extended C period, a drop in fork velocity and, when they give nicely resolved peaks in run out experiments ($pykA_{R32A}$), an increase in the number of origins/cell and the age of initiation is brought forward (Fig. 2 & S5). As all the tested mutations abrogate or strongly reduce PykA metabolic activity, this result indicates that the commitment of PykA in replication gating depends on the protein itself rather than on its metabolic activity and/or its role in cellular metabolism. PykA/dNTP synthesis/dNTP pools by metabolomic. We infer from this that the PykA domain encoding the catalytic

activity (but not the catalytic activity on its on) is required to prevent slow replication fork and possibly early initiation in the MC medium. Three separate regions in the Cat domain appeared to be important for these activities. They encompass amino-acids R32, DG245-246 and residues 208-234 that are deleted in $PykA_{JP}$. Interestingly, the R32 and DG245-246 mutations affect the binding of phosphoenolpyruvate (Fig. S1) suggesting that proper replication gating may depend on phosphoenolpyruvate binding to PykA. Given the strong genetic link between amino-acids 208-234 in one hand and DnaC, DnaG and DnaE in the other hand, and given that this link may depend on conformational changes in replication enzymes (Jannière et al., 2007), our results further suggest that replication gating may involve conformational changes in DnaC, DnaG and/or DnaE orchestrated somehow by residues 208-234 of the Cat domain of PykA.



B

Genotype	Ori/ter	C period (min)	Fork speed (bp/sec)	Origins/cell	Age of initiation (min)
WT	2.86 +/- 0.09	40.0 +/- 0.8	830	6.5 +/- 0.3	9.2
ΔpykA	3.07 +/- 0.10	45.9 +/- 1.9	730	6.0 +/- 0.2	10.1
$pykA_{1cat}$	3.15 +/- 0.06	55.1 +/- 1.6	600	Not yet tested	
$pykA_{JP}$	3.09 +/- 0.04	51.2 +/- 0.3	650	Poor resol?	
$pykA_{R32A}$	3.18 +/- 0.07	49.0 +/- 2.0	680	6.72 +/- 0.2	5.0
$pykA_{R73A}$	2.73 +/- 0.06	41.6 +/- 1.2	800	6.48 +/- 0.2	8.9
$pykA_{K220A}$	2.86 +/- 0.04	45.1 +/- 1.0	740	Not yet tested	
$pykA_{DG245-6AA}$	3.42 +/- 0.05	54.5 +/- 1.0	610	Poor resol?	

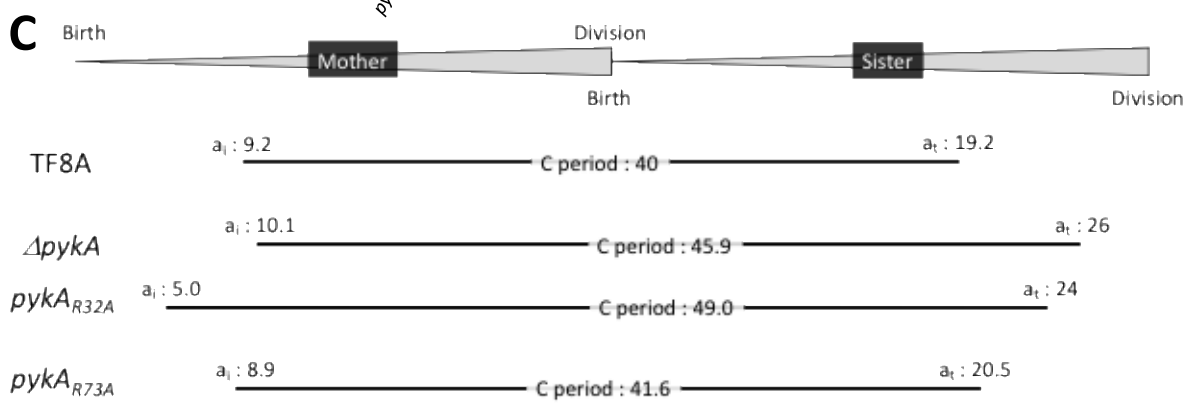


Fig.2: Distinct regions in the catalytic domain of PykA are important for proper replication gating. A. *Ori/ter* ratio analysis. B. Cell cycle parameters. C. Schematic representation of cell cycles. See Fig. 1 for details.

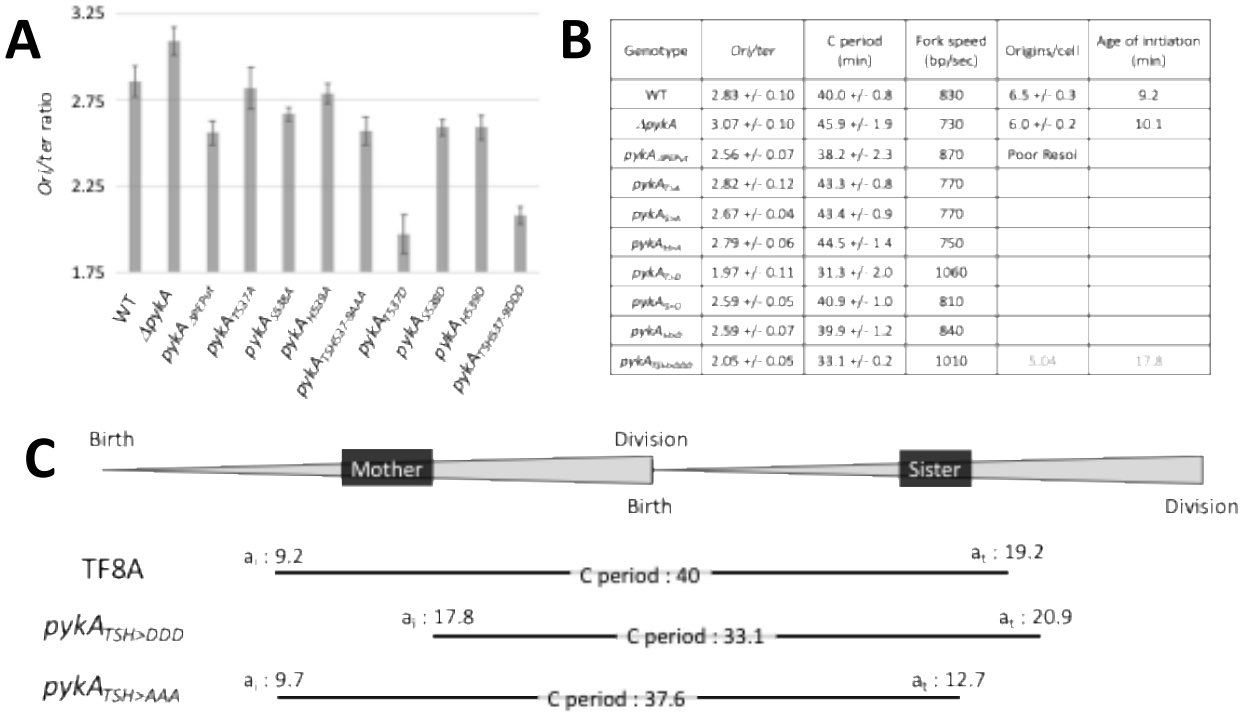
3. PEPut is committed in replication gating and phosphorylation of the T residue in the TSH motif of PEPut delays initiation and increases fork speed beyond a mere absence of PEPut

Next, cell cycle parameters were characterized in mutants of PEPut. These included a strain encoding a PykA protein deleted for PEPut ($pykA_{\Delta PEPut}$) and mutants of the TSH motifs in which individual amino-acids are replaced by A or D, the latter amino-acid mimicking phosphorylation. We also tested mutants in which the three residues of the TSH motif are replaced by A or D. Again, these mutations do not affect growth in MC and have metabolomic profiles similar to wild-type cells in a minimal media supplemented with malate (Fig.

S6A and S7). This is consistent with previously data showing that PEPut plays no major function in the *B. subtilis* pyruvate kinase metabolic activity (Sakai, 2004). Additionally, PEPut mutations have no detectable effect on DNA segregation/compaction (Fig. S6B).

The moderate *ori/ter* ratio decrease found *in* $pykA_{\Delta PEP}$ cells showed that PEPut is also important for replication (Fig. 3A). As this domain has no detectable metabolic activity (Fig. S7) (Sakai, 2004), this result confirms that PykA commitment in replication gating depends on the protein itself rather than on its metabolic activity and/or its effect on cellular metabolism. Interestingly, the $pykA_{TSH>AAA}$ mutant exhibits a similar *ori/ter* ratio decrease than $pykA_{\Delta PEP}$ cells indicating that this phenotype mostly depends on the TSH motif (Fig. 3A). As individual mutations have no ($pykA_{T>A}$ and $pykA_{H>A}$) or a marginal effect on the ratio ($pykA_{S>A}$), 2 or the 3 residues of the motif are important for proper *ori/ter* ratio in the MC medium. Surprisingly, a dramatic drop in the ratio was found in $pykA_{TSH>DDD}$ cells and the $pykA_{T>D}$ mutation is sufficient to account for this drop (Fig. 3A). The D residue mimicking amino-acid phosphorylation, this suggests that permanent phosphorylation of T in the TSH motif dramatically interferes with replication. The moderate ratio decrease observed in $pykA_{S>D}$ and $pykA_{H>D}$ cells indicates that permanent phosphorylation of the S and H residues of the TSH motif may also impact on the ratio however this effect is much less pronounced than in the $pykA_{T>D}$ mutant. Determination of the remaining cell cycle parameters showed that the $pykA_{TSH>DDD}$ mutation strongly delays initiation and dramatically stimulates replication speed while the $pykA_{TSH>AAA}$ mutation only slightly increases fork velocity (Fig. 3BC and Fig. S8). *Strong elongation phenotype/ dNTP pools by metabolomic in TSH mutants*.

Overall, results show that the Cat and PEPut domains have opposite effects on replication timing in the MC medium. Where the Cat domain prevents early initiation and slow replication forks, PEPut, and more specifically transient phosphorylation of the T residue of TSH, prevents initiation delay and fast replication forks. The S and H residues of the TSH motif are also engaged in replication.



Genotype	Ori/ter	C period (min)	Fork speed (bp/sec)	Origins/cell	Age of initiation (min)
WT	2.83 +/- 0.10	40.0 +/- 0.8	830	6.5 +/- 0.3	9.2
ΔpykA	3.07 +/- 0.10	45.9 +/- 1.9	730	6.0 +/- 0.2	10.1
pykA ΔPEPut	2.56 +/- 0.07	38.2 +/- 2.3	870	Poor Resol	
pykA T>A	2.82 +/- 0.12	43.3 +/- 0.8	770		
pykA S>A	2.67 +/- 0.04	43.4 +/- 0.9	770		
pykA H>A	2.79 +/- 0.06	44.5 +/- 1.4	750		
pykA T>D	1.97 +/- 0.11	31.3 +/- 2.0	1060		
pykA S>D	2.59 +/- 0.05	40.9 +/- 1.0	810		
pykA H>D	2.59 +/- 0.07	39.9 +/- 1.2	840		
pykA TSH>DDD	2.05 +/- 0.05	33.1 +/- 0.2	1010	5.04	17.8

Fig. 3: The PEPut sequence and phosphorylation of the T residue in the TSH motif of PEPut are important for proper replication gating. A. *Ori/ter* ratio analysis. B. Cell cycle parameters. C. Schematic representation of cell cycles. See Fig. 1 for details.

4. Free PEPut molecules have an inhibitory activity on the *ori/ter* ratio that depends on the H residue of the TSH motif, on amino-acids 208-234 of the Cat domain and on a limiting inhibitor

In order to better understand the functioning of PEPut and its interplay with Cat, this peptide was expressed from an IPTG inducible construct cloned at an ectopic position (in the *amyE* gene). The construction, termed *hyPEPut* hereafter, was introduced in wild-type cells and in various *pykA* mutants. Cells with key mutations in the TSH motif of *hyPEPut* (*hyPEPut*$_{H>A}$ and *hyPEPut*$_{T>D}$) were also constructed.

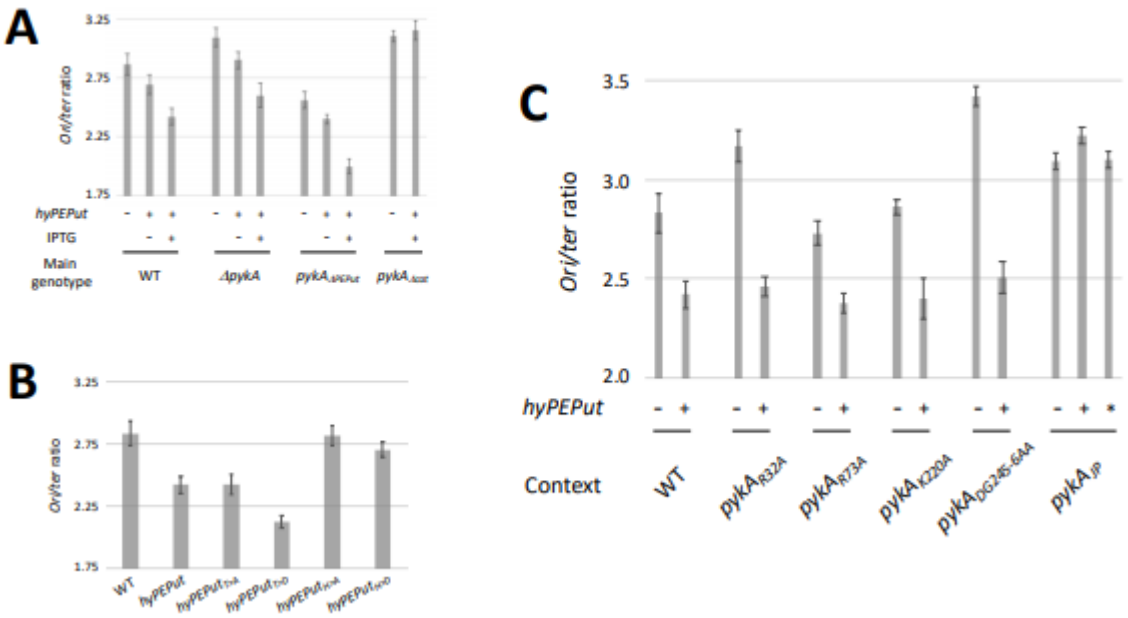
Measurements of the *ori/ter* ratio in the wild-type, *ΔpykA* and *PykA*$_{\Delta PEPut}$ contexts showed that *hyPEPut* expression is associated with a ratio decrease and that this phenotype is accentuated at high expression level (Fig. 4A). In contrast, this expression has no effect on the ratio in *pykA*$_{\Delta cat}$ cells (Fig. 4A). As the inducible promoter of *hyPEPut* drives synthesis of much less molecules per cell than the native promoter of *pykA* (about 3.000 versus about 30.000) (Eymann et al., 2004; Le Chatelier et al., 2004), this result showed that free PEPut molecules have a negative effect on the *ori/ter* ratio at moderate concentration (strains *hyPEPut*, *hyPEPut ΔpykA* and *hyPEPut pykA*$_{\Delta PEPut}$) but not at high concentration (strains *hyPEPut pykA*$_{\Delta cat}$ and *pykA*$_{\Delta cat}$) and that this effect is dominant over high concentration of Cat (strain *hyPEPut pykA*$_{\Delta PEPut}$) and high concentration of PEPut fused to Cat (WT strain). They also show that free PEPut does not complement *pykA*$_{\Delta PEPut}$ as the *ori/ter*

ratio in the *pykA*$_{\Delta PEPut}$ *hyPEPut* double mutant is lower than in the wild-type strain (Fig. 4A). These results suggest that the Cat and PEPut domain must be fused in a single macromolecule for proper functioning in replication. The opposite outcome of *hyPEPut* expression in WT (inhibition) and *pykA*$_{\Delta cat}$ (no inhibition) cells further suggests that *B. subtilis* encodes a replication factor (unidentified so far) that free but not Cat-fused PEPut molecules can sense. This factor is likely limiting as it is able to convey a signal for reducing the *ori/ter* ratio at mid but not high concentration of free PEPut molecules. We thus propose that most of the PEPut peptides receive the signal from the limiting factor at mid concentration switching on a system causing a decrease in the *ori/ter* ratio while, at high concentration, the system remains in the off status because only a weak proportion of free PEPut molecules receive the signal from the limiting factor.

Next, the effect on the *ori/ter* ratio of TSH mutations in *hyPEPut* was analyzed (Fig. 4C). This work, carried out in a WT context, showed a low ratio in *hyPEPut*$_{T>A}$ cells and an even lower value in the *hyPEPut*$_{T>D}$ mutant. However, the ratio is nearly wild-type in *hyPEPut*$_{H>A}$ and *hyPEPut*$_{H>D}$ strains. This suggests that the H residue of the TSH motif in free PEPut molecules is essential for receiving and/or conveying the signal from the limiting inhibitor to the *ori/ter* ratio. This assigns a function to this residue that was not clearly detected in PykA (see Fig. 3A). Alike in the Cat-fused context, the T>D mutation in the TSH motif enhanced the inhibitory effect of free PEPut molecules further highlighting the importance of T phosphorylation in PEPut functioning. Taken together these results suggest that the T and H residues of the TSH motif of PEPut ensure two separate functions important for replication.

We then analyzed the effect of *hyPEPut* expression in catalytic mutants of PykA (Fig. 4D). We found that, alike in the WT strain, expression of *hyPEPut* is associated with a reduction in the *ori/ter* ratio in the *pykA*$_{R32A}$, *pykA*$_{R73A}$, *pykA*$_{K220A}$ and *pykA*$_{DG245\text{-}6AA}$ contexts. However, this inhibitory activity is abrogated in the *pykA*$_{JP}$ context and this abrogation is very strong as it still operates in the *pykA*$_{JP}$ *hyPEPut*$_{T>D}$ double mutant (Fig. 4D). This established a tight functional relationship between amino-acids 208-234 of the Cat domain of PykA (these residues are deleted in PykA$_{JP}$) and the H residue of the TSH motif of PEPut.

Overall, data show that PEPut can affect replication (likely initiation) on its on (i.e. independently of its fusion to Cat) and suggest that this activity is regulated by Cat (test pykA at low concentration). Although it is unclear how PEPut affects replication, this activity somehow depends on the H and T residues of the TSH motif of PEPut, on the level of T phosphorylation in the TSH motif, on residues 208-234 of Cat and on a limiting inhibitor. Potential targets in the replication machinery of PEPut are the DnaC, DnaG and DnaE replication enzymes as these enzymes are genetically linked to the 208-234 residues of the Cat domain of PykA (Jannière et al., 2007). Interestingly, these proteins form a ternary complex and play key functions in initiation and elongation of replication (Dervyn et al., 2001; Paschalis et al., 2017; Rannou et al., 2013; Sanders et al., 2010).



5. **The purified PykA protein modulates replication activities of DnaE and DnaC through protein-protein interactions**

The results presented above uncovered metabolic determinants important for replication gating. They however do not provide clear insights on the mechanism at play. To investigate this mechanism, the effect of PykA on replication enzymes activities was investigated *in vitro*. To that end, the PykA protein was purified and controlled experiments showed that the isolated protein is metabolically active and forms the expected stable tetramer (Fig. S9). The effect of PykA was then tested on DnaC, DnaG and DnaE as these replication enzymes are genetically linked to PykA (Jannière et al., 2007). Previous studies showed that DnaC, DnaG and DnaE form a ternary complex that ensures important roles during replication initiation and elongation (Dervyn et al., 2001; Paschalis et al., 2017; Rannou et al., 2013; Sanders et al., 2010). The DnaC helicase melts the duplex DNA at *oriC* during initiation and separates the DNA strands in replication forks during elongation. The DnaG primase synthesizes RNA primers at the origin and in replication forks. These primers are used by DNA polymerases for polymerizing DNA. The DnaE polymerase extends RNA primers produced by DnaG at the origin as well as in replication forks to ensure partial or complete synthesis of the lagging strand. The three replication enzymes were purified and their activities were tested in the presence or absence of PykA as previously (Paschalis et al., 2017; Rannou et al., 2013).

Replication assays with polymerase DnaE were first conducted. These assays were carried at a DnaE concentration (10 nM) that produces low amount of replication products in order to ease the detection of

stimulatory effects. In reaction mixtures containing M13 ssDNA annealed to a 20-mer DNA oligonucleotide and DnaE in combination or not with equimolar concentrations of PykA, a substantial stimulation of DnaE polymerase activity by PykA was found in term of both product sizes and product amounts (Fig. 5A, left panel). A similar stimulation was observed with different templates (M13 ssDNA annealed to a 60mer DNA primer and a 110-mer annealed to a 15-mer primer) and at a lower DnaE concentration (2 nM) even though nearly no DNA synthesis was observed at this low DnaE concentration in the absence of PykA (not shown). The stimulation was specific to PykA as it was not observed with equimolar amount of BSA (Fig. 5A, right panel). The lack of DnaE stimulation by BSA was confirmed at a 10-fold excess concentration over DnaE (Fig. S10A). The marginal stimulation observed at very high (100-fold) BSA excess is likely artifactual. Finally, titration experiments and gel shift assays showed that the stimulation of DNA synthesis by PykA was not due to a stimulation of DnaE binding to primed template (Fig. S10B). It is inferred from these experiments that PykA stimulates DnaE polymerase activity probably via a direct interaction between these two proteins. Since the purified PEPut domain does not change DnaE activity (not shown), the stimulation probably depends on a direct contact between DnaE and the Cat domain of PykA.

The helicase activity of DnaC was assayed by monitoring the displacement of a labelled 104-mer oligonucleotide annealed onto M13 ssDNA and forming a double-fork substrate with poly(dT) tails as previously (Rannou et al., 2013). To assemble a functional DnaC hexamer onto the DNA substrate, the helicase loader DnaI was added to reaction mixtures at equimolar concentrations. Results showed that DnaC is marginally inhibited in the presence of PykA (Fig. 5B, panel (1)). As we previously showed that DnaC activity is stimulated by DnaG (Rannou et al., 2013), DnaG was added to the above-mentioned reaction mixtures at equimolar concentrations. This analysis confirmed the stimulation of helicase activity by DnaG and showed that this stimulation is cancelled by PykA. Thus, in some contexts, PykA can significantly inhibit the helicase activity of DnaC. All together, these results suggest that PykA can modulate replication enzyme activities likely through direct interactions with replication enzymes.

Possible new result sections: suppression of dnaTs mutants by PykA mutations; epifluorescence microscopy to look for colocalization of PykA or PykA domains with replication origins or replication forks; TSH phosphorylation in vivo; cell cycle parameters and artificial changes in phosphoenolpyruvate concentration…

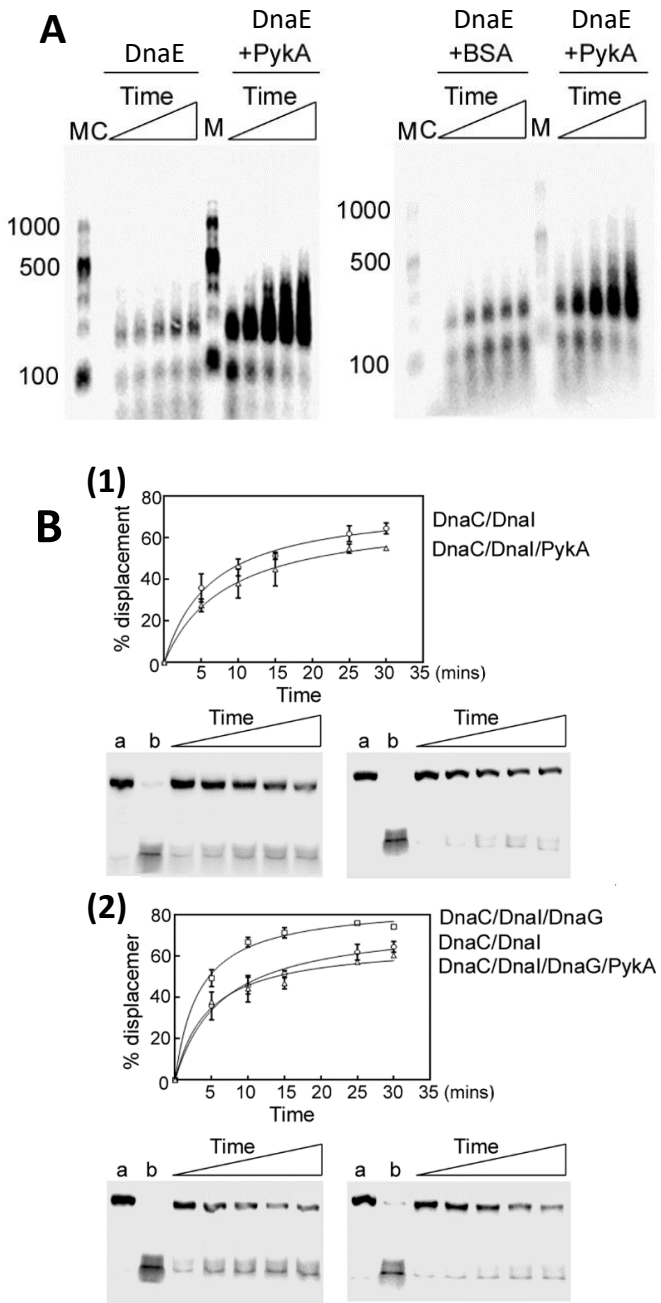


Fig. 5: PykA stimulates the DNA polymerase activity of DnaE and inhibits the helicase activity of the helicase DnaC via the primase DnaG. A. Representative alkaline agarose gels showing DnaE primer extension time courses alone and in the presence or absence of PykA or BSA. 10 nM DnaE were premixed with equimolar concentrations of pykA or BSA. Reactions were carried out as described in Material and Methods. B. Monitoring of DnaC helicase activity. The helicase activity of DnaC/DnaI (500 nM each protein) in the presence or absence of PykA (x nM) and/or DnaG (250 nM) is shown as a function of time. It was assayed by monitoring the displacement of a primer annealed to M13 ssDNA onto non-denaturing polyacrylamide gel as described in Material and Method. Data were plotted as a percentage of displaced primer versus time.

DISCUSSION

The pyruvate kinase is a highly conserved and highly abundant protein that catalyzes the last step of glycolysis by converting phosphoenolpyruvate and ADP into pyruvate and ATP. In *B. subtilis*, this enzyme is composed of a N-terminal domain that encodes the catalytic activity and a C-terminal domain that is dispensable for PykA catalytic and has no known function. Interestingly, this peptide is strongly homologous to PEP utilizers imbedded in metabolic enzyme and endowed with kinase activity (Herzberg et al., 1996; Sakai, 2004; Teplyakov et al., 2006). Different evidences suggest that PykA is important for DNA replication. First, its deletion suppress thermosensitive mutations in enzymes essential for replication initiation and elongation in LB (the DnaC helicase, the DnaG primase and the DnaE DNA polymerase) and this suppression was proposed to occur through conformational changes in replication enzymes (Jannière et al., 2007; Paschalis et al., 2017). Second, the metabolic control that mainly depends on DnaA and *oriC* in wild-type cells, was found DnaA-dependent and *oriC*-independent in cells lacking PykA and grown in LB (Murray & Koh, 2014). Third, *pykA* deletion was found to stimulate replication initiation in rich media and this phenotype was proposed to result from a modulation of the functional recruitment of DnaC, DnaG and/or DnaE to *oriC* (Nouri et al., 2018). Fourth, the *pykA* and *dnaE* genes are supposed functionally related as they are synthenic along a significant range of bacterial species (Overbeek et al., 1999).

To further investigate the link between *pykA* and replication in *B. subtilis*, we analyzed here different replication parameters in cells encoding a PykA protein mutated in the Cat or PEPut domain. This study was carried out in a gluconeogenic medium where the enzyme is strongly expressed but not required for metabolism and growth (Nicolas et al., 2012). Significant replication phenotypes were found in 13 of the 19 mutants tested (Fig. 2 and 3). These phenotypes are not due to changes in pyruvate kinase activity as (i) some catalytic mutations that disrupt the metabolic activity of PykA are not associated with replication defect and as (ii) PEPut mutations do not affect PykA metabolic activity (Sakai, 2004). In contrast, these results suggest that replication phenotypes in PykA mutants result from changes in the PykA protein on its own. It is inferred from these data that the *B. subtilis* PykA protein encodes non-metabolic function for moonlighting in DNA replication. A literature survey shows that PykA and other CCM enzymes are multifunctional proteins that moonlight in various key cellular processes all along the tree of life (Boukouris et al., 2016; Commichau & Stülke, 2007; Snaebjornsson & Schulze, 2018).

Depending on the position of the mutations in the protein, opposite replication defects are generally observed (Fig. 2 and 3). In mutants of the catalytic site, the speed of replication forks is significantly decreased (from 830 nt/s in the wild-type strain to 600 nt/s in mutants) causing a dramatic enlargement of the C period (from 40 to 55 min). In at least one of these mutants, the age of initiation is also put forward. Conversely, in PEPut mutants, the speed of the replication forks is remarkably accelerated (up to 1060 nt/s) causing an

impressive reduction of the C period (up to 31 min) and the age initiation can be strongly delayed (from 9.2 to 17.8 min in a generation time of 30 min). It is inferred from this that both the Cat and PEPut domain of PykA moonlight in replication. Given that cells encoding only the Cat or the PEPut domain have different phenotypes, the results also suggest that the moonlighting activity of each domain is, at least in part, independent of each other. Finally, as replication defects in PykA mutants change the speed of replication forks and the age of initiation, our result shows that the moonlighting activity of PykA helps the cells to temporalize DNA replication in the cell cycle.

Results uncover amino-acids or groups of amino-acids in the PykA protein important for its moonlighting activity in replication. In the Cat domain, three such elements were identified (Fig. 2). They correspond to amino-acids R32, DG245-246 and residues at position 208-234. Interestingly, residues R32 and DG245-246 are important for phosphoenolpyruvate binding (Schormann et al., 2019) (Fig. S1) suggesting that replication gating may depend on phosphoenolpyruvate binding to PykA. Moreover, the moonlighting activity of Cat may operate through replication enzymes DnaG, DnaC and/or DnaE as residues 208-234 are genetically linked to these enzymes and may impact their conformation (Jannière et al., 2007). The *in vitro* modulation of replication activity of DnaC and DnaE by Cat (Fig. 5) strongly enforce this hypothesis and suggests that the temporalization of replication involves direct physical interactions between Cat and DnaC, DnaG and/or DnaE. Test colocalization Cat- replication machinery in vivo.

Point mutations in the PEPut domain clearly showed that the TSH motif play a key role in the moonlighting activity of PykA (Fig. 3). As these residues have been reported in the literature to be phosphorylated and as the phosphorylated H residue can transfer its phosphoryl group to a target protein (Burnell and Hatch, 1984; Herzberg et al., 1996; Mäder et al., 2012; Teplyakov et al., 2006; Tolentino et al., 2013), this result strongly argues that the moonlighting activity of PEPut may involve phosphorylation events. This hypothesis is supported by results showing that mutations mimicking phosphorylation at T and, to a lesser extent, S and H residues affects replication (Fig. 3). Alike in the PTS system, the H residue may be phosphorylated at the expense of phosphoenolpyruvate. As PykA binds this metabolite to catalyze the conversion of phosphoenolpyruvate to pyruvate, and as PEPut peptides are mobile, H phosphorylation may occur through an interaction between the PEPut and Cat domain of PykA. This hypothesis is supported by data suggesting that the binding of phosphoenolpyruvate to Cat is important for PykA moonlighting activity. However, despite our efforts, we were unable to detect autophosphorylation of PykA in the presence phosphoenolpyruvate *in vitro* (not shown). We speculate that something was missing in our reaction mixtures or that phosphorylation of the H residue of PEPut is carried out through *trans* interaction with other metabolic enzymes containing PEPut, namely the pyruvate phosphate dikinase and the PEP synthase. So far, we do not known how the T and S residues of PEPut may be phosphorylated. A survey of the literature

however suggests that this phosphorylation may be carried out by one or another serine-threonine kinases encoded by *B. subtilis* (Cousin et al., 2013).

Additional determinants of the moonlighting activity of PykA have been uncovered when cells express free PEPut molecules from an ectopic locus at a moderate concentration (about 3.000 molecules/cell in comparison to about 30.000 molecules/cell for PykA; the construction is termed *hyPEPut* hereafter) (Fig. 4). First, we found that the production of free PEPut molecules from *hyPEPut* is associated with an *ori/ter* ratio decrease in *ΔpykA* context confirming that PEPut can express a moonlighting activity in replication independently of Cat. Interestingly, the ratio decrease observed at moderate concentrations of free PEPut molecules in cells lacking *pykA* is still expressed in cells producing PykA (the Cat-PEPut fusion) at high concentration, but not in cells producing high concentration of free PEPut molecules from the natural expression signals of *pykA*. This indicates that free PEPut molecules are dominant over PEPut molecules fused to Cat and that the activity of PEPut is modulated by its fusion to Cat. The contrasted response of the *ori/ter* ratio to the concentration of free PEPut molecules suggests that *B. subtilis* encodes a replication factor (unidentified so far) that free but not Cat-fused PEPut molecules can sense and that this factor is limiting as it is able to convey a signal for reducing the *ori/ter* ratio at mid but not high concentration of free PEPut molecules. We thus propose that most of the PEPut peptides receive the signal from the limiting factor at mid concentration, switching on a system causing a decrease in the *ori/ter* ratio while, at high concentration, the system remains in the off status because only a weak proportion of free PEPut molecules receive the signal from the limiting factor. The identification of the limiting factor is awaited to determine whether it can modulate the moonlighting activity of PEPut when fused to Cat.

The effect of moderate concentration of wild-type or mutated free PEPut molecules in various genetic contexts was also investigated. While, the negative effect of free PEPut on the *ori/ter* was found in most of the tested contexts, this phenotype was abolished in cells encoding a PykA protein deleted for amino-acids 208-234 ($PykA_{JP}$) and when PEPut contains the H>A or H>D mutation (Fig. 4). Hence, the inhibitory activity of PEPut at moderate concentration on the *ori/ter* ratio depends on residues 208-234 of the Cat domain and on the H residue of the TSH motif in addition to the limiting factor. This functional link may also include DnaC, DnaG and DnaE as residues 208-234 of Cat are genetically connected to these replication enzymes and as the Cat domain of PykA modulates *in vitro* the activity of DnaE and DnaC likely through direct protein-protein interactions (Fig. 5). Additionally, the moonlighting activity of PykA may also depend on the binding of phosphoenolpyruvate to Cat as mutations known to disrupt this interaction affect PykA activity on replication and as the H residue in metabolic enzymes containing a PEP utilizer is phosphorylated at the expense of phosphoenolpyruvate. The phosphorylation of the T residue of the TSH motif may also contributes to PykA moonlighting activity as the T>D mutation has dramatic effect on replication (Fig. 3 and 4). Interestingly, results suggest that the T and H residues operate independently indicating that the PEPut is the site of more

than one signaling system. Further studies are awaited to better understand how the determinants uncovered here operate for conferring a moonlighting activity to PykA in DNA replication.

Results reported here show that PykA in *B. subtilis* moonlights in replication to help the temporal control of replication in the cell cycle. This may assign a function to PykA in genetic stability as an impressive core of data shows that defects in replication control increase the risk of double-stranded breaks which are chromosomal lesions with a high mutagenic potential (Michel et al., 2001; Rodgers & McVey, 2016). In cells with aberrant timing of initiation, re-initiation of replication may increase the risk of double-stranded breaks by collision between forks or by forks encountering gaps in the DNA template. Such a risk may also increase in cells with a lengthened C period in consequence of a higher probability of fork breakage over time and when replication forks are fast by reducing the time for error corrections. The potential involvement of CCM mutations in genetic stability through their effect on the temporal control of replication may be important in the field of cancer. Indeed, two hallmarks of cancer cells are aerobic glycolysis and genetic instability (Gatenby & Gillies, 2004; Loeb et al., 2003; Tubbs & Nussenzweig, 2017). As early as 1956, Otto Warburg has postulated that aerobic glycolysis is the root cause of cancer (Warburg, 1956). However, more recent hypotheses suggest that mutations in oncogenes and tumor suppressor genes are responsible for malignant transformation, and that aerobic glycolysis is a result of these mutations rather than their cause (Macheret & Halazonetis, 2015). Although the Warburg hypothesis (i.e. cancer results from impaired mitochondrial metabolism) has proven incorrect (aerobic glycolysis results from stimulation of glycolysis and acetate production), the hypothesis that changes in metabolism pave the way to mutations and tumorigenesis has been rekindled by a series of studies from bacteria to human cells. By assigning a moonlighting activity in replication control to CCM enzymes, this work suggests that CCM mutations can increase genetic instability and pave the way to tumorigenesis by altering replication control.

MATERIAL AND METHODS

Strains and plasmids

Strains and plasmids are listed in Supplementary Table S1. Table S2 presents primers used for strain constructions. In order to ease the construction of *pykA* mutants, we first inserted the *tet* gene of plasmid pTB19 few bp downstream of *pykA*. To that goal, we amplified the *tet* gene and two chromosomal sequences flanking the site of insertion. Next, we separated the PCR products from the parental genomic DNA by gel electrophoresis and the PCR fragments were gel purified using the Monarch DNA Gel Extraction Kit (New England Biolabs, Evry, Fr). The purified fragments were then mixed and fused together using assembly PCR. This reaction depends on 20 bp sequences that were added to insertion site primers and that are homologous to the 5' and 3' tails of the *tet* fragment. Competent cells of the wild-type strain cured of prophages TF8A were then transformed with the PCR product and double cross-over integrating events were selected on Tet

plates after induction of Tet resistance (before plating, cells were incubated 1h at 37°C in the presence of 2.5 µg/mL Tet). A representative transformant was selected by sequencing and named DGRM295. To generate *pykA* mutants, two PCR reactions were performed on strain DGRM295 (*pykA-tet*). In each reaction, one external primer and one mutagenic specific primer were used to generate PCR products with the desired *pykA* mutation at one end. Upon gel purification, the products were fused by assembly PCR and used to transformed TF8A competent cells as above (however, transformants thought to encode an inactive pyruvate kinase were selected on LB + Tet + Malate 0,4%). The DGRM1009 strain containing the *Lactococcus lactis pykA* open reading frame in place off and downstream from the expression signals of the native *pykA* gene was constructed using the same principal as for generating the *pykA-tet* (DGRM295) strain. First, we generated 3 PCR fragments: one contained sequences upstream of and ending at the start codon of *pykA*; another one contained the *tet* gene and downstream *B. subtilis* sequences (they were both generated using as template the DNA of the *pykA-tet* strain); and the last one corresponded to the *L. lactis pykA* open reading frame (it was generated using as template the *L. lactis* genome and primers having short homologies with the flanking PCR products). The fragments were then fused and the combined PCR product was used to transform TF8A competent cells as above. The DGRM1010 strain that contains the *pykA L. lactis* fused to the PEPut of PykA was constructed in a similar way. However, in that case the *tet* PCR fragment contained 5' to the *tet* gene the sequences of the PEPut domain. The genotype of transformants was verified by PCR and DNA sequencing. A slow growth on the rich LB medium which confers a glycolytic regimen confirmed that mutations in the catalytic domain of PykA (R32A, R73A, K220A and DG245-246AA) inactivate (or strongly reduce) its metabolic activity.

The GM2840 strain (provided by D. Le Coq, INRA, Jouy en Josas) contains at the *amyE* locus a DNA sequence called here *HyPEPut* that encodes (i) the PEPut domain of PykA under the control of the IPTG inducible *Hyspank* promoter and (ii) a CmR marker. The genomic DNA of this strain was used to construct strain DGRM722. DGRM722 derivatives with mutations in *HyPEPut* were generated as *pykA* mutants. For generating double mutants, competent cells of *pykA* mutants were transformed with the genomic DNA of *HyPEP* strains. The genotype of CmR transformants were verified by PCR and DNA sequencing.

Growth conditions

Routinely, *B. subtilis* cells were grown at 37°C in LB supplement or not with antibiotics at the following concentrations: spectinomycin (Sp, 60 µg/mL); tetracycline (Tet, 5 µg/ml); chloramphenicol (Cm, 5 µg/mL); phleomycin (Phl, 10 µg/mL). When cells are deficient in pyruvate kinase activity, 0.4 % malate was added to LB. For replication analysis, cells were grown in MC, a minimal medium (K_2HPO_4: 80 mM; KH_2PO_4: 44 mM; $(NH_4)_2SO_4$: 15 mM; $C_6H_5Na_3O_7$ $2H_20$: 3, 4 mM; $CaCl_2$: 50 mM; $MgSO_4$: 2 mM; FeIII citrate: 11 µg/mL; $MnCl_2$: 10

µM; $FeSO_4$: 1 µM; $FeCl_3$:4 µg/mL; Trp 50 µg/mL) supplemented with casein hydrolysate 0.2%, malate 0.4% and tryptophan 0.01%. Other genetic and molecular biology procedures were standards.

Quantitative PCR

Cells were grown overnight at 30°C in MC medium supplemented with antibiotic. For accurate measurement, ≥ 4 cultures were inoculated for *ori/ter* ratio analysis and 3 for marker frequency analysis. With a very few exceptions, three independent constructs were used for the cultures. Saturated overnight cultures were then diluted 1000-fold in MC without antibiotic and growth at 37°C was carefully monitored using spectrophotometry. Samples for qPCR analysis were collected at low cell concentration (OD600 nm 0.06 to 0.25) to ensure that cell cycle parameters are determined in steady-state cells and are not affected by the approach to the stationary phase or by changes in medium composition. The genomic DNA was extracted as previously (Nouri et al., 2018) or using the PureLink Genomic DNA mini kit (Invitrogen by Thermo Fisher Scientific, Courtaboeuf, Fr). Every qPCR reaction was carried out using two technical repeats of 4 serial dilutions. For *ori/ter* ratio monitoring, a non-replicating control DNA (stage II sporlets[230]) was analyzed simultaneously with the samples in about 1/4 of the qPCR plates. Reactions of qPCR for *ori/ter* ratio and C period measurements were carried out as previously (Nouri et al., 2018). The C period was extracted from marker frequency data and the slope of the straight line of the plot $(\log_2 Nm)\,\tau$ against *m* where *Nm* is the relative copy number of a sequence *N* at position *m* (with $0 < m < 1$ along the *ori-ter* axis) and τ the generation time in minutes (Zheng et al., 2016). The mean fork speed was calculated using the C period (min) and the actual size of the TF8A genome (4,000,631 bp; this genome had been reduced compared to the reference 168 strain by deleting prophages SPβ, PBSX and *skin*).

Flow cytometry analysis

Strains were grown as indicated in the previous section and at OD_{600nm} = 0.1–0.15, chloramphenicol (200 µg/mL final) was added to the cultures to impede replication initiation, cell division and allow completion of ongoing rounds of replication (Séror et al., 1994). After 4 hours of drug treatment, 10^8 cells were fixed in filtered ethanol 70% and stored at 4°C. Stored cells were then washed in filtered 0.1M of phosphate buffer (PB) pH 9, resuspended in 1mL of Tris-buffered saline buffer (TBS 150, filtered) (20mM Tris-HCl pH 7.5, 150mM NaCl) and stained with Hoechst 33258 (1.5 µg/mL) for at least 30 min as described elsewhere (Morigen et al., 2009). Flow cytometry analysis was performed using a MoFlow Astrios cell sorter (Beckman Coulter, Life Sciences) equipped with a 355 nm krypton laser and a 448/59 nm bandpass filter used to collect Hoechst 33258 fluorescence. Data were analyzed with the Kaluza software (Beckman Coulter, Life Sciences). We counted 50,000 to 100,000 events. In most of the tested samples, DNA histograms showed 2 main peaks

with a 2^n distribution. The number of origins/cell was obtained from the proportion of cells with 1, 2, 4, 8 and 16 chromosomes. The age of initiation was obtained from the fraction F of non-initiated cells (left peak in DNA histograms) using the formula $F = 2 - 2^{((\tau - ai)/\tau)}$ were τ is the generation time expressed in minutes (Morigen et al., 2009).

Microscopy analysis

In order to analyze cell size and nucleoid compaction and distribution, cells grown as above were collected at OD_{650nm} = 0.08-0.1. The nucleoid and membrane of 1mL of cells were then stained with X µL of 4',6'-diamidino-2-phenylindole (DAPI, Sigma) 10 mg/ mL and X µL of XXXX (FM4-64, Sigma) XXX µg/mL. Stained cells were then centrifuged gently and resuspended in 10 µL and mounted on glass slides (VWR) covered with 1.2 % agarose (low melting point, Sigma) in MC and a 0.17 mm glass coverslip (VWR) was placed on top. Images were acquired on an epifluorescence microscope (Zeiss, Axio Observer.Z1) with a 100x magnification oil-immersion objective (Plan-APOCHROMAT Ph3) and a CMOS camera (Orca Flash 4.0 LT Hamamatsu). The images were analyzed using the Zen software (ref) and FiJi (Ref). Cell lengths and DAPI fluorescence were exported to Excel for statistical analysis. All experiments were independently performed at least three times and representative pictures are shown.

Biochemistry

Replication enzymes DnaC, DnaG and DnaE were purified an tested as previously (Paschalis et al., 2017; Rannou et al., 2013). PykA expression, purification and biochemistry.

Metabolomic by mass spectrometry

REFERENCES

Barańska, S., Glinkowska, M., Herman-Antosiewicz, A., Maciąg-Dorszyńska, M., Nowicki, D., Szalewska-Pałasz, A., Węgrzyn, A., & Węgrzyn, G. (2013). Replicating DNA by cell factories: Roles of central carbon metabolism and transcription in the control of DNA replication in microbes, and implications for understanding this process in human cells. *Microbial Cell Factories*, *12*(1), 55. https://doi.org/10.1186/1475-2859-12-55

Baxi, M. D., & Vishwanatha, J. K. (1995). Uracil DNA-glycosylase/glyceraldehyde-3-phosphate dehydrogenase is an Ap4A binding protein. *Biochemistry*, *34*(30), 9700–9707. https://doi.org/10.1021/bi00030a007

Bipatnath, M., Dennis, P. P., & Bremer, H. (1998). Initiation and Velocity of Chromosome Replication in Escherichia coli B/r and K-12. *J. BACTERIOL.*, *180*, 9.

Boukouris, A. E., Zervopoulos, S. D., & Michelakis, E. D. (2016). Metabolic Enzymes Moonlighting in the Nucleus: Metabolic Regulation of Gene Transcription. *Trends in Biochemical Sciences*, *41*(8), 712–730. https://doi.org/10.1016/j.tibs.2016.05.013

Boye, E., & Nordström, K. (2003). Coupling the cell cycle to cell growth: A look at the parameters

that regulate cell-cycle events. *EMBO Reports*, *4*(8), 757–760. https://doi.org/10.1038/sj.embor.embor895
Buchakjian, M. R., & Kornbluth, S. (2010). The engine driving the ship: Metabolic steering of cell proliferation and death. *Nature Reviews Molecular Cell Biology*, *11*(10), 715–727. https://doi.org/10.1038/nrm2972
Burnell, Jim N. (2010). Cloning and characterization of Escherichia coli DUF299: A bifunctional ADP-dependent kinase - Pi-dependent pyrophosphorylase from bacteria. *BMC Biochemistry*, *11*(1), 1. https://doi.org/10.1186/1471-2091-11-1
Burnell, Jim N., & Chastain, C. J. (2006). Cloning and expression of maize-leaf pyruvate, Pi dikinase regulatory protein gene. *Biochemical and Biophysical Research Communications*, *345*(2), 675–680. https://doi.org/10.1016/j.bbrc.2006.04.150
Burnell, J.N., & Hatch, M. D. (1984). Regulation of C4 photosynthesis: Identification of a catalytically important histidine residue and its role in the regulation of pyruvate,Pi dikinase. *Archives of Biochemistry and Biophysics*, *231*(1), 175–182. https://doi.org/10.1016/0003-9861(84)90375-8
Burnetti, A. J., Aydin, M., & Buchler, N. E. (2016). Cell cycle Start is coupled to entry into the yeast metabolic cycle across diverse strains and growth rates. *Molecular Biology of the Cell*, *27*(1), 64–74. https://doi.org/10.1091/mbc.E15-07-0454
Chang, V. K., Donato, J. J., Chan, C. S., & Tye, B. K. (2004). Mcm1 Promotes Replication Initiation by Binding Specific Elements at Replication Origins. *Molecular and Cellular Biology*, *24*(14), 6514–6524. https://doi.org/10.1128/MCB.24.14.6514-6524.2004
Chang, Victoria K., Fitch, M. J., Donato, J. J., Christensen, T. W., Merchant, A. M., & Tye, B. K. (2003). Mcm1 Binds Replication Origins. *Journal of Biological Chemistry*, *278*(8), 6093–6100. https://doi.org/10.1074/jbc.M209827200
Chen, Y., & Tye, B. K. (1995). The yeast Mcm1 protein is regulated posttranscriptionally by the flux of glycolysis. *Molecular and Cellular Biology*, *15*(8), 4631–4639. https://doi.org/10.1128/MCB.15.8.4631
Cheng, X., Friesen, R. H. E., & Lee, J. C. (1996). Effects of Conserved Residues on the Regulation of Rabbit Muscle Pyruvate Kinase. *Journal of Biological Chemistry*, *271*(11), 6313–6321. https://doi.org/10.1074/jbc.271.11.6313
Chien, A.-C., Hill, N. S., & Levin, P. A. (2012). Cell Size Control in Bacteria. *Current Biology*, *22*(9), R340–R349. https://doi.org/10.1016/j.cub.2012.02.032
Commichau, F. M., & Stülke, J. (2007). Trigger enzymes: Bifunctional proteins active in metabolism and in controlling gene expression: Trigger enzymes in transcription regulation. *Molecular Microbiology*, *67*(4), 692–702. https://doi.org/10.1111/j.1365-2958.2007.06071.x
Cousin, C., Derouiche, A., Shi, L., Pagot, Y., Poncet, S., & Mijakovic, I. (2013). Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation. *FEMS Microbiology Letters*, *346*(1), 11–19. https://doi.org/10.1111/1574-6968.12189
Deutscher, J., Aké, F. M. D., Derkaoui, M., Zébré, A. C., Cao, T. N., Bouraoui, H., Kentache, T., Mokhtari, A., Milohanic, E., & Joyet, P. (2014). The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiology and Molecular Biology Reviews*, *78*(2), 231–256. https://doi.org/10.1128/MMBR.00001-14
Dickinson, J. R., & Williams, A. S. (1987). The cdc30 Mutation in Saccharomyces cerevisiae Results in a Temperature-sensitive Isoenzyme of Phosphoglucose Isomerase. *Microbiology*, *133*(1), 135–140. https://doi.org/10.1099/00221287-133-1-135
Ewald, J. C. (2018). How yeast coordinates metabolism, growth and division. *Current Opinion in Microbiology*, *45*, 1–7. https://doi.org/10.1016/j.mib.2017.12.012
Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam, L. T., Büttner, K., Buurman, G., Scharf, C., Venz, S., Völker, U., & Hecker, M. (2004). A comprehensive proteome map of growingBacillus subtilis cells. *PROTEOMICS*, *4*(10), 2849–

2876. https://doi.org/10.1002/pmic.200400907
Flåtten, I., Fossum-Raunehaug, S., Taipale, R., Martinsen, S., & Skarstad, K. (2015). The DnaA Protein Is Not the Limiting Factor for Initiation of Replication in Escherichia coli. *PLOS Genetics*, *11*(6), e1005276. https://doi.org/10.1371/journal.pgen.1005276
Fornalewicz, K., Wieczorek, A., Węgrzyn, G., & Łyżeń, R. (2017). Silencing of the pentose phosphate pathway genes influences DNA replication in human fibroblasts. *Gene*, *635*, 33–38. https://doi.org/10.1016/j.gene.2017.09.005
Gatenby, R. A., & Gillies, R. J. (2004). Why do cancers have high aerobic glycolysis? *Nature Reviews Cancer*, *4*(11), 891–899. https://doi.org/10.1038/nrc1478
Görke, B., & Stülke, J. (2008). Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients. *Nature Reviews Microbiology*, *6*(8), 613–624. https://doi.org/10.1038/nrmicro1932
Goss, N. H., Evans, C. T., & Wood, H. G. (1980). Pyruvate phosphate dikinase: Sequence of the histidyl peptide, the pyrophosphoryl and phosphoryl carrier. *Biochemistry*, *19*(25), 5805–5809. https://doi.org/10.1021/bi00566a022
Grosse, F., Nasheuer, H.-P., Scholtissek, S., & Schomburg, U. (1986). Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-alpha-primase complex. *European Journal of Biochemistry*, *160*(3), 459–467. https://doi.org/10.1111/j.1432-1033.1986.tb10062.x
Hartwell, L. H. (1973). GENETIC CONTROL OF THE CELL DIVISION CYCLE I N YEAST: *Genetics*, *74*(2), 267–286.
Helmstetter, C.E. (1996). *Timing of synthetic activities in the cell cycle*. In *Escherichia Coli* And *Salmonella*, F.C. Neidhart, R.I. Curtis, E.C. Ingraham, K.B. Lin, and L.E.C.C., eds. (Washington DC: ASM Press), pp. 1591–1605.
Hernandez, J., & Bremer, H. (1993). Characterization of RNA and DNA Synthesis in Escherichia coli Strains Devoid of ppGpp. *Journal of Biological Chemistry*, *268*(15), 10851–19862.
Herzberg, O., Chen, C. C., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., & Dunaway-Mariano, D. (1996). Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proceedings of the National Academy of Sciences*, *93*(7), 2652–2657. https://doi.org/10.1073/pnas.93.7.2652
Huang, T.-S., & Nagy, P. D. (2011). Direct Inhibition of Tombusvirus Plus-Strand RNA Synthesis by a Dominant Negative Mutant of a Host Metabolic Enzyme, Glyceraldehyde-3-Phosphate Dehydrogenase, in Yeast and Plants. *Journal of Virology*, *85*(17), 9090–9102. https://doi.org/10.1128/JVI.00666-11
Jameson, K., & Wilkinson, A. (2017). Control of Initiation of DNA Replication in Bacillus subtilis and Escherichia coli. *Genes*, *8*(1), 22. https://doi.org/10.3390/genes8010022
Jannière, L., Canceill, D., Suski, C., Kanga, S., Dalmais, B., Lestini, R., Monnier, A.-F., Chapuis, J., Bolotin, A., Titok, M., Chatelier, E. L., & Ehrlich, S. D. (2007). Genetic Evidence for a Link Between Glycolysis and DNA Replication. *PLoS ONE*, *2*(5), e447. https://doi.org/10.1371/journal.pone.0000447
Jindal, H. K., & Vishwanatha, K. (1990). Functional Identity of a Primer Recognition Protein as Phosphoglycerate Kinase. *Journal of Biological Chemistry*, *265*(12), 6540–6543.
Kaplan, Y., & Kupiec, M. (2007). A role for the yeast cell cycle/splicing factor Cdc40 in the G1/S transition. *Current Genetics*, *51*(2), 123–140. https://doi.org/10.1007/s00294-006-0113-y
Kim, B. H., & Gadd, G. M. (2008). *Bacterial Physiology and Metabolism*. Cambridge University Press.
Klevecz, R. R., Bolen, J., Forrest, G., & Murray, D. B. (2004). A genomewide oscillation in transcription gates DNA replication and cell cycle. *Proceedings of the National Academy of Sciences*, *101*(5), 1200–1205. https://doi.org/10.1073/pnas.0306490101
Konieczna, A., Szczepańska, A., Sawiuk, K., Węgrzyn, G., & Łyżeń, R. (2015). Effects of partial silencing of genes coding for enzymes involved in glycolysis and tricarboxylic acid cycle on

the enterance of human fibroblasts to the S phase. *BMC Cell Biology*, *16*(1). https://doi.org/10.1186/s12860-015-0062-8

Kumble, K. D., Iversen, P. L., & Vishwanatha, J. K. (1992). The role of primer recognition proteins in DNA replication: Inhibition of cellular proliferation by antisense oligodeoxyribonucleotides. *Journal of Cell Science*, *101*, 35–41.

Larsen, T. M., Benning, M. M., Rayment, I., & Reed, G. H. (1998). Structure of the Bis(Mg $^{2+}$)−ATP−Oxalate Complex of the Rabbit Muscle Pyruvate Kinase at 2.1 Å Resolution: ATP Binding over a Barrel †,‡. *Biochemistry*, *37*(18), 6247–6255. https://doi.org/10.1021/bi980243s

Li, X., Qian, X., Jiang, H., Xia, Y., Zheng, Y., Li, J., Huang, B.-J., Fang, J., Qian, C.-N., Jiang, T., Zeng, Y.-X., & Lu, Z. (2018). Nuclear PGK1 Alleviates ADP-Dependent Inhibition of CDC7 to Promote DNA Replication. *Molecular Cell*, *72*(4), 650-660.e8. https://doi.org/10.1016/j.molcel.2018.09.007

Loeb, L. A., Loeb, K. R., & Anderson, J. P. (2003). Multiple mutations and cancer. *Proceedings of the National Academy of Sciences*, *100*(3), 776–781. https://doi.org/10.1073/pnas.0334858100

Macheret, M., & Halazonetis, T. D. (2015). DNA Replication Stress as a Hallmark of Cancer. *Annual Review of Pathology: Mechanisms of Disease*, *10*(1), 425–448. https://doi.org/10.1146/annurev-pathol-012414-040424

Maciąg, M., Nowicki, D., Janniere, L., Szalewska-Pałasz, A., & Węgrzyn, G. (2011). Genetic response to metabolic fluctuations: Correlation between central carbon metabolism and DNA replication in Escherichia coli. *Microbial Cell Factories*, *10*(1), 19. https://doi.org/10.1186/1475-2859-10-19

Maciąg, M., Nowicki, D., Szalewska-Pałasz, A., & Węgrzyn, G. (2012). Central carbon metabolism influences fidelity of DNA replication in Escherichia coli. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *731*(1–2), 99–106. https://doi.org/10.1016/j.mrfmmm.2011.12.005

Mader, U., Schmeisky, A. G., Florez, L. A., & Stulke, J. (2012). SubtiWiki—A comprehensive community resource for the model organism Bacillus subtilis. *Nucleic Acids Research*, *40*(D1), D1278–D1287. https://doi.org/10.1093/nar/gkr923

Mathews, C. K. (2015). Deoxyribonucleotide metabolism, mutagenesis and cancer. *Nature Reviews Cancer*, *15*(9), 528–539. https://doi.org/10.1038/nrc3981

Mattevi, A., Valentini, G., Rizzi, M., Speranza, M. L., Bolognesi, M., & Coda, A. (1995). Crystal structure of Escherichia coli pyruvate kinase type I: Molecular basis of the allosteric transition. *Structure*, *3*(7), 729–741. https://doi.org/10.1016/S0969-2126(01)00207-6

Michel, B., Flores, M.-J., Viguera, E., Grompone, G., Seigneur, M., & Bidnenko, V. (2001). Rescue of arrested replication forks by homologous recombination. *Proceedings of the National Academy of Sciences*, *98*(15), 8181–8188. https://doi.org/10.1073/pnas.111008798

Monahan, L. G., Hajduk, I. V., Blaber, S. P., Charles, I. G., & Harry, E. J. (2014). Coordinating Bacterial Cell Division with Nutrient Availability: A Role for Glycolysis. *MBio*, *5*(3). https://doi.org/10.1128/mBio.00935-14

Muirhead, H., Clayden, D. A., Barford, D., Lorimer, C. G., Fothergill-Gilmore, L. A., Schiltz, E., & Schmitt, W. (1986). The structure of cat muscle pyruvate kinase. *The EMBO Journal*, *5*(3), 475–481. https://doi.org/10.1002/j.1460-2075.1986.tb04236.x

Murray, H., & Koh, A. (2014). Multiple Regulatory Systems Coordinate DNA Replication with Cell Growth in Bacillus subtilis. *PLoS Genetics*, *10*(10), e1004731. https://doi.org/10.1371/journal.pgen.1004731

Nguyen, C. C., & Saier, M. H. (1995). Phylogenetic analysis of the putative phosphorylation domain in the pyruvate kinase of Bacillus stearothermophilus. *Research in Microbiology*, *146*(9), 713–719. https://doi.org/10.1016/0923-2508(96)81067-9

Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., Becher, D., Bisicchia, P., Botella, E., Delumeau,

O., Doherty, G., Denham, E. L., Fogg, M. J., Fromion, V., Goelzer, A., … Noirot, P. (2012). Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in Bacillus subtilis. *Science*, *335*(6072), 1103–1106. https://doi.org/10.1126/science.1206848

Nouri, H., Monnier, A.-F., Fossum-Raunehaug, S., Maciąg-Dorszyńska, M., Cabin-Flaman, A., Képès, F., Węgrzyn, G., Szalewska-Pałasz, A., Norris, V., Skarstad, K., & Janniere, L. (2018). Multiple links connect central carbon metabolism to DNA replication initiation and elongation in *Bacillus subtilis*. *DNA Research*, *25*(6), 641–653. https://doi.org/10.1093/dnares/dsy031

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D., & Maltsev, N. (1999). The use of gene clusters to infer functional coupling. *Proceedings of the National Academy of Sciences*, *96*(6), 2896–2901. https://doi.org/10.1073/pnas.96.6.2896

Papagiannakis, A., Niebel, B., Wit, E. C., & Heinemann, M. (2017). Autonomous Metabolic Oscillations Robustly Gate the Early and Late Cell Cycle. *Molecular Cell*, *65*(2), 285–295. https://doi.org/10.1016/j.molcel.2016.11.018

Paschalis, V., Le Chatelier, E., Green, M., Képès, F., Soultanas, P., & Janniere, L. (2017). Interactions of the *Bacillus subtilis* DnaE polymerase with replisomal proteins modulate its activity and fidelity. *Open Biology*, *7*(9), 170146. https://doi.org/10.1098/rsob.170146

Popanda, O., Fox, G., & Thielmann, H. W. (1998). Modulation of DNA polymerases a , d and ´ by lactate dehydrogenase and 3-phosphoglycerate kinase. *Biochimica et Biophysica Acta*, *1397*, 102–117.

Prasanth, K. R., Huang, Y.-W., Liou, M.-R., Wang, R. Y.-L., Hu, C.-C., Tsai, C.-H., Meng, M., Lin, N.-S., & Hsu, Y.-H. (2011). Glyceraldehyde 3-Phosphate Dehydrogenase Negatively Regulates the Replication of Bamboo Mosaic Virus and Its Associated Satellite RNA. *Journal of Virology*, *85*(17), 8829–8840. https://doi.org/10.1128/JVI.00556-11

Prasanth, K. Reddisiva, Chuang, C., & Nagy, P. D. (2017). Co-opting ATP-generating glycolytic enzyme PGK1 phosphoglycerate kinase facilitates the assembly of viral replicase complexes. *PLOS Pathogens*, *13*(10), e1006689. https://doi.org/10.1371/journal.ppat.1006689

Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., Gruissem, W., & Baginsky, S. (2009). Large-Scale Arabidopsis Phosphoproteome Profiling Reveals Novel Chloroplast Kinase Substrates and Phosphorylation Networks. *Plant Physiology*, *150*(2), 889–903. https://doi.org/10.1104/pp.109.138677

Rigden, D. J., Phillips, S. E. V., Michels, P. A. M., & Fothergill-Gilmore, L. A. (1999). The structure of pyruvate kinase from Leishmania mexicana reveals details of the allosteric transition and unusual effector specificity. *Journal of Molecular Biology*, *293*(3), 745–749. https://doi.org/10.1006/jmbi.1999.3170

Rodgers, K., & McVey, M. (2016). Error-Prone Repair of DNA Double-Strand Breaks: ERROR-PRONE REPAIR OF DNA DSBs. *Journal of Cellular Physiology*, *231*(1), 15–24. https://doi.org/10.1002/jcp.25053

Sakai, H. (2004). Possible Structure and Function of the Extra C-Terminal Sequence of Pyruvate Kinase from Bacillus stearothermophilus. *Journal of Biochemistry*, *136*(4), 471–476. https://doi.org/10.1093/jb/mvh152

Schaechter, M., MaalOe, O., & Kjeldgaard, N. O. (1958). Dependency on Medium and Temperature of Cel Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *J.gen.Microbiol.*, *19*, 592–606.

Schormann, N., Hayden, K. L., Lee, P., Banerjee, S., & Chattopadhyay, D. (2019). An overview of structure, function, and regulation of pyruvate kinases. *Protein Science*, *28*(10), 1771–1784. https://doi.org/10.1002/pro.3691

Sharpe, M. E., Hauser, P. M., Sharpe, R. G., & Errington, J. (1998). Bacillus subtilis Cell Cycle as Studied by Fluorescence Microscopy: Constancy of Cell Length at Initiation of DNA Replication and Evidence for Active Nucleoid Partitioning. *J. BACTERIOL.*, *180*, 9.

Shor, E., Warren, C. L., Tietjen, J., Hou, Z., Müller, U., Alborelli, I., Gohard, F. H., Yemm, A. I., Borisov, L., Broach, J. R., Weinreich, M., Nieduszynski, C. A., Ansari, A. Z., & Fox, C. A.

(2009). The Origin Recognition Complex Interacts with a Subset of Metabolic Genes Tightly Linked to Origins of Replication. *PLoS Genetics*, *5*(12), e1000755. https://doi.org/10.1371/journal.pgen.1000755

Siddiqui, K., On, K. F., & Diffley, J. F. X. (2013). Regulating DNA Replication in Eukarya. *Cold Spring Harbor Perspectives in Biology*, *5*(9), a012930–a012930. https://doi.org/10.1101/cshperspect.a012930

Skarstad, K., Boye, E., & Steen, H. B. (1986). Timing of initiation of chromosome replication in individual Escherichia coli cells. *EMBO Journal*, *5*(7), 1711–1717.

Slavov, N., & Botstein, D. (2011). Coupling among growth rate response, metabolic cycle, and cell division cycle in yeast. *Molecular Biology of the Cell*, *22*(12), 1997–2009. https://doi.org/10.1091/mbc.e11-02-0132

Snaebjornsson, M. T., & Schulze, A. (2018). Non-canonical functions of enzymes facilitate cross-talk between cell metabolic and regulatory pathways. *Experimental & Molecular Medicine*, *50*(4), 34. https://doi.org/10.1038/s12276-018-0065-6

Sprague, G. F. (1977). Isolation and Characterization of a Saccharomyces cerevisiae Mutant Deficient in Pyruvate Kinase Activity. *J. BACTERIOL.*, *130*, 10.

Suzuki, K., Ito, S., Shimizu-Ibuka, A., & Sakai, H. (2008). Crystal Structure of Pyruvate Kinase from Geobacillus stearothermophilus. *Journal of Biochemistry*, *144*(3), 305–312. https://doi.org/10.1093/jb/mvn069

Teplyakov, A., Lim, K., Zhu, P.-P., Kapadia, G., Chen, C. C. H., Schwartz, J., Howard, A., Reddy, P. T., Peterkofsky, A., & Herzberg, O. (2006). Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proceedings of the National Academy of Sciences*, *103*(44), 16218–16223. https://doi.org/10.1073/pnas.0607587103

Tian, M., Sasvari, Z., Gonzalez, P. A., Friso, G., Rowland, E., Liu, X.-M., van Wijk, K. J., Nagy, P. D., & Klessig, D. F. (2015). Salicylic Acid Inhibits the Replication of *Tomato bushy stunt virus* by Directly Targeting a Host Component in the Replication Complex. *Molecular Plant-Microbe Interactions*, *28*(4), 379–386. https://doi.org/10.1094/MPMI-09-14-0259-R

Tolentino, R., Chastain, C., & Burnell, J. (2013). Identification of the amino acid involved in the regulation of bacterial pyruvate, orthophosphate dikinase and phosphoenolpyruvate synthetase. *Advances in Biological Chemistry*, *03*(03), 12–21. https://doi.org/10.4236/abc.2013.33A003

Tu, B. P. (2005). Logic of the Yeast Metabolic Cycle: Temporal Compartmentalization of Cellular Processes. *Science*, *310*(5751), 1152–1158. https://doi.org/10.1126/science.1120499

Tubbs, A., & Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell*, *168*(4), 644–656. https://doi.org/10.1016/j.cell.2017.01.002

Tymecka-Mulik, J., Boss, L., Maciąg-Dorszyńska, M., Matias Rodrigues, J. F., Gaffke, L., Wosinski, A., Cech, G. M., Szalewska-Pałasz, A., Węgrzyn, G., & Glinkowska, M. (2017). Suppression of the Escherichia coli dnaA46 mutation by changes in the activities of the pyruvate-acetate node links DNA replication regulation to central carbon metabolism. *PLOS ONE*, *12*(4), e0176050. https://doi.org/10.1371/journal.pone.0176050

Wang, J. D., & Levin, P. A. (2009). Metabolism, cell growth and the bacterial cell cycle. *Nature Reviews Microbiology*, *7*(11), 822–827. https://doi.org/10.1038/nrmicro2202

Warburg, O. (1956). On the Origin of Cancer Cells. *Science*, *123*(3191), 309–314. https://doi.org/10.1126/science.123.3191.309

Weart, R. B., Lee, A. H., Chien, A.-C., Haeusser, D. P., Hill, N. S., & Levin, P. A. (2007). A Metabolic Sensor Governing Cell Size in Bacteria. *Cell*, *130*(2), 335–347. https://doi.org/10.1016/j.cell.2007.05.043

Zhong, W., Cui, L., Goh, B. C., Cai, Q., Ho, P., Chionh, Y. H., Yuan, M., Sahili, A. E., Fothergill-Gilmore, L. A., Walkinshaw, M. D., Lescar, J., & Dedon, P. C. (2017). Allosteric pyruvate kinase-based "logic gate" synergistically senses energy and sugar levels in Mycobacterium tuberculosis. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-02086-y

Supplementary material Chapter 2

A Coordinates of the catalytic residues of the Cat domain of PykA.

Function	B.subtilis	PKM2	Mycobacterium tuberculosis
Binding site (PEP)	R32	R79	R33
Metal binding (K)	N34	N75	N35
Metal binding (K)	S36	S77	S37
Metal binding (K)	D66	D113	D67
Metal binding (K)	T67	T114	-
Binding site (ATP)	R73	R120	R74
Binding site (ATP)	K156	K207	K155
Transition state stabilizer	K220	K270	K218
Metal binding (Mg)	E222	E272	E220
Binding site (PEP)	G245	G295	G243
Metal binding (Mg)	D246	D296	D244
Binding site (PEP)	D246	D296	D244
Binding site (PEP)	T278	T328	T276

B The PEPut domain of PykA is highly conserved and contains the TSH motif.

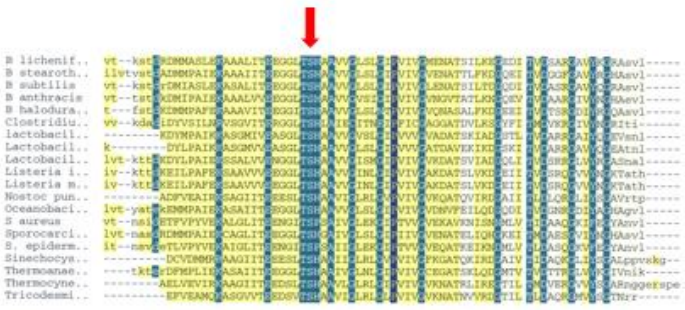


Fig. S1 : important residues in PykA in *B. subtilis* (A) Coordinates of key catalytic residues in pyruvate kinases in *B. subtilis,* human PKM2 and *M. tuberculosis* Taken from UniProt (B) The conserved TSH motif in the PEPut domain
Fig. S2 : Metabolomic analysis of WT and *ΔpykA*

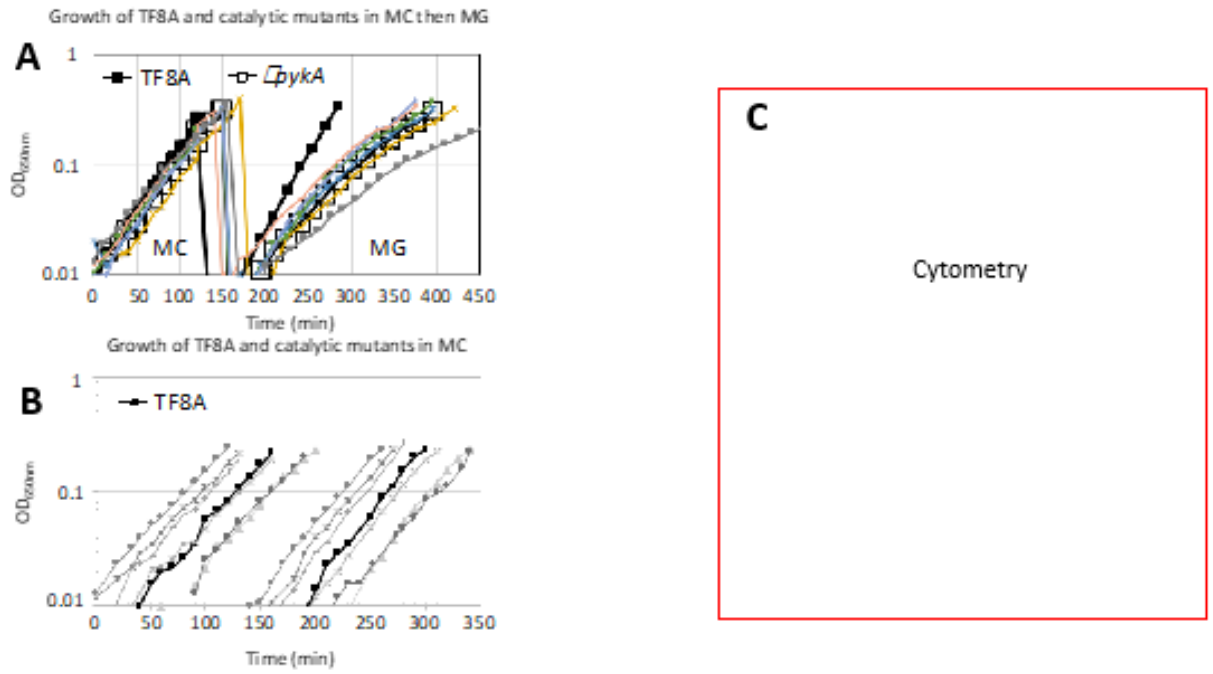


Fig. S3 : A: Growth of catalytic mutants of PykA in MC. The tested mutants are the following: *pykA*$_{Dcat}$, *pykA*$_{JP}$, *pykA*$_{R32A}$, *pykA*$_{R73A}$, *pykA*$_{K220A}$ *and pykA*$_{DG246-6AA}$. The wild-type (TF8A) and *DpykA* were used as controls. B: Growth of the TF8A strain and Cat mutants in MC. C: Microscopy.

Fig. S4 : Metabolomic analysis of catalytic mutants

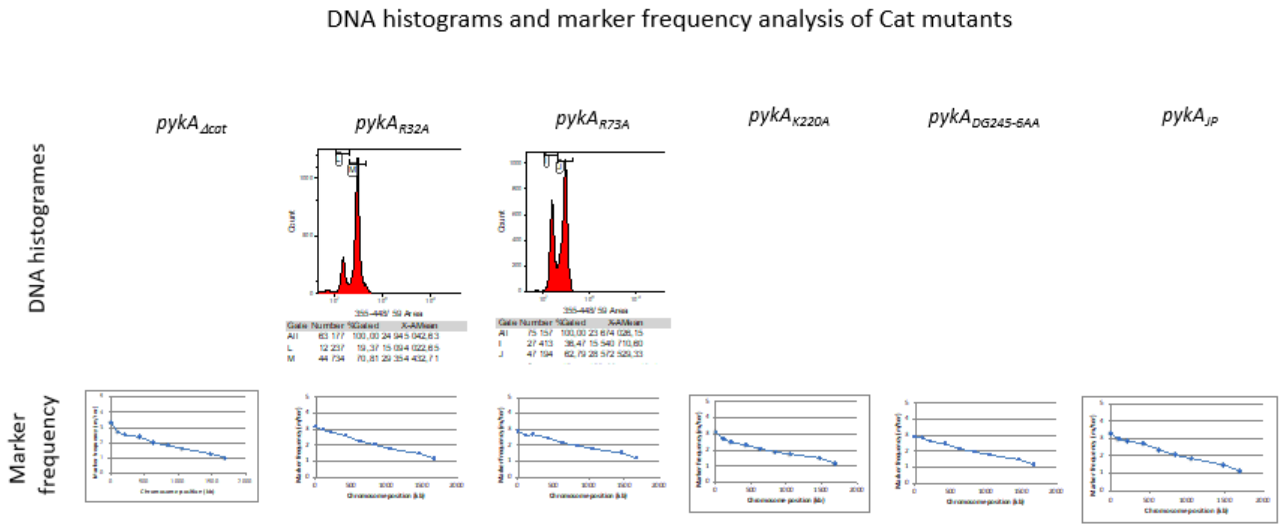


Fig. S5 : Raw data on replication parameters in cat mutants DNA histograms, marker frequencies and microscopic images will be shown

Fig. S6 : General properties of PEPut mutants: growth data and microscopy

Fig. S7 : Metabolomic analysis of PEPut mutants

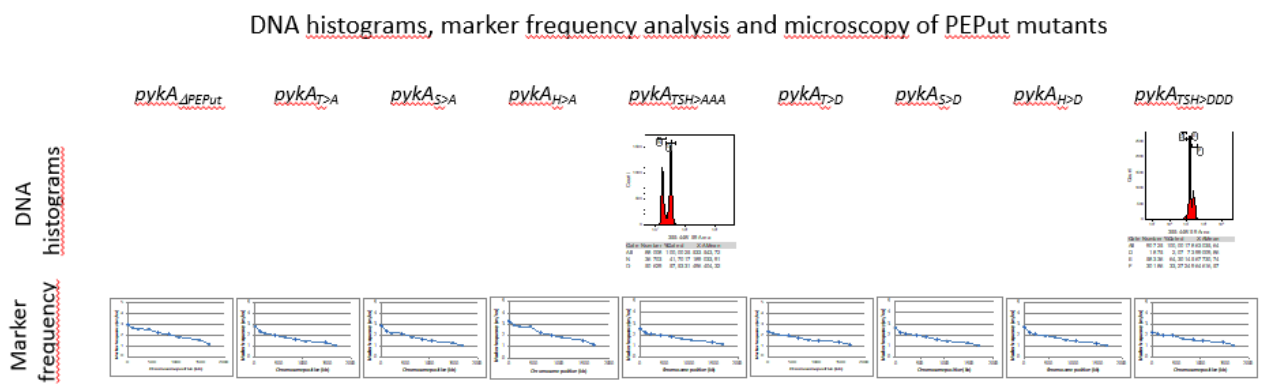


Fig. S8: Raw data on replication parameters in PEPut mutants DNA histograms, marker frequencies and microscopic images will be shown

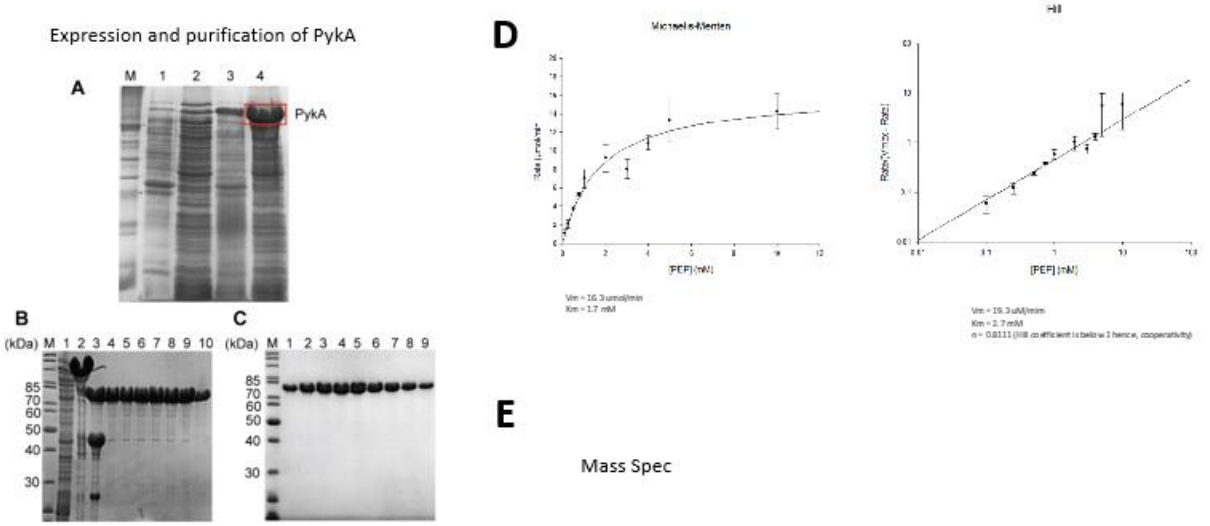


Fig. S9

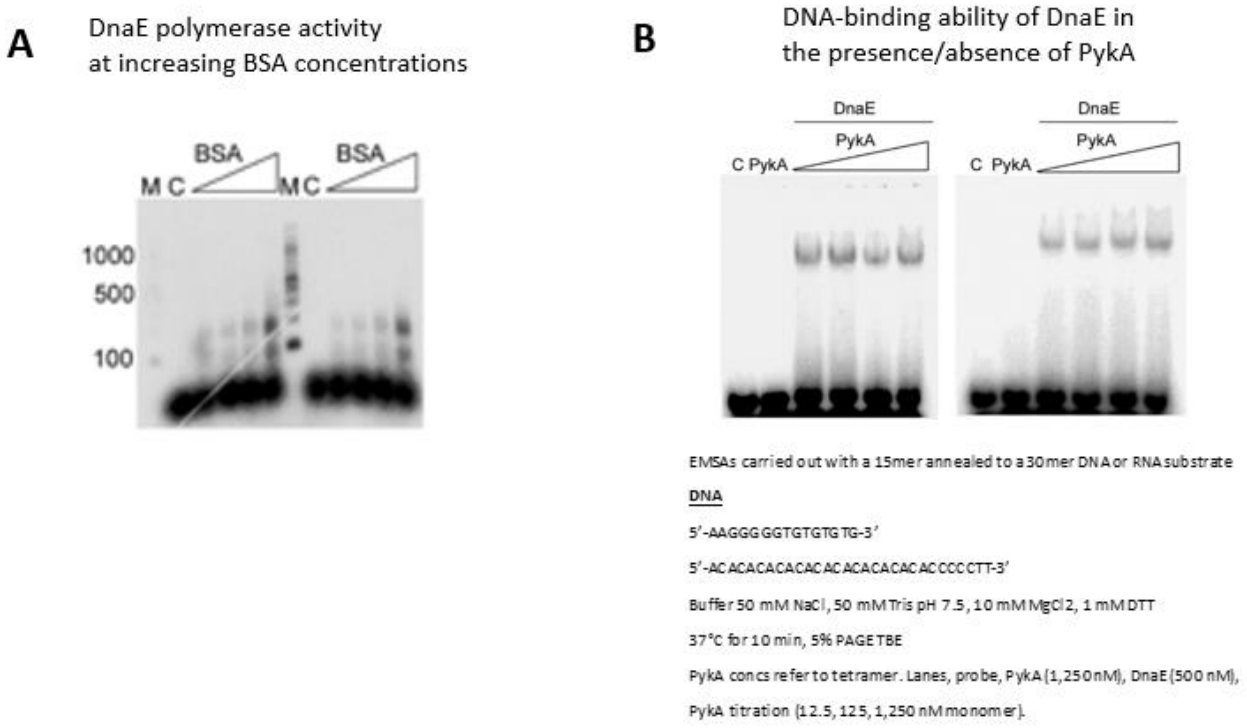


Fig S10

Table S1: Genetic constructs used in this work

Experiment	Name	Context	Genotype	Main phenotypes	Source of construction	Obtained by transformation with:
Parental strains	TF8A		*trpC2*	wild-type	Jannière et al, 2007	
	GM2840	BSB168	*amyE::HyPEPut-cat*	PEPut domain expressed at *amyE* from the IPTG dependent *Hyspank* promoter, CmR	Personal gift from Dominique Le Coq	
	DGRM295	TF8A	*pykA-tet*	Insertion of the TetR marker just downstream of *pykA*	This work, Laurent	PCR product[1]
	DGRM25	TF8A	*pykA::spc trpC2*	Deletion of *pykA*, SpR	Jannière et al, 2007	
PEP utilizer domain mutant	DGRM296N	TF8A	$pykA_{1la}$-*tet trpC2*	Encodes the catalytic domain of PykA (PEPut deletion), TetR	This work, Steff	PCR product
	DGRM298	TF8A	$pykA_{H539A}$-*tet trpC2*	Mutation H539A in PEPut, TetR	This work, Laurent	PCR product
	DGRM299	TF8A	$pykA_{T537A}$-*tet trpC2*	Mutation T537A in PEPut, TetR	This work, Laurent	PCR product
	DGRM302	TF8A	$pykA_{TSH537-539AAA}$-*tet trpC2*	Mutation TSH537-539AAA in PEPut, TetR	This work, Laurent	PCR product
	DGRM303	TF8A	$pykA_{S538A}$-*tet trpC2*	Mutation S538A in PEPut, TetR	This work, Steff	PCR product
	DGRM1015	TF8A	$pykA_{H539D}$-*tet trpC2*	Mutation H539D in PEPut, TetR	This work, Steff	PCR product
	DGRM1016	TF8A	$pykA_{S538D}$-*tet trpC2*	Mutation S538D in PEPut, TetR	This work, Steff	PCR product
	DGRM1017	TF8A	$pykA_{PEPut}$-*tet trpC3*	Encodes the PEPut domain of PykA (deletion of the catalytic domain), TetR	This work, Steff	PCR product
	DGRM1018	TF8A	$pykA_{T537D}$-*tet trpC2*	Mutation T537D in PEPut, TetR	This work, Steff	PCR product
	DGRM1019	TF8A	$pykA_{TSH537-539DDD}$-*tet trpC3*	Mutation TSH537-539DDD in PEPut, TetR	This work, Steff	PCR product
Catalytic domain mutant	DGRM24	TF8A	*pykAJP-prm trpC2*	Deletion of 27 amino acids in the catalytic site of PykA, PhlR	Jannière et al, 2007	
	DGRM1009	TF8A	$PykA_{1la}$-*tet trpC2*	Pyruvate kinase of *B. subtilis* replaced by the one of *L. lactis*, TetR	This work, Steff	PCR product
	DGRM1010	TF8A	$pykA_{1la\text{-}PEPut}$-*tet trpC2*	ruvate kinase of *B. subtilis* replaced by the one of *L. lactis* fused to PEPut of *B. subtilis PykA*, T	This work, Steff	PCR product
	DGRM1094	TF8A	$pykA_{R32A}$-*tet trpC2*	Mutation R32A in the phophoenolpyruvate binding site of PykA, TetR	This work, Steff	PCR product
	DGRM1095	TF8A	$pykA_{R73A}$-*tet trpC2*	Mutation R37A in the ADP binding site of PykA, TetR	This work, Steff	PCR product
	DGRM1096	TF8A	$pykA_{K220A}$-*tet trpC2*	Pleiotropic mutation K220A in the catalytic site of PykA, TetR	This work, Steff	PCR product
	DGRM1097	TF8A	$pykA_{GD245-246AA}$-*tet trpC2*	Mutation DG245-246AA in the Mg^{2+} binding site of PykA, TetR	This work, Steff	PCR product
hyPEP mutants	DGRM722	TF8A	*amyE::HyPEPut-cat*	PEPut expressed at *amyE* from the IPTG dependent *Hyspank* promoter, CmR	This work, Laurent	gGM2840 ➤ TF8A (CmR)[2]
	DGRM733	TF8A	*amyE::HyPEPut-cat* $pykA_{1la}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding the catalytic domain of PykA, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM296N (CmR)
	DGRM734	TF8A	*amyE::HyPEPut-cat* $pykA_{H539A}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{H539A}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM298 (CmR)
	DGRM735	TF8A	*amyE::HyPEPut-cat* $pykA_{TSH537-539AAA}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{TSH537-539AAA}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM302 (CmR)
	DGRM1020	TF8A	*amyE::HyPEPut-cat pykA::spc*	PEPut at *amyE* under *Hyspank* in a strain lacking *pykA*, CmR, SpcR	This work, Steff	gDGRM722 ➤ DGRM25 (CmR)
	DGRM1021	TF8A	*amyE::HyPEPut-cat pykAJP-prm*	PEPut at *amyE* under *Hyspank* in a strain encoding PykAJP, CmR, PhlR	This work, Steff	gDGRM722 ➤ DGRM24 (CmR)
	DGRM1022	TF8A	*amyE::HyPEPut-cat* $pykA_{R32A}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{R32A}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM1094 (CmR)
	DGRM1023	TF8A	*amyE::HyPEPut-cat* $PykA_{R73A}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{R73A}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM1095 (CmR)
	DGRM1024	TF8A	*amyE::HyPEPut-cat* $PykA_{K220A}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{K220A}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM1096 (CmR)
	DGRM1025	TF8A	*amyE::HyPEPut-cat* $PykA_{DG245-246AA}$-*tet*	PEPut at *amyE* under *Hyspank* in a strain encoding $PykA_{DG245-246AA}$, CmR, TetR	This work, Laurent	gDGRM722 ➤ DGRM1097 (CmR)
	DGRM1030	TF8A	*amyE::HyPEP* $_{H539A}$-*cat*	$PEPut_{H539A}$ at *amyE* under *Hyspank*, CmR	This work, Laurent	PCR product
	DGRM1040	TF8A	*amyE::HyPEP* $_{T537A}$-*cat*	$PEPut_{T537A}$ at *amyE* under *Hyspank*, CmR	This work, Laurent	PCR product
	DGRM1060	TF8A	*amyE::HyPEP* $_{H539D}$-*cat*	$PEPut_{H539D}$ at *amyE* under *Hyspank*, CmR	This work, Laurent	PCR product
	DGRM1070	TF8A	*amyE::HyPEP* $_{T537D}$-*cat*	$PEPut_{T537D}$ at *amyE* under *Hyspank*, CmR	This work, Laurent	PCR product

Table S2: Primers for strain construction

	Primer name	Construction	Orientation	Sequence
Construction primers PEP utilizer mutants	CS3	H→ A mutation	Forward	GGT TTG ACT AGC gcT GCT GCG GTA G
	CS4	H → A mutation	Reverse	CTA CCG CAG C AGC GCT AGT CAA ACC
	LJ101	T → A mutation	Forward	gaaggcggtttgGcCagccatgctgcg
	LJ102	T → A mutation	Reverse	CGC AGC ATG GCT GGC CAA ACC GCC TTC
	LJ109	TSH → AAA mutation	Forward	GGT TTG gCT gct gcT GCT GCG GTA G
	LJ110	TSH → AAA mutation	Reverse	CTA CCG CAG C Agc agc AGc CAA ACC
	LJ103	S → A mutation	Forward	ggcggtttgactGCGcatgctgcgg
	LJ104	S → A mutation	Reverse	CCG CAG CAT GCG CAG TCA AAC CGC C
	LJ563	H-→ D mutation	Forward	GGTTTGACTAGCgATGCTGCGGTAG
	LJ564	H → D mutation	Reverse	CTACCGCAGCATCGCTAGTCAAACC
	pSH262	S → D mutation	Forward	ggcggtttgactGATcatgctgcgg
	pSH263	S → D mutation	Reverse	CCGCAGCATGATCAGTCAAACCGCC
	pSH268	PEP	Forward	CTTGGCCTTTAGCGATGATATCCATTTGGTTCACTTCCTTCTGAAATCTTCA
	pSH269	PEP	Reverse	GATATCATCGCTAAAGGCCAAG
	pSH260	T-→ D mutation	Forward	gaaggcggtttggaTagccatgctgcg
	pSH261	T-→ D mutation	Reverse	CGCAGCATGGCTATCCAAACCGCCTTC
	pSH277	TSH-→ DDD mutation	Forward	gaaggcggtttggaTGATGatGCTGCGGTAG
	pSH278	TSH-→ DDD mutation	Reverse	CTACCGCAGCATCATCATCCAAACCGCCTTC
Construction primers catalytic mutants	LJ610	PykAR32A	Forward	TGGCTGCATTAAACTTTTCTCACGGAG
	LJ611	PykAR32A	Reverse	AGAAAAGTTTAATGCAGCCACGTTCATTCCTGACTC
	LJ612	PykAR73A	Forward	TGAAATCGCTACACATACAATGGAAAACGG
	LJ613	PykAR73A	Reverse	TTGTATGTGTTGCGATTTCAGGACCTTTTGTATC
	LJ614	PykAK220A	Forward	AATCATCCCTGCAATCGAAAACCAAGAGG
	LJ615	PykAK220A	Reverse	TTTTCGATTGCAGGGATGATTTGAATATCC
	LJ616	PykAG245AD246A	Forward	GTTGCACGCGCAGCTTTAGGTGTGGAAATTCCAGCTG
	LJ617	PykAG245AD246A	Reverse	CCTAAAGCTGCGCGTGCAACCATTAAGCCGTCAG
	LJ604	PykALL/PykALLPEP	Forward	GATTTCAGAAGGAAGTGAACCAAATGAATAAACGTGTAAAAATCGTCTC
	LJ618	PykALL	Forward	CTCTACTGAATTTAAAACTTAG
	LJ619	PykALL	Reverse	CTAAGTTTTAAATTCAGTAGAGTTATTTAACAGTGCGGATACGC
	LJ608	PykALLPEP	Forward	GATATCATCGCTAAAGGCC
	LJ609	PykALLPEP	Reverse	GGCCTTTAGCGATGATATCTTTAACAGTGCGGATACGC
External primers PykA region	CS1	PykA constructions	Reverse	CATACAATGGAAAACGGCGG
	CS10	PykA constructions	Forward	GCCGGATCGTGTGACAGCTAAC
	BD154	PykA constructions	Reverse	ATGAAACGAATAGGGGTATTAACG
External primers amyE locus	LJ161	hyPEP mutations	Forward	TTCTCCAGTCTTCACATCGGTTTGA
	LJ165	hyPEP mutations	Reverse	TGCCGCTTCCAATCACCC
External primers PykA region HiFi assembly	LJ621	PykA constructions	Forward	ATGGCTCGAGTTTTTCAGCAAGATGCC GGA TCG TGT GAC AGC TAA C
	LJ622	PykA constructions	Reverse	TGTAGGAGATCTTCTAGAAAGATTTATCGGTGGAGACGGTTCC
	LJ623	PykA constructions	Forward	ATGGCTCGAGTTTTTCAGCAAGATTTAGGACTGATTGCGGGAC

Table S3: Primers for qPCR analysis

Name	Primer name	Chromosome arm	Position (degree)	Orientation	Sequence	Product size
Ori	LJ114	oriC	0	Forward	TCCGAGATAATAAAGCCGTCGA	101
	LJ115			Reverse	CTGGGTTTGTTCTTTCCCCG	
R05	LJ518	Right	9	Forward	CGCGAGCAAGTAGAGGATAC	210
	LJ519			Reverse	CGCTACAACTGCTTCATCCT	
R1	LJ520		18	Forward	GGGATTTTAACACCAGCCGA	191
	LJ521			Reverse	ACCATCACTCGTTCACCTGT	
R2	LJ480		36	Forward	TGACCGGGACACAGAATTCA	169
	LJ481			Reverse	ATTGATTTTGACCACCCGCC	
R3	LJ482		54	Forward	CGACGAACTCAATCATGGGG	196
	LJ483			Reverse	ATGTGCCAAGTCTCTCTCCC	
R4	LJ484		72	Forward	TCATTGCAAGACCTCCTCGT	173
	LJ485			Reverse	CAAGACGCCGGATAAAAGCA	
R5	LJ486		90	Forward	CCACAATGCTTCCGGAGATG	194
	LJ487			Reverse	TCGGGGAGATTGGTTTCGAT	
R7	LJ522		126	Forward	TCCAATTCTCCAACCGGTCA	204
	LJ523			Reverse	ACCGGCGTGACTTTATCTGA	
R8	LJ524		144	Forward	ATCTATTGGATCAGGCCGCT	215
	LJ525			Reverse	ATAAGCAGCAAGTACACGCG	
terC	BD280	TerC	180	Forward	TTTCAGTTGTGCCACCATGT	187
	BD281			Reverse	ATTTGCCGTTCTCGGGTTA	

Chapter 3: PykA couples DNA replication timing to metabolism during a shift from gluconeogenic to glycolytic media

université PARIS-SACLAY
ÉCOLE DOCTORALE
Structure et dynamique des systèmes vivants (SDSV)

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

Materials and methods

In general, experiments were performed as in [86]. What follows is a brief explanation of the experimental procedure.

1. Strains and plasmids

Strains and plasmids are listed in Supplementary Table S1. *B. subtilis* cells were grown at 37ºC in LB supplement or not with malate 0.4%. Unless stated otherwise, antibiotics were used at the following concentrations: spectinomycin (Sp, 60 µg/mL); tetracycline (Tet, 10 µg/ml); chloramphenicol (Cm, 5 µg/mL); phleomycin (Pm, 10 µg/mL); ampicillin (100 µg/ml). Liquid cultures were performed in an incubator under strong shaking (200–230rpm).

In order to obtain most point mutations and deletions in the *pykA* gene, two PCR reactions were performed on strain 295 (*pykA::Tet*). In each reaction, one external primer and one construction specific primer was used to introduce the desired mutation and later refuse the PCR products to a mutated version of *pykA.* A list of primers for each construct is found in Supplementary table 2. Next, the PCR reaction products were purified from gel using the Monarch Gel Extraction Kit (NEB) to remove parental genomic DNA. The purified PCR products were then fused together using assembly PCR and the parental PCR fragments were removed using gel purification with the Monarch Gel Extraction Kit. Competent *B. subtilis* strains (TF8A) were then transformed with the PCR products and they were integrated into the chromosome of *B. subtilis* using double cross-over. Positive transformants were selected on Tet. The genotype of the transformants was verified by PCR, phenotypic analysis and/or DNA sequencing. Mutants thought to inhibit PykA activity were selected on LB + Tet + Malate (0,4%) and slow growth on LB (which is glycolytic) was verified. Strains 722 (*hyPEP*) and 1019 (*pykATSH>DDD*) were obtained by transformation with linearized plasmids pIC610 and TSH>DDD-pJET respectively. TSH>DDD-pJET was obtained as follows. First, two PCR reactions were performed on strain 295 (*pykA::Tet*). In each reaction, one external primer (LJ621 and LJ622) and one construction specific primer (pSH277 and pSH278) was used to introduce the desired mutation and refuse the PCR products to a mutated version of *pykA.* The external primers had overhangs homologous to pJET. Next, the PCR reaction products were purified from gel using the Monarch Gel Extraction Kit (NEB) to remove parental genomic DNA. The purified PCR products were then fused together using assembly PCR and the parental PCR fragments were removed using gel purification with the Monarch Gel Extraction Kit. The purified PCR products were then cloned into a linearized *E. coli* plasmid pJET and used for transformation in Stellar E. coli competent cells (Clontech, Saint-Germain-en-Laye,

France). Plasmids from positive transformants were selected on Amp, extracted using the Monarch Plasmid Miniprep Kit (NEB) and verified using agarose gel electrophoresis and DNA sequencing. Plasmids with correct sequences were linearized by digestion with the restriction enzyme FspI (NEB) and used to transform competent *B. subtilis* strains. The TSH>DDD mutations was obtained by double crossover of the linearized plasmid into the chromosome. Positive transformants were selected on Tet. The genotype of the transformants was verified by PCR, phenotypic analysis and/or DNA sequencing.

Strains containing *hyPEP* in other genetic backgrounds were obtained by transforming competent cells of these genetic backgrounds with genomic DNA of strain 722. Positive transformants were selected on Cm and verified using PCR verification and phenotypic analysis.

Strains 1009 (*pykA Lactococcus Lactis*) and 1010 (*pykA Lactococcus Lactis* fused to PEP utilizer from *B. subtilis*) were obtained by fusing *pykA* from *L. Lactis* to DNA sequences homologous to the *pykA* region and containing a Tet marker using assembly PCR. Detailed protocol is as described for the point mutations/deletions in *pykA*. *pykA L.Lactis* was amplified from plasmid pLJH63. Competent *B. subtilis* cells were transformed with the assembled and gel purified PCR fragments. Positive transformants were selected on Tet and verified by PCR, phenotypic analysis and sequencing.

Mutated versions of *hyPEP* were obtained as described for *pykA* mutants using the plasmid pIC610 as a template for generating PCR fragments for assembly PCR. Positive transformants were selected on Cm and verified using PCR verification, DNA sequencing and phenotypic analysis.

2. Quantitative PCR

2.1 Measurement of the *ori-ter* ratio during metabolic shift

Wild-type cells cured of prophages (TF8A) and containing or not CCM mutations were grown overnight in minimal medium (K_2HPO_4: 80 mM; KH_2PO_4: 44 mM; $(NH_4)_2SO_4$: 15 mM; $C_6H_5Na_3O_7$ $2H_20$: 3, 4 mM; $CaCl_2$: 50 mM; $MgSO_4$: 2 mM; FeIII citrate: 11 µg/mL; $MnCl_2$: 10 µM; $FeSO_4$: 1 µM; $FeCl_3$:4 µg/mL; Trp 50 µg/mL) supplemented with casein hydrolysate 0.2%, malate 0.4% and tryptophan 0.01% (MC medium) in triplicates (with a very few exceptions, the three independent cultures were inoculated with different constructs). Saturated cultures were then diluted 1000-fold in the same media and growth was carefully monitored using spectrophotometry. At OD 0.08-0.1, the t_0 sample was taken and 0.8% glucose was added to the MC medium (MCG). In general, samples were collected every 15 min until timepoint 75 min, unless stated otherwise. Cells were generally diluted 1/3 in prewarmed MCG medium after 45 min to maintain exponential growth. The genomic DNA was extracted using the GeneJET genomic DNA extraction kit (Thermoscientific) and every qPCR reaction was carried out using two technical repeats of 4–5

serial dilutions. A nonreplicating control DNA (stage II sporlets[230]) was analysed simultaneously with the samples in about 1/3 of the qPCR plates. Primers for the marker frequency analysis (see Supplementary Table S3 for details) were used at 200nM in 12ml of the 1 x SYBR qPCR Premix Ex Taq (Tli RNaseH Plus) (Ozyme, St Quentin en Yvelines, France) and DNA amplification was carried out as follows: denaturation: 95ºC, 30s; amplification: (33 cycles) 95ºC, 5s and 60ºC, 1min. For every qPCR reaction, the primer efficiency (97 to 103%) was controlled. The *ori/ter* ratio of the non-replicating sporlets was 0.585þ/– 0.019, a value consistent with the size of the PCR products (Supplementary Table S3). The amplifications were carried out on a Mastercycler® ep realplex (Eppendorf, Le Pecq, Fr) and the ΔCq were determined automatically from the cycle threshold.

2.2 Measurement of the C-period during metabolic shift

Cell culture, genomic DNA extraction and qPCR amplification reactions were as above. Primer pairs, their positions on the chromosome map and the size of the PCR products are given in Supplementary Table S3. Results were reported to the *terC* signal and corrected for the size of the PCR products. The C period was extracted from the slope of the straight line of the plot $(\log_2 Nm)\tau$ against m where Nm is the relative copy number of a sequence N at position m (with $0<m<1$ along the ori-ter axis) and τ the generation time in minutes.[231] Control experiments showed that the marker frequency of the different loci in the non-replicating DNA was close to 1 (1.011 þ/– 0.045) as expected. Similar data were obtained with primers on the right and left arm of the chromosome.

3. Flow cytometry analysis

3.1 Measurement of the number of origins/cell during metabolic shift

Strains were grown as indicated in the previous section. At OD 0.08-0.1, 10 ml of cell culture, corresponding to t_0, was added to chloramphenicol (200 µg/mL final) to impede replication initiation, cell division and allow completion of ongoing rounds of replication.[232] 0.8% glucose was added to the rest of the culture, changing the MC medium into an MCG medium. In general, 5 ml of cell culture samples were added to chloramphenicol (200 µg/mL final) every 15 min until timepoint 75 min unless stated otherwise. Cells were generally diluted 1/3 in prewarmed MCG medium after 45 min to maintain exponential growth.

After 4–6 hours of drug treatment, 10^8 cells were fixed in filtered ethanol 70% and stored at 4ºC. Stored cells were then washed in filtered 0.1M of phosphate buffer (PB) pH 9, resuspended in 1mL of Tris-buffered saline buffer (TBS 150, filtered) (20mM Tris-HCl pH 7.5, 150mM NaCl) and stained with Hoechst 33258 (1.5 µg/mL) for at least 30 min as described elsewhere.[233] Flow cytometry analysis was performed using a MoFlow Astrios cell sorter (Beckman Coulter, Life Sciences) equipped with a 355nm krypton laser and a 448/59 nm bandpass filter used to collect

Hoechst 33258 fluorescence. Data were analysed with the Kaluza software (Beckman Coulter, Life Sciences). We counted 50,000 to 100,000 events. In most of the tested samples, DNA histograms showed 1–2 main peaks with a 2^n distribution. The number of origins/cell was obtained from the proportion of cells with 1, 2, 4, 8 and/or 16 chromosomes.

3.2 Age of initiation

The age of initiation was calculated from the equation $a_i=\tau - \tau * \ln(2-F) / \ln(2)$where τ is the generation time and F the fraction of uninitiated cells (Fraction of cells in the lowest peak).[82]

4 Microscopy analysis

In order to analyze cell size and the presence of a nucleoid, cells were grown overnight at 30ºC in LB and diluted 1000-fold in the same medium for further growth. At OD 0.080-0,1 ,cells were collected and the nucleoid was stained with 4',6'-diamidino-2-phenylindole (DAPI, 10mg/ mL) (Sigma). Stained cells were then mounted on glass slides (VWR) covered with 1.2% agarose (low melting point, Sigma) in minimal medium supplemented with malate 0.4% and casa 0.2% and a 0.17mm glass coverslip (VWR) was placed on top. Images were acquired on an epifluorescence microscope (Zeiss, Axio Observer.Z1) with a 100x magnification oil-immersion objective (Plan-APOCHROMAT Ph3) and a CMOS camera (Orca Flash 4.0 LT Hamamatsu). The images were analysed using FiJi. Cell lengths and DAPI fluorescence were exported to Excel for statistical analysis. All experiments were independently performed at least three times.

Results

1. Dynamic adaptation of cell cycle parameters in *B. subtilis* to a shift from MC medium to MCG medium

The goal of this work is to demonstrate that PykA communicates information about cellular metabolism to the replication machinery. To achieve this, the effect of a metabolic shift from rich gluconeogenic medium (MC: see previous chapter) to rich glycolytic medium (MC + 0.8% Glucose =MCG) on replication parameters was studied in wild type (TF8A) and *pykA* mutants. Concretely, 0.8% of glucose was added to cell cultures growing in MC and samples were collected every 15 min for *ori-ter* ratio analysis, marker frequency analysis and run-out flow cytometry. *Ori-ter* ratio analysis in TF8A during the shift reveals that the *ori-ter* ratio remains constant for 15 min after the shift and starts to increase afterwards to reach a plateau at 45-60 min after the shift (Fig. 21A). This plateau *ori-ter* ratio is the same as the *ori-ter* ratio observed of TF8A in steady state MCG. However, determination of replication parameters by qPCR and flow cytometry reveals a complex

multiphase adaptation of replication parameters to the shift (fig 21B-D). Between 0 and 15 min after the shift, none of the cell cycle parameters changes in TF8A. Between 15 min and 45 minutes, the C-period (time to replicate the chromosome) increases from 40 to 55 min (fig. 21B). Next, from 30-45 minutes onwards, the initiation frequency (number of origins/cell) increases (fig. 21C) and the age of initiation is pushed forward (fig. 21D). To our knowledge, this behaviour has not yet been described in literature.

A caveat to this study is that many of the cell cycle parameters require the growth rate to be calculated and it is not possible to quantify growth rate changes during the shift. The generation time of TF8A in MC is about 30 min and this drops to 23 min in MCG (data not shown). Therefore, the growth rate might change during the shift experiment.

However, it is known that the average cell size of *B. subtilis* is connected to the growth rate[75]. Therefore, the cell size at specific time points was used as a proxy to assess whether the growth rate changes after a specific time point. This analysis demonstrated that the cell size of TF8A is smaller in steady state MC (5.67µm) than in steady state MCG (6.2 µm), consistent with the known connection between cell size and growth rate[75]. Interestingly though, the cell size didn't change after 45 min. Therefore, the growth rate was assumed constant between 0 min and 45 min after the shift.

In summary, TF8A (wild type) cells adapt their replication parameters to a metabolic shift from MC to MCG medium in a complex, multiphase manner involving in order (i) a short period in which replication doesn't change (ii) a temporary slowdown of replication fork speed (iii) increase in initiation frequency and advancement of the initiation age and finally (iii) increase of the cell size/growth rate.

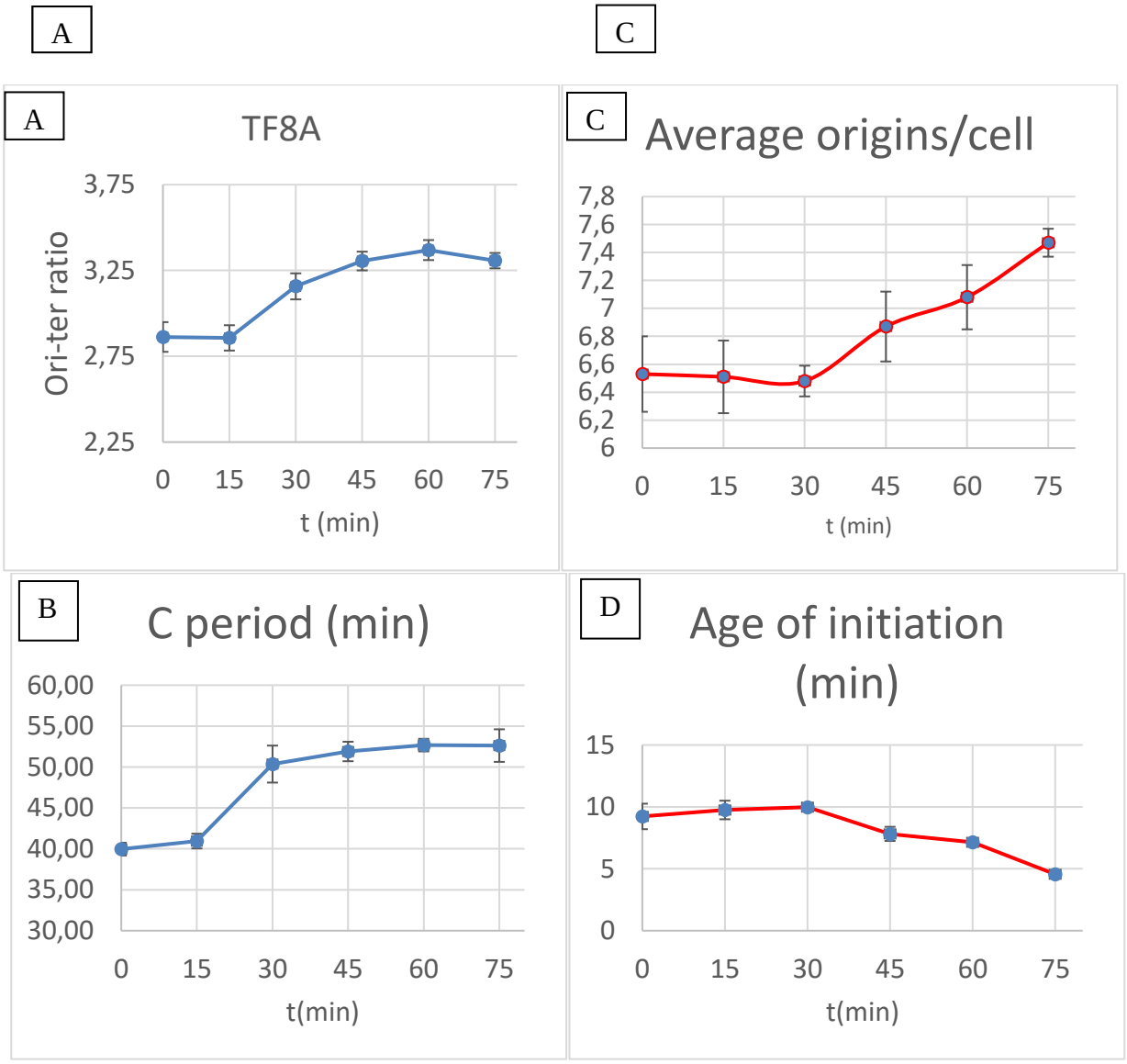


Figure 22: Response of replication parameters during a metabolic shift from MC to MCG media in TF8A

(A) Ori-ter ratio analysis during the shift (B) C-period during the shift (C) Number of origins per cell during the shift (D) Initiation age during the shift. C-period and initiation age were calculated assuming a constant growth rate during the shift, which may not hold true, especially at later stages of the shift.

2. Evidence of PykA mediated coupling between DNA replication timing and metabolism

2.1 Theoretical framework for understanding the results of metabolic shift experiments

Given the unexpected nature of the results in wild type TF8A, it is helpful to develop a theoretical framework to evaluate the results of the metabolic shift experiments. In other words, what kind of *ori-ter* ratio curves would show that PykA couples DNA replication timing to the metabolic state of the cell.

The theoretical framework chosen here is shown in fig. 23.

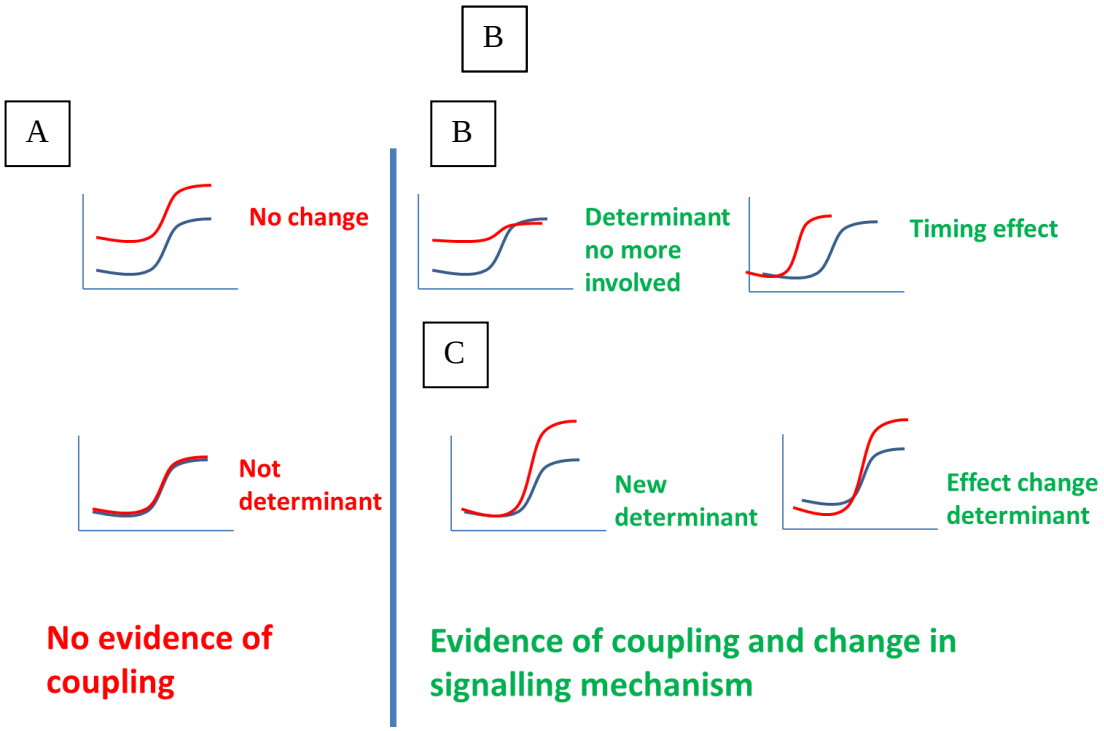


Figure 23: Theoretical framework for evaluating replication parameter curves obtained from metabolic shift experiments (Blue curves) Theoretical WT curve (Red curves) pykA mutant curves. Examples of curves that don't provide evidence for coupling are shown in fig. 22A, whereas examples of curves that doe provide evidence for coupling are shown fig. 22B.

Two types of curves are shown in figure 23 that do not provide evidence for coupling (fig.23A). First, if the effect of a *pykA* mutation does not change throughout the metabolic shift, the replication parameter curve does not provide evidence for a coupling role of PykA. This is similar to the observed effects of *yabA* or *soj* deletion, which have been shown to impact DNA replication similarly in different growth conditions[110]. Trivially, if the curve of a *pykA* mutant is completely identical to wild type, it does not provide evidence for coupling either.

Conversely, any other *pykA* mutant curve would suggest coupling and examples are shown in Fig. 23B and C. Loss of function of a mutation or an effect on the timing of the replication response both provide evidence that the regulatory behaviour of PykA changes in response to the metabolic shift (fig 23B). Moreover, a positive new role for a *pykA* mutant at later stages of the shift indicates that the regions involved in DNA replication timing respond to metabolism and that the PykA signals through a different molecular mechanism in different metabolic contexts (fig. 23C).

In the next sections, the *ori-ter* ratio curves are examined for different *pykA* mutants using this framework. Other replication parameters are scheduled to be measured but were not yet finished at the time of writing.

2.2 Several cat and PEPut mutants lose their moonlighting role at later stages of the metabolic shift, indicating coupling.

The *ori-ter* ratio curves of most cat and PEPut mutants were measured during the metabolic shift and

B

compared to wild type and classified into the categories mentioned in the previous section. These results are shown in figure 24.

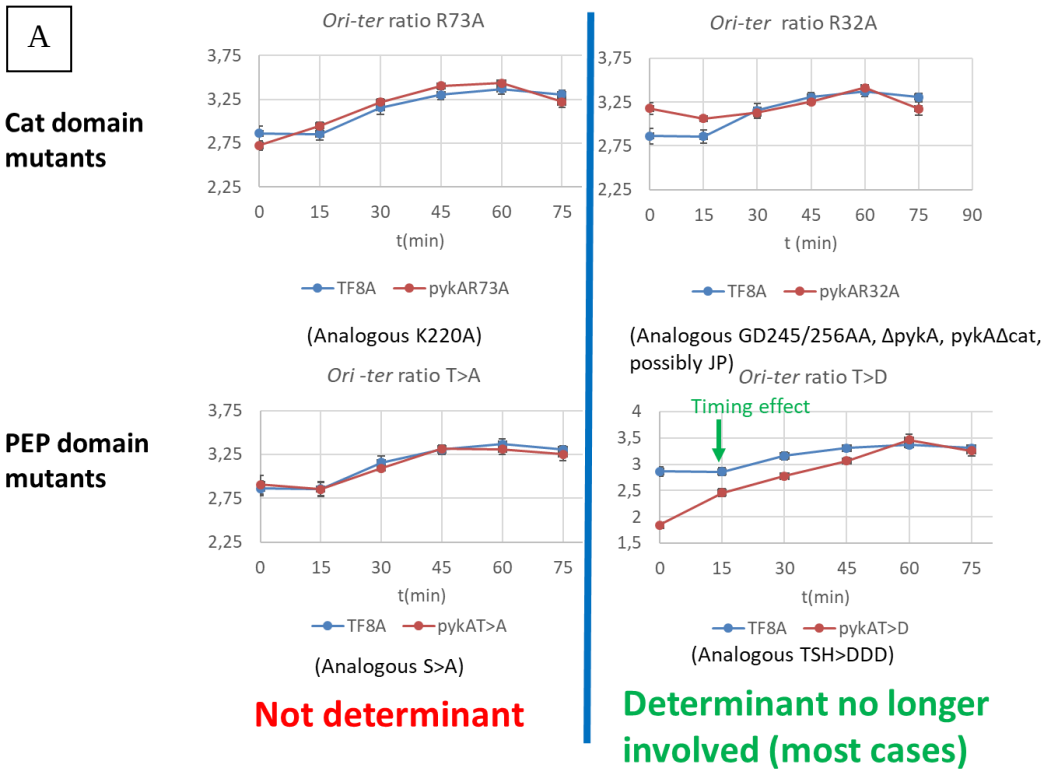


Figure 24: Effect of cat and PEPut mutations on ori-ter ratio during shift : (A) Not determinants : pykA mutation that neither had an effect in MC nor acquired one during the shift. These do not indicate coupling. (B) Determinant no longer involved : these mutations affected ori-ter ratio in MC, but gradually lost their effect during the shift, indicating coupling. In both cases, cat mutations are shown at the top, whereas PEPut mutations were shown at the bottom.

Fig. 24A shows examples of cat and PEPut mutation that were not determinants i.e. they neither affected the *ori-ter* ratio in MC nor at later stages of the shift. Hence, these results do not provide evidence for coupling.

By contrast, many of the cat and PEPut mutations that were determinants in MC are no longer involved at later stages of the shift (fig. 23B). This is the case for the rest of the cat mutants, as well as the T>D and TSH>DDD mutants. As argued in the previous chapter, these results therefore strongly suggest PykA mediated coupling between DNA replication timing and metabolism.

2.3 The signalling mechanism behind PykA mediated DNA replication gating responds to metabolism: a new role for phosphorylated H539.

A few mutations in the TSH motif did not fall in the categories from section 2.2 i.e. their effect on replication changed during the shift. These results are summarized in figure 25.

Figure 25: A changing role for the TSH motif on DNA replication gating during the shift

Surprisingly, the effect of TSH>AAA lowers the *ori-ter* ratio in MC, but causes an overshoot at later stages of the shift, indicating a changing involvement of the TSH motif at later stages of the shift. Replacement of individual amino acids by alanine revealed that H539 and not T537 was responsible for the overshoot phenotype, i.e. H539A is a new determinant at late stages of the shift. Replacement of H539 by aspartate suppressed the effect of H539A, indicating that H539 phosphorylation plays a key role in DNA replication gating during the shift. Furthermore, the effect of H539D does not change during the shift.

3. Chapter summary

In the previous chapter, PykA was demonstrated to moonlight in DNA replication gating in ways that could not be attributed to its metabolic activity. Here, it was shown that the moonlighting PykA, unlike its traditional counterparts YabA and Soj, couples its DNA replication gating to metabolic signals. This was demonstrated by studying the effect of *pykA* mutations on DNA replication parameters during a metabolic shift from the rich gluconeogenic medium MC to the MCG medium. First, the dynamic response of DNA replication parameters in wild type TF8A was assessed. This response was found to consists of multiple phases involving (i) no change of replication parameters during the first 15 min of the shift, (ii) an increase of the C-period between 15 and 45 min, (iii) a change in initiation parameters after 30 min and (iv) a change in cell size/growth rate sometime after 45 min after the shift.

Next, a theoretical framework was developed to infer coupling by comparing the *ori-ter* ratio curves of different *pykA* mutants to wild type. Using this framework, most cat and PEPut mutants were shown to lose their moonlighting effects on DNA replication gating at later stages of the shift, indicating that PykA regulatory activity is coupled to metabolism. Furthermore, the role of the TSH motif changed during the shift and phosphorylation of H539 became important at later stages of the shift. This suggests that the signalling mechanism behind PykA mediated DNA replication gating also depends on metabolism. Taken together, these results suggest for the first time that PykA, unlike its classical counterparts YabA and Soj, couples its activity to the metabolic state of the cell.

Chapter 4: Discussion

The goal of this research project was to demonstrate that (i) PykA moonlights in DNA replication gating beyond its metabolic activity and (ii) that PykA can communicate information about the metabolic state of the cell to the replication machinery.

1. PykA moonlights in DNA replication gating

To demonstrate that PykA has a role in regulating replication gating beyond its metabolic activity, the effect of *pykA* mutations on replication control was analyzed in a rich gluconeogenic medium (MC) in which PykA is strongly expressed[207–209], but not required for metabolism. Using this approach, the results presented in this work put forward 5 main arguments that the effect of PykA on DNA replication gating can't be explained by its metabolic activity.

First, deletion of *pykA* impacted replication speed and initiation parameters without impacting growth rate in the MC medium. Second, mutations in the PEPut domain and its conserved TSH motif, which are dispensable for the metabolic activity of PykA[220], affected replication speed and initiation parameters. Third, ectotopic expression of free PEPut affected *ori-ter* ratio, even in a wild type context. Fourth, precise mutagenesis in the *pykA* catalytic site revealed that some, but not all mutations that inhibit catalytic activity also lead to defects in replication gating. Finally, purified PykA could modulate DnaE polymerase and DnaC helicase activity *in vitro*. Taken together, these results strongly suggest that the effect of PykA on DNA replication gating is not due to its metabolic activity, but is indicative of a moonlighting function, similar to the main function of the proteins YabA[43] and Soj[47,48].

In support of this hypothesis, non-canonical functions of CCM enzymes (moonlighting functions) are increasingly being found to play a major role in cellular decision making across the tree of life, in particular in the development of cancer (see for a review [234,235]).

For example, the moonlighting functions of pyruvate kinase isoform 2 (PKM2) in humans have been subject of intense research due to its perceived connections with tumorigenesis (for a recent review, see [236]). Recent results have directly associated PKM2 to the regulation of cell cycle progression[203,237,238], and DNA repair[239], although direct stimulation of replication has not yet been demonstrated to the best of our knowledge. Additionally, PKM2 is involved in the other hallmarks of cancer as well including promotion of angiogenesis[240], regulation of apoptosis[241], immunity[242] or growth factor regulation[243]. As a side note, several studies also indicate that PKM2 is excreted and used for cell-cell signalling in humans, promoting growth, migration and matrix adhesion of endothelial cells [244], enhancing cell migration in colon cancer cells [245], promoting wound healing [246] and associating with MHC complex of B-cells [242]. PKM2 is but one of many examples of moonlighting CCM enzymes[234], but it illustrates the extent to which CCM enzymes are increasingly found at the heart of cellular decision making. Hence, a moonlighting role for PykA in DNA replication gating is consistent with this growing body of work.

2. Mechanisms behind PykA moonlighting

Though still incomplete, this work has also generated insight into the mechanisms behind PykA moonlighting.

2.1 Mechanisms behind PEPut moonlighting

PykA moonlighting in DNA replication gating occurs in part through its PEPut domain. Despite many speculations about its role prior to this work[129,219], this is the first work to identify a role for the domain, hence it is also the first to provide some insights into its functioning. Our results confirmed the speculated importance of the conserved TSH motif and the effects of phosphomimic mutations in H, S and especially T suggest that phosphorylation of the TSH motif may play a role in PEPut functioning. This is consistent with observations in PEPut from EI, PEPS and PPDK, where these residues were found to be phosphorylated as well[223–226] . In these examples, H is (auto)phosphorylated using PEP as a donor[223] and T using AMP[226]. However, the absence of any strong effects for alanine mutations in the T, S and H residues suggests that the PEP utilizer is only phosphorylated in a limited way under MC conditions.

Ectotopic expression of mutated *hyPEPut* largely confirmed these results. Mutation of the T in *hyPEPut* confirmed the results obtained for wild type (*hyPEPut*$_{T>D}$ and *hyPEPut*$_{T>A}$). Additionally, *hyPEPut* does require H in its TSH motif (*hyPEPut*$_{H>A}$ and *hyPEPut*$_{H>D}$), underscoring the importance of this residue as well.

Importantly, H>A does not have the same effect on replication gating than T>D. This suggests that T>D affects replication beyond regulating the phosphorylation on H as is observed for plant PPDK and PEPS PEP utilizers[225]. Even though more research is needed to fully understand the mechanisms behind PEP utilizer functioning, these initial results indicate that the mechanism might be different from what is known from EI, PPDK and PEPS.

In summary, several lines of evidence suggest that PEPut is involved in the regulation of DNA replication gating and that its activity depends on phosphorylation events of its TSH motif, particularly involving T and H. This is visualized in figure 26.

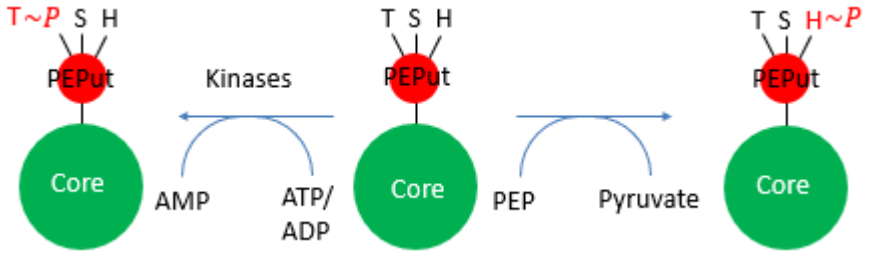


Figure 26: Scheme of PEPut regulation

2.2 Mechanisms behind cat moonlighting

Additionally, the cat domain of PykA also seems to have independent moonlighting functions in DNA replication gating. PEP binding mutants and deletion of the catalytic domain affected replication speed, initiation frequency and the age of initiation. Therefore, these results might mean that PEP binding is required for cat moonlighting. As pointed out before, the PEP/pyruvate ratio is an important indicator of metabolic activity[137,138] that controls PTS activity as well[136,137]. These results thus suggest, but do not prove, that the activity of the catalytic domain in replication control is coupled to the PEP/pyruvate ratio and the metabolic state of the cell. Furthermore, the effect of the JP mutation on DNA replication parameters confirms its previously observed involvement in DNA replication[162].

Additionally, the *in vitro* effects of PykA on DNA replication suggest that PykA might modulate DNA replication through direct protein-protein interactions. Like PykA, there are indications that PKM2 can regulate its targets through protein-protein interactions. PKM2 interacts with important actors for signal transduction and gene expression such as tyrosine kinase receptors like A-Raf and Fibroblast growth factor receptor 1 (FGFR1) [243,247,248] and RNA binding proteins such as HuR and tritetraprolin [249,250] with documented consequences for tumour development. Additionally, PKM2 is also known to translocate into the nucleus and interact with transcription factors like NF-kB and HIF-1α to promote angiogenesis there [240].

In summary, the catalytic domain of PykA is also though to moonlight in replication through direct protein-protein interactions and this activity is thought to depend on PEP binding.

2.3 Functional interplay between cat and PEPut moonlighting

Conflicting results were obtained about the interplay between cat and PEPut moonlighting. On the one hand, ectotopically expressed PEPut does not require the cat domain to interfere with replication (*hyPEPut ΔpykA*). However, ectotopic expression of PEPut in a background with only the catalytic domain did not recover the gating phenotype, suggesting that PEPut needs to be fused to the cat domain to function properly (*hyPEPut pykAΔPEPut*). In line with that, the JP mutation could fully suppress the effects of ectotopic expression of *hyPEPut*, even when the latter was mutated T>D (*pykAJP hyPEPut* and *pykAJP hyPEPut*$_{T>D}$). This functional connection between JP and the TSH motif might suggest that the contact between PEPut and the cat domain is also important for moonlighting, since both mutations are near the interface between PEPut and the cat domain[214]. More research is needed to better understand these results.

2.4 A model for PykA moonlighting in DNA replication gating control

The extensive genetic analysis performed in this work has enabled the creation of a working model of the PykA moonlighting system, shown in figure 27. In this model, PEP and possibly AMP are reused as sentinel molecules to determine PykA moonlighting activity. This is consistent with the role of PEP in regulating CCM activity[129]. Both the cat and the PEP domains act as sensors of metabolic activity. They convey the metabolic status of the cell sent by the sentinel molecules to the initiation and elongation machinery via phosphorylation events and protein-protein interactions, similar to how CcpA/HpR senses metabolic information to drive metabolic gene expression[129]. This regulatory system involves PEP binding and the JP region of Cat, the TSH residues of PEPut and kinases. Together, these regulatory signals precisely gate replication in the cell cycle and may be part of the metabolic control of replication.

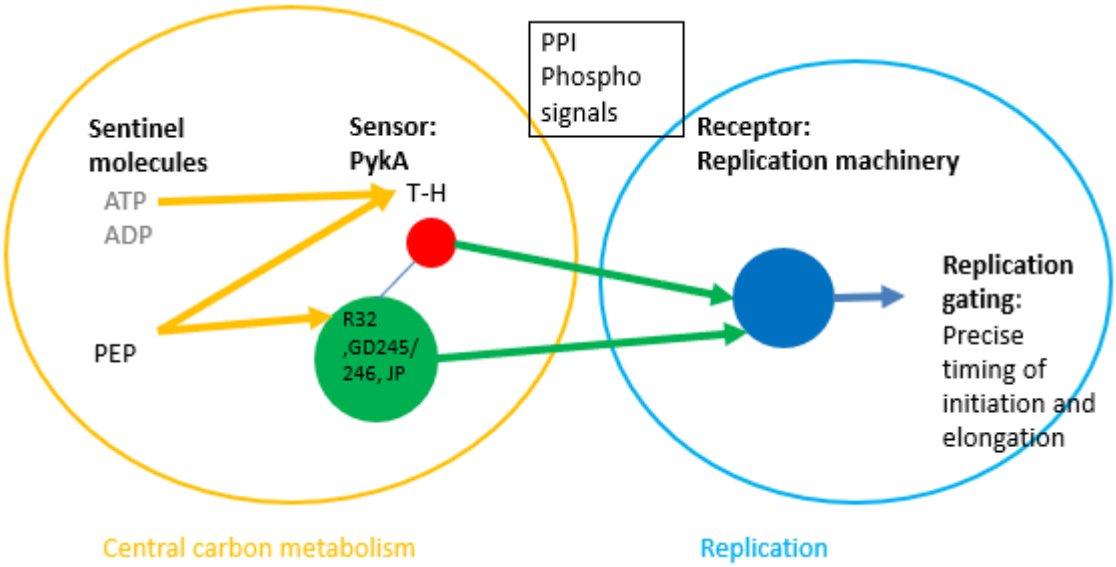


Figure 27: A model for PykA moonlighting in DNA replication gating : PykA senses metabolic information using the sentinel metabolites AMP and PEP. The cat and PEPut domain then transduce this signal to the replication machinery using PPI and phosphorylation events, leading to precise DNA gating

2.5 Future directions

Despite these advances, many questions remain about the PykA signaling system discovered in this work.

To demonstrate moonlighting more strongly, the effect of *pykA* mutations on the physiology of *B. subtilis* in MC needs to be examined in greater detail. The goal of this work was to demonstrate that the effect of PykA on DNA replication gating was due to moonlighting beyond its metabolic activity. OD measurements of all *pykA* mutations as well as a competitive growth experiment of

ΔpykA suggest that no mutation significantly affects *B. subtilis* physiology under MC conditions, but small effects on physiology can't be ruled out by this method and might explain some of the weaker effect sizes, especially for mutants in the cat domain. However, the fact that some mutants, but not all affect replication suggests that the replication phenotype is distinct from catalysis and strong effect size PEPut mutations are unlikely to be explained through possible weak effects on physiology anyway. Regardless, the metabolic state of several key *pykA* mutants will be assayed using mass spectrometry to strengthen this point, which is reviewed here[251].

Additionally, the PykA signaling pathway remains largely unknown. This work has not clearly identified the targets of PykA nor the metabolic sensing mechanism.

First, the replication target of PykA (initiation/elongation) needs to be established. The genetic analysis presented in this work was very effective at (i) demonstrating the existence of the moonlighting system (ii) identifying important regions in the PykA protein for replication control and (iii) determining how they are functionally connected. However, as stated in section 1.1 of the introduction, proper temporal control of cell cycle events is vital for cell fitness. Therefore, defects in the timing of cell cycle events create an evolutionary advantage for secondary effects that restore proper timing. For this reason, the genetic analysis presented here is inherently insufficient to figure out the exact replication target and effects of PykA. This limitation can be overcome by testing the effect of different *pykA* mutations on *ori-ter* ratio in an *oriN ΔoriC* and *oriN ΔdnaA* background, as done in our previous work. [86]

Next, concrete molecular regulatory targets of PykA need to be established as well. These can partially be found through temperature sensitive complementation assays in *pykA* mutants [162]. Moreover, performing these experiments in MC bypasses any potential physiological effects of *pykA* mutations, making it a more powerful tool than before. Since our *in vitro* results suggest that PykA moonlighting might operate through direct protein-protein interactions, a good start would also be to identify potential interaction partners. Protein-protein interactions can be inferred from evolutionary information, sequence features and structural features using a variety of computational methods, reviewed in [252]. Additionally, many experimental procedures have been developed to detect and validate protein-protein interactions, both *in vitro* (NMR, isothermal calorimetry, analytical ultracentrifugation etc) and *in vivo* (fluorescence colocalization, FRET, FCCS, yeast-2-hybrid, etc)[253].

Finally, the fact that PykA senses metabolic information (i.e. coupling between DNA replication gating and metabolism) and the mechanisms by which it does this remain to be established as well. Coupling DNA replication gating can be demonstrated either by studying the effect of *pykA* mutations in different growth conditions[110] or by studying their effect on a metabolic shift (see below). *In vitro* and *in vivo* tests of protein-protein interactions in PEP binding mutants could be

used to further establish a sentinel role for PEP. Alternatively, studying the effect of varying PEP concentration on replication *in vitro* demonstrates that as well. TSH phosphorylation has to be established, using *in vivo* and *in vitro* methods[254] and the mechanisms responsible for this phosphorylation have to be identified.

In the next part, the question about PykA coupling DNA replication gating to metabolism is taken up.

3. PykA couples its moonlighting activity in DNA replication gating control to metabolism

The goal of the second part of the thesis was (i) to demonstrate that PykA couples metabolism to replication and (ii) identify the role that the different replication determinants play in this coupling. To determine coupling, the effect of a metabolic shift from the gluconeogenic MC to the glycolytic MCG medium on *ori-ter* ratio in wild type and *pykA* mutants was assessed. As stated before, the importance of replication control proteins for cellular fitness may cause secondary compensatory mutations to be accumulated during overnight growth in liquid cultures. The metabolic shift approach presented here avoids this problem by studying in real time the effect of the metabolic shift, before secondary mutations can occur.

Before analyzing the effect of the metabolic shift on *ori-ter* ratio in *pykA* mutants, it was determined in a wild type strain cured from prophages (TF8A). Surprisingly, the adaptation of DNA replication parameters did not occur instantaneously after the addition of glucose, but consisted of four overlapping phases. (I) Between 0 and 15 min after the shift, none of the cell cycle parameters changed. (II) Between 15 and 45 minutes, the replication speed dropped and *ori-ter* ratio increases. (III) Between 30 and 45 minutes after shift, the initiation frequency increased and the age of initiation advanced. Finally, (IV) sometime after 45 min, the cell size (a proxy for growth rate in this work) increased. To the best of our knowledge, this complex response of DNA replication parameters to a metabolic shift has not been described in literature and merits further investigation. It is also worth pointing out that the assumption that cell size is a good proxy for cell growth (i) is a law derived in steady state conditions[75] and thus(i) may not hold and (ii) may not hold in *pykA* mutants. For this reason, this work uses the shift as a tool to investigate coupling between DNA replication timing and metabolism by PykA, even though a full investigation of the phenotype is of interest.

Next, a framework was established to interpret *ori-ter* ratio curves of *pykA* mutants and determine whether or not they indicate that PykA couples DNA replication gating to metabolism. The work of Murray and Koh was used as an inspiration to decide whether or not PykA couples DNA replication timing to metabolism[110]. In order to decide whether or not YabA and Soj were able to couple DNA

replication timing to metabolism, the *ori-ter* ratio of deletion mutants of these two proteins was analyzed in different media by Murray and Koh[110]. They found that the effect of the deletions remained coupled to the growth rate in the media they tested and concluded that therefore the function of YabA and Soj was independent from metabolism[110]. Thus, if the *ori-ter* ratio curve of a *pykA* mutant is either identical to wild type or runs parallel to it, the *ori-ter* ratio curve does not count as evidence for coupling and any other situation suggests coupling.
Using this framework, we classified the *ori-ter* ratio curves of all the *pykA* mutants in different categories.
The first category held *pykA* mutants that had curves identical to wild type and did not have a moonlighting role in MC nor at any other point during the shift. This included the cat mutants K220A and R73A and the PEPut mutants T>A and S>A. Hence, these mutants do not provide evidence for coupling. The second category held *pykA* mutants that had an effect in MC, but lost that effect at later stages of the shift. This group included *ΔpykA, Δcat*, the PEP binding mutants in the cat domain and T>D and TSH>DDD in the PEPut domain. The third category held *pykA* mutants that acquired a new effect during the metabolic shift. This group included TSH>AAA and H>A. Based on our framework, we decided that the second and third category provided strong evidence for PykA mediated coupling between DNA replication gating and metabolism. This behavior is not described for any replication control protein in literature to our knowledge. All in all, this work presents evidence that PykA can communicate metabolic signals to the replication machinery, making PykA a new type of regulator of DNA replication.

4. Mechanisms behind PykA coupling metabolism to DNA replication gating

While this part of the research is in its infancy, the coupling experiments have generated some new insights into the PykA moonlighting mechanism.

4.1 H phosphorylation becomes important during the shift

Interestingly, coupling experiments showed that TSH>AAA and H>A acquire a new effect at later stages of the shift and that H>D suppresses this. This suggests that, while H phosphorylation doesn't play a major role in MC, it does at later stages of the shift and confirms that H is important for PykA moonlighting.
Some observations suggest that PKM2 moonlights as a protein kinase and phosphorylate histones [255], transcription factors [256] and other proteins in the nucleus [237,238,257]. These last results however are in conflict with other studies in mouse embryonic cells, casting doubt on the protein kinase function of PKM2 [258]. Additionally, PEP utilizers like the one in EI, PPDK or PEP synthase are

known to transfer phosphates from one side of the protein to another using their conserved histidine[223]. However, this work does not support a role for PykA as a protein kinase, since H539D seems to be able to suppress the effect of H539A. If the mechanism were linked to phosphoryl transfer, H539D was expected to have the same effect as H539A. Phosphoryl transfer by H therefore doesn't seem to be required for the effect of H on replication during the shift. The effects of phosphomimic mutants do however indicate that phosphorylation plays an important role in PykA moonlighting activity and these results still establish a metabolic condition in which H is likely phosphorylated.

4.2 Future directions

Despite the advances made in this chapter, much work remains to be done.
A weakness of this coupling part of the work is that the effect of the metabolic shift on cell physiology in general and the growth rate in particular are insufficiently understood. This is particularly problematic for catalytic mutations of *pykA*, since they are expected to impact physiology in glycolytic media such as MCG. However, given that PEPut mutations are not expected to impact metabolism under any condition[220], it is reasonable to assume that comparison of the curves with wild type still yields meaningful insights, even if the consequences of the metabolic shift are not fully understood. Particularly, the effect of the metabolic shift on initiation parameters and growth rate needs to be established.
The majority of *ori*-ter ratio curves presented in this work seem to indicate that the outcome of the PykA signaling system depends on metabolic signals. It is currently unclear how this is achieved, but it may involve a structural transition in the PykA monomer/dimer/tetramer. The translocation of PKM2 inside the nucleus is accompanied by a switch in oligomerization state from metabolically active tetramer to non-metabolically active dimer[259]. A structural transition in PykA induced by metabolism could also explain why the same mutation (e.g. T>D or H>A) has an enormous effect in one medium, but not the other.

5. Theoretical implications of CCM enzymes coupling DNA replication gating to metabolism

5.1 Implications for the coordination between replication, growth and the availability of nutrients

Currently, several groups have started to argue that growth rate dependent control of replication timing is a multifactorial process, which may involve sensing cellular metabolism and communicating this information to the replication machinery[10,70,86–88]. However, all of the mechanisms that were proposed to be responsible for this coupling have been challenged in some

form [110–112], especially under fast growth conditions (see introduction section 1.4 for a discussion). The discovery that PykA communicates metabolic signals to the replication machinery in fast growth conditions therefore supports the view that growth rate dependent control of replication over a wide range of nutritional conditions involves sensing cellular metabolism[10,70,86–88] and establishes for the first time a mechanism by which this may occur. Furthermore, a previous publication of our group identified multiple, medium dependent links between CCM enzymes and replication[86]. This suggests that other CCM enzymes may also communicate metabolic information to the replication machinery, leading to the hypothesis that growth rate dependent control of DNA replication is achieved at least in part through non-metabolic activities of CCM enzymes. This hypothesis is likely important for understanding replication in a wide range of organisms since (i) CCM enzymes are strongly conserved throughout the tree of life, (ii) CCM-replication links have been found in diverse model organisms, including *E. coli*, *S. cerevisae* and even human cells (see introduction section 2.3 for a full discussion) and (iii) numerous non-metabolic functions of CCM enzymes are described in a wide range of organisms (for reviews, see [234,235]). This hypothesis will later be referred to as the CCM-coupling hypothesis. Hence, this work extends the knowledge on the long enigmatic coupling between DNA replication gating and metabolism by providing the first example of a CCM enzyme responsible for it and pointing to other CCM enzymes as likely future targets. Furthermore, the CCM-coupling hypothesis has important implications for understanding and treating cancer, which is discussed in the next section.

5.2 Theoretical implications of the CCM coupling hypothesis for the origin of cancer

The CCM-coupling hypothesis has important implications for our understanding of the origin of cancer. The Warburg hypothesis was the first major theory about the origin of cancer[260]. It was based on the observation that many tumors have a high rate of glucose consumption and excrete lactate, instead of fully oxidizing it using the TCA cycle[260]. This led him to conclude that cancer cells originate from dysfunctions in the mitochondria[260]. This theory has later been disproven on the basis that most tumor cells also have functional respiration, but the debate on the reason behind the observation of disrupted metabolism in most tumor cells continues to date[261,262]. Nowadays, cancer is thought to require the simultaneous acquisition of many mutations to obtain the so-called hallmarks of cancer[263]. These include sustained proliferation, evasion of immune responses, evasion of growth factors, evasion of induced cell-death, induction of angiogenesis and activation of metastasis[264]. The chance of this occurring in a healthy cell is incredibly small, which is why cancer nowadays is thought to originate from mutations that promote further mutagenesis i.e. genetic instability[263,265]. Genetic instability in turn is caused by mutations in any process that is

required for faithful transmission of genetic information, such as DNA repair, DNA replication (see [266]), chromosome segregation, cell cycle checkpoint control and others[265]. The CCM-coupling hypothesis therefore suggests that cancer could originate from defects in the metabolic signaling pathways that couple these processes to metabolism as shown in figure 28.

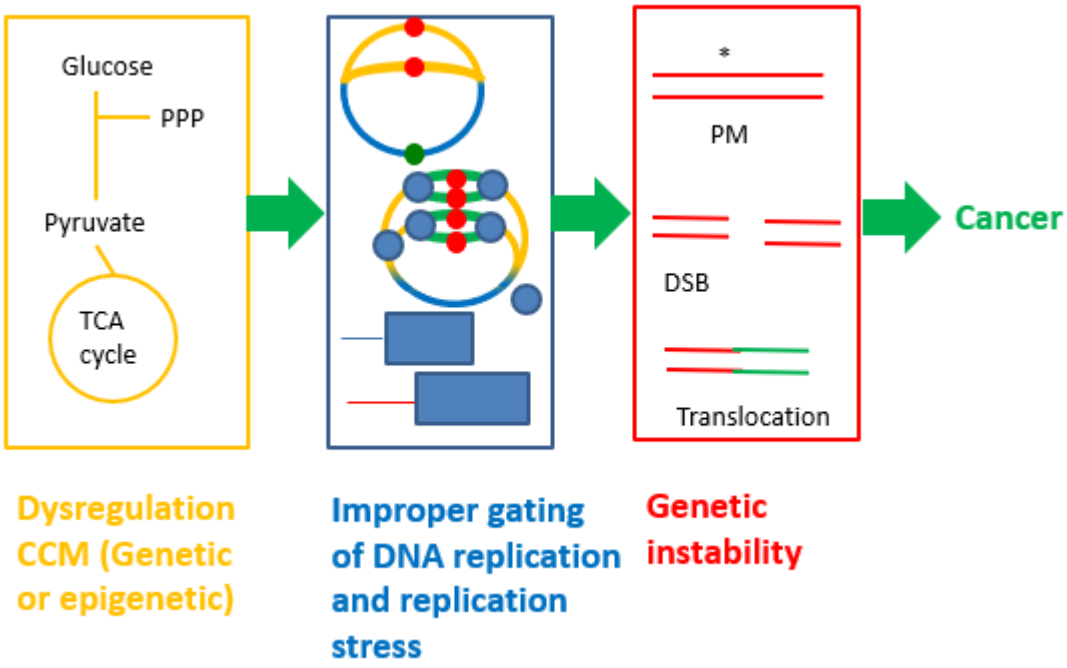


Figure 28: Coupling CCM-DNA replication may lead to cancer

To conclude, the work presented in this thesis provides evidence for the first time not only that PykA in *B. subtilis* is a replication control protein, but also that it couples DNA replication timing to metabolism. It also identified for the first time a role for the long enigmatic PEP utilizer domain. While more research is needed to understand the details of PykA signalling to replication, these results are consistent with an emerging body of work that puts non-metabolic functions of CCM enzymes at the heart of cellular decision making and the development of cancer. This emerging understanding of cell biology will likely lead to exciting new treatment options, such as the therapeutic targeting of non-metabolic functions of CCM enzymes (See [267]) or the improvement of innovative treatments such as immunotherapy[268] and manipulation of the microbiome[269].

Bibliography

1. Vermeulen, K., Van Bockstaele, D. R. & Berneman, Z. N. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif.* **36**, 131–149 (2003).
2. Lodish, H., Berk, A. & Zipursky, S. Overview of the Cell cycle and its control. in *Molecular Cell Biology* vol. 13.1 (W.H.Freeman, 2000).

3. Brehm, A. *et al.* Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597–601 (1998).
4. Voitenleitner, C., Fanning, E. & Nasheuer, H.-P. Phosphorylation of DNA polymerase a-primase by Cyclin A-dependent kinases regulates initiation of DNA replication in vitro. *Oncogene.* **14**, 1611-1615 (1997).
5. Siliciano, J. D. *et al.* DNA damage induces phosphorylation of the amino terminus of p53. *Genes Dev.* **11**, 3471–3481 (1997).
6. Adams, D. W. & Errington, J. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat. Rev. Microbiol.* **7**, 642–653 (2009).
7. Badrinarayanan, A., Le, T. B. K. & Laub, M. T. Bacterial Chromosome Organization and Segregation. *Annu. Rev. Cell Dev. Biol.* **31**, 171–199 (2015).
8. Reyes-Lamothe, R. & Sherratt, D. J. The bacterial cell cycle, chromosome inheritance and cell growth. *Nat. Rev. Microbiol.* **17**, 467–478 (2019).
9. Robert, L. Size sensors in bacteria, cell cycle control, and size control. *Front. Microbiol.* **6**, (2015).
10. Boye, E. & Nordström, K. Coupling the cell cycle to cell growth: A look at the parameters that regulate cell-cycle events. *EMBO Rep.* **4**, 757–760 (2003).
11. Kreuzer, K. N. DNA Damage Responses in Prokaryotes: Regulating Gene Expression, Modulating Growth Patterns, and Manipulating Replication Forks. *Cold Spring Harb. Perspect. Biol.* **5**, a012674–a012674 (2013).
12. Lindås, A.-C. & Bernander, R. The cell cycle of archaea. *Nat. Rev. Microbiol.* **11**, 627–638 (2013).
13. Grogan, D. W., Bernander, R. & Poplawski, A. Altered patterns of cellular growth, morphology, replication and division in conditional-lethal mutants of the thermophilic archaeon Sulfolobus acidocaldarius. *Microbiology* **146**, 749–757 (2000).
14. Hjort, K. & Bernander, R. Cell cycle regulation in the hyperthermophilic crenarchaeon

Sulfolobus acidocaldarius: S. acidocaldarius cell cycle. *Mol. Microbiol.* **40**, 225–234 (2001).

15. Gabrielli, B., Brooks, K. & Pavey, S. Defective Cell Cycle Checkpoints as Targets for Anti-Cancer Therapies. *Front. Pharmacol.* **3**, (2012).
16. Jameson, K. & Wilkinson, A. Control of Initiation of DNA Replication in Bacillus subtilis and Escherichia coli. *Genes* **8**, 22 (2017).
17. Fuller, R. S., Funnell, B. E. & Kornberg, A. The dnaA Protein Complex with the E. coli Chromosomal Replication Origin (OK) and Other DNA Sites. *Cell* **38**, 889–900 (1984).
18. Zorman, S., Seitz, H., Sclavi, B. & Strick, T. R. Topological characterization of the DnaA–oriC complex using single-molecule nanomanipuation. *Nucleic Acids Res.* **40**, 7375–7383 (2012).
19. Erzberger, J. P., Mott, M. L. & Berger, J. M. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* **13**, 676–683 (2006).
20. Kowalski, D. & Eddy, M. J. The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. *EMBO J.* **8**, 4335–4344 (1989).
21. Bramhill, D. & Kornberg, A. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome. *Cell* **52**, 743–755 (1988).
22. Bruand, C., Ehrlich, S. D. & Jannière, L. Primosome assembly site in Bacillus subtilis. *EMBO J.* **14**, 2642–2650 (1995).
23. Zhang, W. *et al.* The Bacillus subtilis DnaD and DnaB Proteins Exhibit Different DNA Remodelling Activities. *J. Mol. Biol.* **351**, 66–75 (2005).
24. Smits, W. K., Goranov, A. I. & Grossman, A. D. Ordered association of helicase loader proteins with the *Bacillus subtilis* origin of replication *in vivo*. *Mol. Microbiol.* **75**, 452–461 (2010).
25. Fang, L., Davey, M. J. & O'Donnell, M. Replisome Assembly at oriC, the Replication Origin of E. coli, Reveals an Explanation for Initiation Sites outside an Origin. *Mol. Cell* **4**, 541–553 (1999).

26. Reyes-Lamothe, R., Sherratt, D. J. & Leake, M. C. Stoichiometry and Architecture of Active DNA Replication Machinery in Escherichia coli. *Science* **328**, 498–501 (2010).

27. Robinson, A., J. Causer, R. & E. Dixon, N. Architecture and Conservation of the Bacterial DNA Replication Machinery, an Underexploited Drug Target. *Curr. Drug Targets* **13**, 352–372 (2012).

28. Beattie, T. R. & Reyes-Lamothe, R. A Replisomeâ€™s journey through the bacterial chromosome. *Front. Microbiol.* **6**, (2015).

29. Voet, D. & Voet, J. DNA Replication, Repair, and Recombination. in *Biochemistry* 1171–1259 (Wiley, 2011).

30. Corn, J. E., Pelton, J. G. & Berger, J. M. Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. *Nat. Struct. Mol. Biol.* **15**, 163–169 (2008).

31. Corn, J. E. & Berger, J. M. Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions. *Nucleic Acids Res.* **34**, 4082–4088 (2006).

32. Sanders, G. M., Dallmann, H. G. & McHenry, C. S. Reconstitution of the B. subtilis Replisome with 13 Proteins Including Two Distinct Replicases. *Mol. Cell* **37**, 273–281 (2010).

33. Hill, T. M. Arrest of Bacterial DNA Replication. *Annual Reviews on Microbiology **46***, 602-633 (1992).

34. Lewis, P. J., Ralston, G. B., Christopherson, R. I. & Wake, R. G. Identification of the Replication Terminator .Protein Binding Sites in the Terminus Region of the Bacillus subtilis Chromosome and Stoichiometry of the Binding. *Journal of Molecular Biology*. **214**, 73-84 (1990).

35. Vivian, J. P., Porter, C. J., Wilce, J. A. & Wilce, M. C. J. An Asymmetric Structure of the Bacillus subtilis Replication Terminator Protein in Complex with DNA. *J. Mol. Biol.* **370**, 481–491 (2007).

36. Langley, D. B., Smith, M. T., Lewis, P. J. & Wake, R. G. Protein?nucleoside contacts in the interaction between the replication terminator protein of Bacillus subtilis and the DNA

terminator. *Mol. Microbiol.* **10**, 771–779 (1993).

37. Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H. & Ogasawara, N. Cloning of an autonomously replicating sequence (*ars*) from the *Bacillus subtilis* chromosome. *Mol. Microbiol.* **6**, 309–315 (1992).
38. Dervyn, E. Two Essential DNA Polymerases at the Bacterial Replication Fork. *Science* **294**, 1716–1719 (2001).
39. Hidaka, M., Kobayashi, T., Takenaka, S., Takeya, H. & Horiuchi, T. Purification of a DNA Replication Terminus (ter) Site-binding Protein in Escherichia coli and Identification of the Structural Gene. *J. Biol. Chem.* **267**, 21031–21037 (1989).
40. Hill, T. M., Henson, J. M. & Kuempel, P. L. The terminus region of the Escherichia coli chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc. Natl. Acad. Sci.* **84**, 1754–1758 (1987).
41. Neylon, C., Kralicek, A. V., Hill, T. M. & Dixon, N. E. Replication Termination in Escherichia coli: Structure and Antihelicase Activity of the Tus-Ter Complex. *Microbiol. Mol. Biol. Rev.* **69**, 501–526 (2005).
42. Hayashi, M., Ogura, Y., Harry, E. J., Ogasawara, N. & Moriya, S. *Bacillus subtilis* YabA is involved in determining the timing and synchrony of replication initiation. *FEMS Microbiol. Lett.* **247**, 73–79 (2005).
43. Felicori, L. *et al.* Tetramerization and interdomain flexibility of the replication initiation controller YabA enables simultaneous binding to multiple partners. *Nucleic Acids Res.* **44**, 449–463 (2016).
44. Cho, E., Ogasawara, N. & Ishikawa, S. The functional analysis of YabA, which interacts with DnaA and regulates initiation of chromosome replication in Bacillus subtils. *Genes Genet. Syst.* **83**, 111–125 (2008).
45. Merrikh, H. & Grossman, A. D. Control of the replication initiator DnaA by an anti-cooperativity factor: Control of DnaA by an anti-cooperativity factor. *Mol. Microbiol.* **82**,

434–446 (2011).

46. Scholefield, G., Errington, J. & Murray, H. Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA: Soj inhibits DnaA helix formation. *EMBO J.* **31**, 1542–1555 (2012).
47. Scholefield, G., Whiting, R., Errington, J. & Murray, H. Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation: Oligomeric status switches Soj regulatory activity. *Mol. Microbiol.* **79**, 1089–1100 (2011).
48. Murray, H. & Errington, J. Dynamic Control of the DNA Replication Initiation Protein DnaA by Soj/ParA. *Cell* **135**, 74–84 (2008).
49. Leonard, T. A., Butler, P. J. & Löwe, J. Bacterial chromosome segregation: structure and DNA binding of the Soj dimer - a conserved biological switch. *EMBO J.* **24**, 270–282 (2005).
50. Okumura, H. *et al.* Regulation of chromosomal replication initiation by oriC-proximal DnaA-box clusters in Bacillus subtilis. *Nucleic Acids Res.* **40**, 220–234 (2012).
51. Ishida, T. *et al.* DiaA, a Novel DnaA-binding Protein, Ensures the Timely Initiation of *Escherichia coli* Chromosome Replication. *J. Biol. Chem.* **279**, 45546–45555 (2004).
52. Terradot, L. & Zawilak-Pawlik, A. Structural insight into *Helicobacter pylori* DNA replication initiation. *Gut Microbes* **1**, 330–334 (2010).
53. Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E. & Leonard, A. C. Escherichia coli prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA: Regulation of bacterial pre-RC assembly. *Mol. Microbiol.* **51**, 1347–1359 (2004).
54. Cassler, M. R., Grimwade, J. E. & Leonard, A. C. Cell cycle-specific changes in nucleoprotein complexes at a chromosomal replication origin. *EMBO J.* **14**, 5833–5841 (1995).
55. Grimwade, J. E., Ryan, V. T. & Leonard, A. C. IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of Escherichia coli. *Mol. Microbiol.* **35**, 835–844 (2000).
56. Rasmussen, K. V. & Schaechter, M. SeqA limits DnaA activity in replication from oriC in Escherichia coli. *Mol. Microbiol.* **14**, 763–772 (1994).

57. Lu, M. SeqA: A negative modulator of replication initiation in E. coli. *Cell* **77**, 413–426 (1994).

58. Campbell, J. L. & Kleckner, N. E. coli oriC and the dnaA Gene Promoter Are Sequestered from dam Methyltransfferase Following the Passage of the Chromosomal Replication Fork. *Cell* **62**, 967–979 (1990).

59. Han, J. S., Kang, S., Kim, S. H., Ko, M. J. & Hwang, D. S. Binding of SeqA Protein to Hemi-methylated GATC Sequences Enhances Their Interaction and Aggregation Properties. *J. Biol. Chem.* **279**, 30236–30243 (2004).

60. Waldminghaus, T. & Skarstad, K. The Escherichia coli SeqA protein. *Plasmid* **61**, 141–150 (2009).

61. Kurokawa, K. *et al.* Rapid Exchange of Bound ADP on the *Staphylococcus aureus* Replication Initiation Protein DnaA. *J. Biol. Chem.* **284**, 34201–34210 (2009).

62. Su'etsugu, M., Nakamura, K., Keyamura, K., Kudo, Y. & Katayama, T. Hda Monomerization by ADP Binding Promotes Replicase Clamp-mediated DnaA-ATP Hydrolysis. *J. Biol. Chem.* **283**, 36118–36131 (2008).

63. Su'etsugu, M. *et al.* The DnaA N-terminal domain interacts with Hda to facilitate replicase clamp-mediated inactivation of DnaA: DnaA domain I binding to Hda. *Environ. Microbiol.* **15**, 3183–3195 (2013).

64. Nozaki, S., Yamada, Y. & Ogawa, T. Initiator titration complex formed at *datA* with the aid of IHF regulates replication timing in *Escherichia coli*. *Genes Cells* **14**, 329–341 (2009).

65. Kasho, K. & Katayama, T. DnaA binding locus datA promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci.* **110**, 936–941 (2013).

66. Fujimitsu, K., Senriuchi, T. & Katayama, T. Specific genomic sequences of E. coli promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* **23**, 1221–1233 (2009).

67. Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T. & Katayama, T. Timely binding of IHF and

Fis to DARS2 regulates ATP–DnaA production and replication initiation. *Nucleic Acids Res.* **42**, 13134–13149 (2014).

68. Zhang, Q. *et al.* Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in Escherichia coli. *Sci. Rep.* **6**, (2016).
69. Wang, J. D. & Levin, P. A. Metabolism, cell growth and the bacterial cell cycle. *Nat. Rev. Microbiol.* **7**, 822–827 (2009).
70. Cooper, S. & Helmstetter, C. Chromosome Replication and the Division Cycle of Escherichia coli B/r. *J Mol Biol* 619–644 (1967).
71. Michelsen, O. Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. *Microbiology* **149**, 1001–1010 (2003).
72. Yoshikawa, H., O'Sullivan, A. & Sueoka, N. SEQUENTIAL REPLICATION OF THE BACILLUS SUBTILIS CHROMOSOME, III. REGULATION OF INITIATION. *Genetics* 973–980 (1964).
73. Donachie, W. D. Relationship between Cell size and Time of Initiation of DNA replication. *Nature* **219**, 1077–1079 (1968).
74. Schaechter, M., MaalOe, O. & Kjeldgaard, N. O. Dependency on Medium and Temperature of Cel Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *J.gen.Microbiol.* **19**, 592–606 (1958).
75. Campos, M. *et al.* A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell* **159**, 1433–1446 (2014).
76. Taheri-Araghi, S. *et al.* Cell-Size Control and Homeostasis in Bacteria. *Curr. Biol.* **25**, 385–391 (2015).
77. Wold, S., Skarstad, K., Steen, H. B., Stokke, T. & Boye, E. The initiation mass for DNA replication in Escherichia coli K-12 is dependent on growth rate. *EMBO J.* **13**, 2097–2102 (1994).
78. Boye, E., Stokke, T., Kleckner, N. & Skarstad, K. Coordinating DNA replication initiation

with cell growth: differential roles for DnaA and SeqA proteins. *Proc. Natl. Acad. Sci.* **93**, 12206–12211 (1996).

79. Bates, D. & Kleckner, N. Chromosome and Replisome Dynamics in E. coli: Loss of Sister Cohesion Triggers Global Chromosome Movement and Mediates Chromosome Segregation. *Cell* **121**, 899–911 (2005).
80. Chien, A.-C., Hill, N. S. & Levin, P. A. Cell Size Control in Bacteria. *Curr. Biol.* **22**, R340–R349 (2012).
81. Hill, N. S., Kadoya, R., Chattoraj, D. K. & Levin, P. A. Cell Size and the Initiation of DNA Replication in Bacteria. *PLoS Genet.* **8**, e1002549 (2012).
82. Weart, R. B. *et al.* A Metabolic Sensor Governing Cell Size in Bacteria. *Cell* **130**, 335–347 (2007).
83. Bipatnath, M., Dennis, P. P. & Bremer, H. Initiation and Velocity of Chromosome Replication in Escherichia coli B/r and K-12. *J BACTERIOL* **180**, 9 (1998).
84. Churchward, G. Growth Rate-Dependent Control of Chromosome Replication Initiation in Escherichia coli. *Journal of Bacteriology.* **145,** 1232–1238 (1981).
85. Nouri, H. *et al.* Multiple links connect central carbon metabolism to DNA replication initiation and elongation in *Bacillus subtilis*. *DNA Res.* **25**, 641–653 (2018).
86. Barańska, S. *et al.* Replicating DNA by cell factories: roles of central carbon metabolism and transcription in the control of DNA replication in microbes, and implications for understanding this process in human cells. *Microb. Cell Factories* **12**, 55 (2013).
87. Zyskind, J. W. & Smith, D. W. DNA replication, the bacterial cell cycle, and cell growth. *Cell* **69**, 5–8 (1992).
88. Chance, B., Estabrook, R. & Ghosh, A. DAMPED SINUSOIDAL OSCILLATIONS OF CYTOPLASMIC REDUCED PYRIDINE NUCLEOTIDE IN YEAST CELLS. *Proc Natl Acad Sci U A* **51**, 1244–1251 (1964).
89. Klevecz, R. R., Bolen, J., Forrest, G. & Murray, D. B. A genomewide oscillation in

transcription gates DNA replication and cell cycle. *Proc. Natl. Acad. Sci.* **101**, 1200–1205 (2004).

90. Tu, B. P. Logic of the Yeast Metabolic Cycle: Temporal Compartmentalization of Cellular Processes. *Science* **310**, 1152–1158 (2005).
91. Burnetti, A. J., Aydin, M. & Buchler, N. E. Cell cycle Start is coupled to entry into the yeast metabolic cycle across diverse strains and growth rates. *Mol. Biol. Cell* **27**, 64–74 (2016).
92. Chiaramello, A. E. & Zyskind, J. W. Expression of Escherichia coli dnaA and mioC genes as a function of growth rate. *J. Bacteriol.* **171**, 4272–4280 (1989).
93. Hanawalt, P. C., Maaløe, O., Cummings, D. J. & Schaechter, M. The normal DNA replication cycle. II. *J. Mol. Biol.* **3**, 156–165 (1961).
94. Ogura, Y., Imai, Y., Ogasawara, N. & Moriya, S. Autoregulation of the dnaA-dnaN Operon and Effects of DnaA Protein Levels on Replication Initiation in Bacillus subtilis. *J. Bacteriol.* **183**, 3833–3841 (2001).
95. Skarstad, K., Lobner-Olesen, A., Atlung, T., von Meyenburg, K. & Boye, E. Initiation of DNA replication in Escherichia coil after overproduction of the DnaA protein. *Mol Gen Genet* 50–56 (1989).
96. Xu, J. C. & Bremer, H. Chromosome replication in Escherichia coli induced by oversupply of DnaA. *Mol Gen Genet* 138–142 (1988).
97. Schaus, N., O'Day, K., Peters, W. & Wright, A. Isolation and Characterization of Amber Mutations in Gene dnaA of Escherichia coli K-12. *J BACTERIOL* **145**, 10 (1981).
98. Barker, M. M., Gaal, T., Josaitis, C. A. & Gourse, R. L. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* **305**, 673–688 (2001).
99. Cashel, M., Gentry, D. R., Hernandez, V. H. & Vinella, D. The Stringent Response. In: Escherichia coli and Salmonella. in *Cellular and Molecular Biology* (ASM, 1996).
100. Chiaramello, A. E. & Zyskind, J. W. Coupling of DNA replication to growth rate in

Escherichia coli: a possible role for guanosine tetraphosphate. *J. Bacteriol.* **172**, 2013–2019 (1990).

101. Samadpour, A. N. & Merrikh, H. DNA gyrase activity regulates DnaA-dependent replication initiation in *Bacillus subtilis*: DNA gyrase regulates DnaA-dependent initiation. *Mol. Microbiol.* **108**, 115–127 (2018).
102. Reece, R. J. & Maxwell, A. DNA Gyrase: Structure and Function. *Crit. Rev. Biochem. Mol. Biol.* **26**, 335–375 (1991).
103. Churchward, G. & Bremer, H. Determination of Deoxyribonucleic Acid Replication Time in Exponentially Growing Escherichia coli B/r. **130**, 8 (1977).
104. Herrick, J. & Sclavi, B. Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol. Microbiol.* **63**, 22–34 (2007).
105. Torrents, E. *et al.* NrdR Controls Differential Expression of the Escherichia coli Ribonucleotide Reductase Genes. *J. Bacteriol.* **189**, 5012–5021 (2007).
106. McKethan, B. L. & Spiro, S. Cooperative and allosterically controlled nucleotide binding regulates the DNA binding activity of NrdR: Characterization of NrdR in *Escherichia coli*. *Mol. Microbiol.* n/a-n/a (2013) doi:10.1111/mmi.12364.
107. DeNapoli, J., Tehranchi, A. K. & Wang, J. D. Dose-dependent reduction of replication elongation rate by (p)ppGpp in *Escherichia coli* and *Bacillus subtilis*: (p)ppGpp modulates replication elongation rates. *Mol. Microbiol.* **88**, 93–104 (2013).
108. Wang, J. D., Sanders, G. M. & Grossman, A. D. Nutritional Control of Elongation of DNA Replication by (p)ppGpp. *Cell* **128**, 865–875 (2007).
109. Murray, H. & Koh, A. Multiple Regulatory Systems Coordinate DNA Replication with Cell Growth in Bacillus subtilis. *PLoS Genet.* **10**, e1004731 (2014).
110. Flåtten, I., Fossum-Raunehaug, S., Taipale, R., Martinsen, S. & Skarstad, K. The DnaA Protein Is Not the Limiting Factor for Initiation of Replication in Escherichia coli. *PLOS Genet.* **11**, e1005276 (2015).

111. Hernandez, J. & Bremer, H. Characterization of RNA and DNA Synthesis in Escherichia coli Strains Devoid of ppGpp. *The Journal of Biological Chemistry.* **268**, 10851–10862 (1993).

112. Saxena, R., Fingland, N., Patil, D., Sharma, A. & Crooke, E. Crosstalk between DnaA Protein, the Initiator of Escherichia coli Chromosomal Replication, and Acidic Phospholipids Present in Bacterial Membranes. *Int. J. Mol. Sci.* **14**, 8517–8537 (2013).

113. SekimizuS, K. & Kornberg, A. Cardiolipin Activation of dnaA Protein, the Initiation Protein of Replication in Escherichia coZi. *J.Biol.Chem* 7131–7135 (1988).

114. Yung, B. Y. & Kornberg, A. Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in Escherichia coli. *Proc. Natl. Acad. Sci.* **85**, 7202–7205 (1988).

115. Castumas, C. E., Crooke, E. & Kornberg, A. Fluid Membranes with Acidic Domains Activate DnaA, the Initiator Protein of Replication in Escherichia coli (. *J. Biol. Chem.* **268**, 24665–24668 (1993).

116. Makise, M., Mima, S., Katsu, T., Tsuchiya, T. & Mizushima, T. Acidic phospholipids inhibit the DNA-binding activity of DnaA protein, the initiator of chromosomal DNA replication in Escherichia coli: Inhibition of DnaA binding to oriC by phospholipids. *Mol. Microbiol.* **46**, 245–256 (2002).

117. Heacock, P. N. & Dowhan, W. Alteration of the Phospholipid Composition of Escherichia coli through Genetic Manipulation. *J. Biol. Chem.* **264**, 14972–14977 (1989).

118. Heacock, P. N. & Dowhan, W. Construction of a Lethal Mutation in the Synthesis of the Major Acidic Phospholipids of Escherichia coli*. *J. Biol. Chem.* **262**, 13044–13049 (1987).

119. Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes. *Mol. Cell* **42**, 426–437 (2011).

120. Huang, T.-S. & Nagy, P. D. Direct Inhibition of Tombusvirus Plus-Strand RNA Synthesis by a Dominant Negative Mutant of a Host Metabolic Enzyme, Glyceraldehyde-3-Phosphate Dehydrogenase, in Yeast and Plants. *J. Virol.* **85**, 9090–9102 (2011).

121. Prasanth, K. R. *et al.* Glyceraldehyde 3-Phosphate Dehydrogenase Negatively Regulates the Replication of Bamboo Mosaic Virus and Its Associated Satellite RNA. *J. Virol.* **85**, 8829–8840 (2011).

122. Kim, B. H. & Gadd, G. M. *Bacterial Physiology and Metabolism*. (Cambridge University Press., 2008).

123. Lunt, S. Y. & Vander Heiden, M. G. Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. *Annu. Rev. Cell Dev. Biol.* **27**, 441–464 (2011).

124. Brasen, C., Esser, D., Rauch, B. & Siebers, B. Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation. *Microbiol. Mol. Biol. Rev.* **78**, 89–175 (2014).

125. Eymann, C. *et al.* A comprehensive proteome map of growingBacillus subtilis cells. *PROTEOMICS* **4**, 2849–2876 (2004).

126. Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci.* **110**, 10039–10044 (2013).

127. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**, 613–624 (2008).

128. Deutscher, J. *et al.* The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiol. Mol. Biol. Rev.* **78**, 231–256 (2014).

129. Alpert, C. A., Frank, R., Stueber, K., Deutscher, J. & Hengstenberg, W. Phosphoenolpyruvate-dependent protein kinase enzyme I of Streptococcus faecalis: purification and properties of the enzyme and characterization of its active center. *Biochemistry* **24**, 959–964 (1985).

130. Gassner, M. *et al.* The Phosphoenolpyruvate-Dependent Phosphotransferase System of Staphylococcus aureus. 2. 1H and 31P Nuclear-Magnetic-Resonance Studies on the Phosphocarrier Protein HPr, Phosphohistidines and Phosphorylated HPr. *Eur. J. Biochem.* **75**,

287–296 (1977).

131. Dorschug, M., Frank, R., Kalbitzer, H. R., Hengstenberg, W. & Deutscher, J. Phosphoenolpyruvate-dependent phosphorylation site in enzyme IIIglc of the Escherichia coli phosphotransferase system. *Eur. J. Biochem.* **144**, 113–119 (1984).
132. Deutscher, J., Beyreuther, K., Sobek, H. M. & Stuber, K. Phosphoenolpyruvate-Dependent Phosphotransferase System of Staphylococcus aureus: Factor IIIlaC, a Trimeric Phospho-Carrier Protein That Also Acts as a Phase Transfer Catalystt. *Biochemistry* **21**, 4867–4873 (1982).
133. Pas, H. H. & Robillard, G. T. S-Phosphocysteine and Phosphohistidine Are Intermediates in the Phosphoenolpyruvate-Dependent Mannitol Transport Catalyzed by Escherichia coli HIMt1. *Biochemistry* **27**, 5835–5839 (1988).
134. Cao, Y. *et al.* Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature* **473**, 50–54 (2011).
135. Bettenbrock, K. *et al.* A Quantitative Approach to Catabolite Repression in *Escherichia coli*. *J. Biol. Chem.* **281**, 2578–2584 (2006).
136. Hogema, B. M. *et al.* Inducer exclusion in Escherichia coli by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Mol. Microbiol.* **30**, 487–498 (1998).
137. Vadeboncoeur, C., Brochu, D. & Reizer, J. Quantitative Determination of the Intracellular Concentration of the Various Forms of HPr, a Phosphocarrier Protein of the Phosphoenolpyruvate: Sugar Phosphotransferase System in Growing Cells of Oral Streptococci. *Anal. Biochem.* **196**, 24–30 (1991).
138. Warner, J. B. & Lolkema, J. S. CcpA-Dependent Carbon Catabolite Repression in Bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 475–490 (2003).
139. Nicholson, W. L. *et al.* Catabolite repression-resistant mutations of the Bacillus subtilis alpha-amylase promoter affect transcription levels and are in an operator-like sequence. *J. Mol. Biol.*

198, 609–618 (1987).

140. Nessler, S. *et al.* HPr Kinase/Phosphorylase, the Sensor Enzyme of Catabolite Repression in Gram-Positive Bacteria: Structural Aspects of the Enzyme and the Complex with Its Protein Substrate. *J. Bacteriol.* **185**, 4003–4010 (2003).
141. Galinier, A. *et al.* New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proc. Natl. Acad. Sci.* **95**, 1823–1828 (1998).
142. Jault, J.-M. *et al.* The HPr Kinase from *Bacillus subtilis* Is a Homo-oligomeric Enzyme Which Exhibits Strong Positive Cooperativity for Nucleotide and Fructose 1,6-Bisphosphate Binding. *J. Biol. Chem.* **275**, 1773–1780 (2000).
143. Reizer, J. *et al.* A novel protein kinase that controls carbon catabolite repression in bacteria. *Mol. Microbiol.* **27**, 1157–1169 (1998).
144. Mijakovic, I. *et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc. Natl. Acad. Sci.* **99**, 13442–13447 (2002).
145. Deutscher, J., Küster, E., Bergstedt, U., Charrier, V. & Hillen, W. Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in Gram-positive bacteria. *Mol. Microbiol.* **15**, 1049–1053 (1995).
146. Jones, B. E. *et al.* Binding of the Catabolite Repressor Protein CcpA to Its DNA Target Is Regulated by Phosphorylation of its Corepressor HPr. *J. Biol. Chem.* **272**, 26530–26535 (1997).
147. Seidel, G., Diel, M., Fuchsbauer, N. & Hillen, W. Quantitative interdependence of coeffectors, CcpA and cre in carbon catabolite regulation of Bacillus subtilis: Regulatory differences of HPrSerP and CrhP. *FEBS J.* **272**, 2566–2577 (2005).
148. Schumacher, M. A., Seidel, G., Hillen, W. & Brennan, R. G. Structural Mechanism for the Fine-tuning of CcpA Function by The Small Molecule Effectors Glucose 6-Phosphate and Fructose 1,6-Bisphosphate. *J. Mol. Biol.* **368**, 1042–1050 (2007).
149. Djordjevic, G. M., Tchieu, J. H. & Saier, M. H. Genes Involved in Control of Galactose

Uptake in Lactobacillus brevis and Reconstitution of the Regulatory System in Bacillus subtilis. *J. Bacteriol.* **183**, 3224–3236 (2001).

150. Poolman, B., Knol, J., Mollet, B., Nieuwenhuis, B. & Sulter, G. Regulation of bacterial sugar-H+ symport by phosphoenolpyruvate-dependent enzyme I/HPr-mediated phosphorylation. *Proc. Natl. Acad. Sci.* **92**, 778–782 (1995).
151. Gunnewijk, M. G. W. & Poolman, B. Phosphorylation State of HPr Determines the Level of Expression and the Extent of Phosphorylation of the Lactose Transport Protein of *Streptococcus thermophilus*. *J. Biol. Chem.* **275**, 34073–34079 (2000).
152. Stulke, J., Arnaud, M., Rapoport, G. & Martin-Verstraete, I. PRD, a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol. Microbiol.* **28**, 865–874 (1998).
153. Martin-Verstraete, I. *et al.* Antagonistic effects of dual PTS-catalysed phosphorylation on the Bacillus subtilis transcriptional activator LevR. *Mol. Microbiol.* **28**, 293–303 (1998).
154. Tortosa, P. *et al.* Sites of positive and negative regulation in the Bacillus subtilis antiterminators LicT and SacY. *Mol. Microbiol.* **41**, 1381–1393 (2001).
155. Krüger, S., Gertz, S. & Hecker, M. Transcriptional analysis of bglPH expression in Bacillus subtilis: evidence for two distinct pathways mediating carbon catabolite repression. *J. Bacteriol.* **178**, 2637–2644 (1996).
156. Lindner, C., Galinier, A., Hecker, M. & Deutscher, J. Regulation of the activity of the Bacillus subtilis antiterminator LicT by multiple PEP-dependent, enzyme I- and HPr-catalysed phosphorylation. *Mol. Microbiol.* **31**, 995–1006 (1999).
157. Lindner, C., Hecker, M., Le Coq, D. & Deutscher, J. Bacillus subtilis Mutant LicT Antiterminators Exhibiting Enzyme I- and HPr-Independent Antitermination Affect Catabolite Repression of the bglPH Operon. *J. Bacteriol.* **184**, 4819–4828 (2002).
158. Bouraoui, H., Ventroux, M., Noirot-Gros, M.-F., Deutscher, J. & Joyet, P. Membrane sequestration by the EIIB domain of the mannitol permease MtlA activates the *Bacillus*

subtilis mtl operon regulator MtlR: *B. subtilis* MtlR sequestration by the EIIB ^{Mtl} domain. *Mol. Microbiol.* **87**, 789–801 (2013).

159. Krin, E., Sismeiro, O., Danchin, A. & Bertin, P. N. The regulation of Enzyme IIAGlc expression controls adenylate cyclase activity in Escherichia coli. *Microbiology* **148**, 1553–1559 (2002).
160. Rothe, F. M., Wrede, C., Lehnik-Habrink, M., Gorke, B. & Stulke, J. Dynamic Localization of a Transcription Factor in Bacillus subtilis: the LicT Antiterminator Relocalizes in Response to Inducer Availability. *J. Bacteriol.* **195**, 2146–2154 (2013).
161. Jannière, L. *et al.* Genetic Evidence for a Link Between Glycolysis and DNA Replication. *PLoS ONE* **2**, e447 (2007).
162. Maciąg, M., Nowicki, D., Janniere, L., Szalewska-Pałasz, A. & Węgrzyn, G. Genetic response to metabolic fluctuations: correlation between central carbon metabolism and DNA replication in Escherichia coli. *Microb. Cell Factories* **10**, 19 (2011).
163. Maciąg-Dorszyńska, M., Ignatowska, M., Jannière, L., Węgrzyn, G. & Szalewska-Pałasz, A. Mutations in central carbon metabolism genes suppress defects in nucleoid position and cell division of replication mutants in Escherichia coli. *Gene* **503**, 31–35 (2012).
164. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci.* **96**, 2896–2901 (1999).
165. Chen, Y. & Tye, B. K. The yeast Mcm1 protein is regulated posttranscriptionally by the flux of glycolysis. *Mol. Cell. Biol.* **15**, 4631–4639 (1995).
166. Chang, V. K. *et al.* Mcm1 Binds Replication Origins. *J. Biol. Chem.* **278**, 6093–6100 (2003).
167. Chang, V. K., Donato, J. J., Chan, C. S. & Tye, B. K. Mcm1 Promotes Replication Initiation by Binding Specific Elements at Replication Origins. *Mol. Cell. Biol.* **24**, 6514–6524 (2004).
168. Sprague, G. F. Isolation and Characterization of a Saccharomyces cerevisiae Mutant Deficient in Pyruvate Kinase Activity. *J BACTERIOL* **130**, 10 (1977).
169. Dickinson, J. R. & Williams, A. S. The cdc30 Mutation in Saccharomyces cerevisiae Results

in a Temperature-sensitive Isoenzyme of Phosphoglucose Isomerase. *Microbiology* **133**, 135–140 (1987).

170. Hartwell, L. H. GENETIC CONTROL OF THE CELL DIVISION CYCLE I N YEAST: *Genetics* **74**, 267–286 (1973).
171. Kaplan, Y. & Kupiec, M. A role for the yeast cell cycle/splicing factor Cdc40 in the G1/S transition. *Curr. Genet.* **51**, 123–140 (2007).
172. Shor, E. *et al.* The Origin Recognition Complex Interacts with a Subset of Metabolic Genes Tightly Linked to Origins of Replication. *PLoS Genet.* **5**, e1000755 (2009).
173. Konieczna, A., Szczepańska, A., Sawiuk, K., Węgrzyn, G. & Łyżeń, R. Effects of partial silencing of genes coding for enzymes involved in glycolysis and tricarboxylic acid cycle on the enterance of human fibroblasts to the S phase. *BMC Cell Biol.* **16**, (2015).
174. Fornalewicz, K., Wieczorek, A., Węgrzyn, G. & Łyżeń, R. Silencing of the pentose phosphate pathway genes influences DNA replication in human fibroblasts. *Gene* **635**, 33–38 (2017).
175. Buckland, R. J. *et al.* Increased and Imbalanced dNTP Pools Symmetrically Promote Both Leading and Lagging Strand Replication Infidelity. *PLoS Genet.* **10**, e1004846 (2014).
176. Tymecka-Mulik, J. *et al.* Suppression of the Escherichia coli dnaA46 mutation by changes in the activities of the pyruvate-acetate node links DNA replication regulation to central carbon metabolism. *PLOS ONE* **12**, e0176050 (2017).
177. Rannou, O. *et al.* Functional interplay of DnaE polymerase, DnaG primase and DnaC helicase within a ternary complex, and primase to polymerase hand-off during lagging strand DNA replication in Bacillus subtilis. *Nucleic Acids Res.* **41**, 5303–5320 (2013).
178. Paschalis, V. *et al.* Interactions of the *Bacillus subtilis* DnaE polymerase with replisomal proteins modulate its activity and fidelity. *Open Biol.* **7**, 170146 (2017).
179. Kilkenny, M. L. *et al.* The human CTF4-orthologue AND-1 interacts with DNA polymerase α/primase via its unique C-terminal HMG box. *Open Biol.* **7**, 170217 (2017).
180. Noirot-Gros, M.-F. *et al.* An expanded view of bacterial DNA replication. *Proc. Natl. Acad.*

Sci. **99**, 8342–8347 (2002).

181. Stein, A. & Firshein, W. Probable Identification of a Membrane-Associated Repressor of Bacillus subtilis DNA Replication as the E2 Subunit of the Pyruvate Dehydrogenase Complex. *J. Bacteriol.* **182**, 2119–2124 (2000).
182. Ronai, Z. Glycolytic enzymes as DNA binding proteins. *Int. J. Biochem.* **25**, 1073–1076 (1993).
183. Boukouris, A. E., Zervopoulos, S. D. & Michelakis, E. D. Metabolic Enzymes Moonlighting in the Nucleus: Metabolic Regulation of Gene Transcription. *Trends Biochem. Sci.* **41**, 712–730 (2016).
184. Baxi, M. D. & Vishwanatha, J. K. Uracil DNA-glycosylase/glyceraldehyde-3-phosphate dehydrogenase is an Ap4A binding protein. *Biochemistry* **34**, 9700–9707 (1995).
185. Grosse, F., Nasheuer, H.-P., Scholtissek, S. & Schomburg, U. Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-alpha-primase complex. *Eur. J. Biochem.* **160**, 459–467 (1986).
186. Popanda, O., Fox, G. & Thielmann, H. W. Modulation of DNA polymerases a , d and ´ by lactate dehydrogenase and 3-phosphoglycerate kinase. *Biochim. Biophys. Acta* **1397**, 102–117 (1998).
187. Jindal, H. K. & Vishwanatha, K. Functional Identity of a Primer Recognition Protein as Phosphoglycerate Kinase. *J. Biol. Chem.* **265**, 6540–6543 (1990).
188. Kumble, K. D., Iversen, P. L. & Vishwanatha, J. K. The role of primer recognition proteins in DNA replication: inhibition of cellular proliferation by antisense oligodeoxyribonucleotides. *J. Cell Sci.* **101**, 35–41 (1992).
189. Li, X. *et al.* Nuclear PGK1 Alleviates ADP-Dependent Inhibition of CDC7 to Promote DNA Replication. *Mol. Cell* **72**, 650-660.e8 (2018).
190. Tian, M. *et al.* Salicylic Acid Inhibits the Replication of *Tomato bushy stunt virus* by Directly Targeting a Host Component in the Replication Complex. *Mol. Plant. Microbe Interact.* **28**,

379–386 (2015).

191. Prasanth, K. R., Chuang, C. & Nagy, P. D. Co-opting ATP-generating glycolytic enzyme PGK1 phosphoglycerate kinase facilitates the assembly of viral replicase complexes. *PLOS Pathog.* **13**, e1006689 (2017).
192. van Noort, V. *et al.* Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Mol. Syst. Biol.* **8**, (2012).
193. Hentchel, K. L. & Escalante-Semerena, J. C. Acylation of Biomolecules in Prokaryotes: a Widespread Strategy for the Control of Biological Function and Metabolic Stress. *Microbiol. Mol. Biol. Rev.* **79**, 321–346 (2015).
194. Garcia-Garcia, T. *et al.* Role of Protein Phosphorylation in the Regulation of Cell Cycle and DNA-Related Processes in Bacteria. *Front. Microbiol.* **7**, (2016).
195. Shi, L. *et al.* Protein-tyrosine phosphorylation interaction network in Bacillus subtilis reveals new substrates, kinase activators and kinase cross-talk. *Front. Microbiol.* **5**, (2014).
196. Petranovic, D. *et al.* Bacillus subtilis strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication: Bacillus subtilis ΔptkA deficient in DNA replication. *Mol. Microbiol.* **63**, 1797–1805 (2007).
197. Mijakovic, I. Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* **34**, 1588–1596 (2006).
198. Ma, Y., Kanakousaki, K. & Buttitta, L. How the cell cycle impacts chromatin architecture and influences cell fate. *Front. Genet.* **6**, (2015).
199. Shi, L. & Tu, B. P. Acetyl-CoA induces transcription of the key G1 cyclin CLN3 to promote entry into the cell division cycle in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **110**, 7318–7323 (2013).
200. Zheng, L., Roeder, R. G. & Luo, Y. S Phase Activation of the Histone H2B Promoter by OCA-S, a Coactivator Complex that Contains GAPDH as a Key Component. *Cell* **114**, 255–266 (2003).

201. Sutendra, G. *et al.* A Nuclear Pyruvate Dehydrogenase Complex Is Important for the Generation of Acetyl-CoA and Histone Acetylation. *Cell* **158**, 84–97 (2014).

202. Yang, W. *et al.* PKM2 Phosphorylates Histone H3 and Promotes Gene Transcription and Tumorigenesis. *Cell* **150**, 685–696 (2012).

203. Sauer, U. & Eikmanns, B. J. The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria, *FEMS Microbiol. Rev.* **29**, 765–794 (2005).

204. Servant, P., Le Coq, D. & Aymerich, S. CcpN (YqzB), a novel regulator for CcpA-independent catabolite repression of Bacillus subtilis gluconeogenic genes: Catabolite repression of gluconeogenesis genes by CcpN. *Mol. Microbiol.* **55**, 1435–1451 (2005).

205. Yoshida, K. -i. Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in Bacillus subtilis. *Nucleic Acids Res.* **29**, 683–692 (2001).

206. Blencke, H.-M. *et al.* Transcriptional profiling of gene expression in response to glucose in Bacillus subtilis: regulation of the central metabolic pathways. *Metab. Eng.* **5**, 133–149 (2003).

207. Diesterhaft, M. & Freese, E. Pyruvate kinase of Bacillus subtilis. *Biochim. Biophys. Acta BBA - Enzymol.* **268**, 373–380 (1972).

208. Diesterhaft, M. & Freese, E. Role of Pyruvate Carboxylase, Phosphoenolpyruvate Carboxykinase, and Malic Enzyme during Growth and Sporulation of Bacillus subtilis. *J. Biol. Chem.* **248**, 6062–6070 (1973).

209. Monahan, L. G., Hajduk, I. V., Blaber, S. P., Charles, I. G. & Harry, E. J. Coordinating Bacterial Cell Division with Nutrient Availability: a Role for Glycolysis. *mBio* **5**, (2014).

210. Dombrauckas, J. D., Santarsiero, B. D. & Mesecar, A. D. Structural Basis for Tumor Pyruvate Kinase M2 Allosteric Regulation and Catalysis †,‡. *Biochemistry* **44**, 9417–9429 (2005).

211. Mattevi, A. *et al.* Crystal structure of Escherichia coli pyruvate kinase type I: molecular basis of the allosteric transition. *Structure* **3**, 729–741 (1995).

212. Zhong, W. *et al.* Allosteric pyruvate kinase-based “logic gate” synergistically senses energy and sugar levels in Mycobacterium tuberculosis. *Nat. Commun.* **8**, (2017).

213. Suzuki, K., Ito, S., Shimizu-Ibuka, A. & Sakai, H. Crystal Structure of Pyruvate Kinase from Geobacillus stearothermophilus. *J. Biochem. (Tokyo)* **144**, 305–312 (2008).

214. Cheng, X., Friesen, R. H. E. & Lee, J. C. Effects of Conserved Residues on the Regulation of Rabbit Muscle Pyruvate Kinase. *J. Biol. Chem.* **271**, 6313–6321 (1996).

215. Muirhead, H. *et al.* The structure of cat muscle pyruvate kinase. *EMBO J.* **5**, 475–481 (1986).

216. Larsen, T. M., Benning, M. M., Rayment, I. & Reed, G. H. Structure of the Bis(Mg $^{2+}$)−ATP−Oxalate Complex of the Rabbit Muscle Pyruvate Kinase at 2.1 Å Resolution: ATP Binding over a Barrel †,‡. *Biochemistry* **37**, 6247–6255 (1998).

217. Rigden, D. J., Phillips, S. E. V., Michels, P. A. M. & Fothergill-Gilmore, L. A. The structure of pyruvate kinase from Leishmania mexicana reveals details of the allosteric transition and unusual effector specificity. *J. Mol. Biol.* **293**, 745–749 (1999).

218. Schormann, N., Hayden, K. L., Lee, P., Banerjee, S. & Chattopadhyay, D. An overview of structure, function, and regulation of pyruvate kinases. *Protein Sci.* **28**, 1771–1784 (2019).

219. Sakai, H. Possible Structure and Function of the Extra C-Terminal Sequence of Pyruvate Kinase from Bacillus stearothermophilus. *J. Biochem. (Tokyo)* **136**, 471–476 (2004).

220. Sakai, H. Mutagenesis of the Active Site Lysine 221 of the Pyruvate Kinase from Bacillus stearothermophilus. *J. Biochem. (Tokyo)* **137**, 141–145 (2005).

221. Bollenbach, T. J., Mesecar, A. D. & Nowak, T. Role of Lysine 240 in the Mechanism of Yeast Pyruvate Kinase Catalysis †. *Biochemistry* **38**, 9137–9145 (1999).

222. Teplyakov, A. *et al.* Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proc. Natl. Acad. Sci.* **103**, 16218–16223 (2006).

223. Herzberg, O. *et al.* Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci.* **93**, 2652–2657 (1996).

224. Burnell, J. N. & Hatch, M. D. Regulation of C4 photosynthesis: Identification of a catalytically important histidine residue and its role in the regulation of pyruvate,Pi dikinase. *Arch. Biochem. Biophys.* **231**, 175–182 (1984).

225. Tolentino, R., Chastain, C. & Burnell, J. Identification of the amino acid involved in the regulation of bacterial pyruvate, orthophosphate dikinase and phosphoenolpyruvate synthetase. *Adv. Biol. Chem.* **03**, 12–21 (2013).

226. Magill, N. G. & Setlow, P. Properties of purified sporlets produced by spoII mutants of Bacillus subtilis. *J. Bacteriol.* **174**, 8148–8151 (1992).

227. Zheng, H. *et al.* Interrogating the *Escherichia coli* cell cycle by cell dimension perturbations. *Proc. Natl. Acad. Sci.* **113**, 15000–15005 (2016).

228. Séror, S. J. *et al.* A mutant cysteinyl-tRNA synthetase affecting timing of chromosomal replication initiation in B. subtilis and conferring resistance to a protein kinase C inhibitor. *EMBO J.* **13**, 2472–2480 (1994).

229. Stokke, C., Flåtten, I. & Skarstad, K. An Easy-To-Use Simulation Program Demonstrates Variations in Bacterial Cell Cycle Parameters Depending on Medium and Temperature. *PLoS ONE* **7**, e30981 (2012).

230. Mazurek, S., Drexler, H. C. A., Troppmair, J., Eigenbrodt, E. & Rapp, U. R. Regulation of Pyruvate Kinase Type M2 by A-Raf: A Possible Glycolytic Stop or Go Mechanism. *ANTICANCER Res.* 9 (2007).

231. Hitosugi, T. *et al.* Tyrosine Phosphorylation Inhibits PKM2 to Promote the Warburg Effect and Tumor Growth. *Sci. Signal.* **2**, ra73–ra73 (2009).

232. Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M. & Cantley, L. C. Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature* **452**, 181–186 (2008).

233. Mukherjee, J. *et al.* PKM2 uses control of HuR localization to regulate p27 and cell cycle progression in human glioblastoma cells: PKM2 Regulates HuR-mediated p27 Expression. *Int. J. Cancer* **139**, 99–111 (2016).

234. Huang, L. *et al.* Interaction with Pyruvate Kinase M2 Destabilizes Tristetraprolin by Proteasome Degradation and Regulates Cell Proliferation in Breast Cancer. *Sci. Rep.* **6**, (2016).

235. Azoitei, N. *et al.* PKM2 promotes tumor angiogenesis by regulating HIF-1α through NF-κB activation. *Mol. Cancer* **15**, (2016).

236. Yang, W. Structural basis of PKM2 regulation. *Protein Cell* **6**, 238–240 (2015).

237. Yang, W. *et al.* Nuclear PKM2 regulates β-catenin transactivation upon EGFR activation. *Nature* **480**, 118–122 (2011).

238. Gao, X., Wang, H., Yang, J. J., Liu, X. & Liu, Z.-R. Pyruvate Kinase M2 Regulates Gene Transcription by Acting as a Protein Kinase. *Mol. Cell* **45**, 598–609 (2012).

239. Jiang, Y. *et al.* PKM2 phosphorylates MLC2 and regulates cytokinesis of tumour cells. *Nat. Commun.* **5**, (2014).

240. Jiang, Y. *et al.* PKM2 Regulates Chromosome Segregation and Mitosis Progression of Tumor Cells. *Mol. Cell* **53**, 75–87 (2014).

241. Keller, K. E., Doctor, Z. M., Dwyer, Z. W. & Lee, Y.-S. SAICAR Induces Protein Kinase Activity of PKM2 that Is Necessary for Sustained Proliferative Signaling of Cancer Cells. *Mol. Cell* **53**, 700–709 (2014).

242. Hosios, A. M., Fiske, B. P., Gui, D. Y. & Vander Heiden, M. G. Lack of Evidence for PKM2 Protein Kinase Activity. *Mol. Cell* **59**, 850–857 (2015).

243. Li, L., Zhang, Y., Qiao, J., Yang, J. J. & Liu, Z.-R. Pyruvate Kinase M2 in Blood Circulation Facilitates Tumor Growth by Promoting Angiogenesis. *J. Biol. Chem.* **289**, 25812–25821 (2014).

244. Yang, P. *et al.* Secreted pyruvate kinase M2 facilitates cell migration via PI3K/Akt and Wnt/β-catenin pathway in colon cancer cells. *Biochem. Biophys. Res. Commun.* **459**, 327–332 (2015).

245. Zhang, Y., Li, L., Liu, Y. & Liu, Z.-R. PKM2 released by neutrophils at wound site facilitates early wound healing by promoting angiogenesis: PKM2 facilitates wound healing. *Wound*

Repair Regen. **24**, 328–336 (2016).

246. Buschow, S. I. *et al.* MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis. *Immunol. Cell Biol.* **88**, 851–856 (2010).
247. Snaebjornsson, M. T. & Schulze, A. Non-canonical functions of enzymes facilitate cross-talk between cell metabolic and regulatory pathways. *Exp. Mol. Med.* **50**, 34 (2018).
248. Commichau, F. M. & Stülke, J. Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression: Trigger enzymes in transcription regulation. *Mol. Microbiol.* **67**, 692–702 (2007).
249. Warburg, O. On the Origin of Cancer Cells. *Science* **123**, 309–314 (1956).
250. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).
251. Xu, X. D. *et al.* Warburg Effect or Reverse Warburg Effect? A Review of Cancer Metabolism. *Oncol. Res. Treat.* **38**, 117–122 (2015).
252. Bertram, J. S. The molecular biology of cancer. *Mol. Aspects Med.* 57 (2001).
253. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
254. Loeb, L. A., Loeb, K. R. & Anderson, J. P. Multiple mutations and cancer. *Proc. Natl. Acad. Sci.* **100**, 776–781 (2003).
255. Macheret, M. & Halazonetis, T. D. DNA Replication Stress as a Hallmark of Cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 425–448 (2015).
256. Hsu, M.-C. & Hung, W.-C. Pyruvate kinase M2 fuels multiple aspects of cancer cells: from cellular metabolism, transcriptional regulation to extracellular signaling. *Mol. Cancer* **17**, (2018).
257. Sizemore, S. T. *et al.* Pyruvate kinase M2 regulates homologous recombination-mediated DNA double-strand break repair. *Cell Res.* **28**, 1090–1102 (2018).
258. Liang, J. *et al.* Mitochondrial PKM2 regulates oxidative stress-induced apoptosis by

stabilizing Bcl2. *Cell Res.* **27**, 329–351 (2017).

259. Anastasiou, D. *et al.* Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.* **8**, 839–847 (2012).
260. Buck, M. D., Sowell, R. T., Kaech, S. M. & Pearce, E. L. Metabolic Instruction of Immunity. *Cell* **169**, 570–586 (2017).
261. Vivarelli, S. *et al.* Gut Microbiota and Cancer: From Pathogenesis to Therapy. *Cancers* **11**, 38 (2019).

Annex I: Periodic organization of stress response regulons in *E. coli*

1. Introduction

1.1 The connection between chromosome structure, gene position and gene expression

Bacterial chromosomes have to be compacted 500-1000 times to fit inside the bacterial cell, whilst staying accessible to the action of DNA processing proteins such as DNA polymerases, RNA polymerases and other DNA binding proteins[270]. This is achieved by folding the chromosome into a remarkably intricate structure called the nucleoid of which the most detailed features are only now being rapidly discovered[270]. This intricate structure is summarized in figure 1. The lowest level of chromosome structure is defined by supercoiling or the twisting of the DNA helix[270]. DNA supercoiling is balanced by dedicated enzymes called topoisomerases and gyrases, that can relax and add DNA twists, respectively. DNA topology is continuously altered by the action of DNA processing enzymes like DNA and RNA polymerases, which overwind the DNA in front of them and underwind the DNA behind[271]. On a higher structural level, abundant nucleoid associated proteins (NAP's) organize the chromosome by locally constraining supercoiling in topologically isolated domains called microdomains. Experimental studies in *E. coli* have shown that its chromosome has 400 of these microdomains of average length between 8-12 kb[272]. Recently, new experimental studies have revealed that the *E. coli* chromosome is organized in domains of 100-200 kb long, in which DNA elements physically interact more frequently[273] These domains represent yet another level of chromosome structural organization. At the highest level of organization (~1 Mbp), the chromosome is divided in four structured macrodomains (Ori, Left, Right, Ter) and two non-structured ones (NS-Left, NS-Right)[274]. The Ter macrodomain is organized by the MatP protein[275] and the Ori one by the MaoP protein[276], but the other macrodomains seem to be defined by the replication process itself[277]. Figure 1 was assembled from figures in [270,278–280].

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

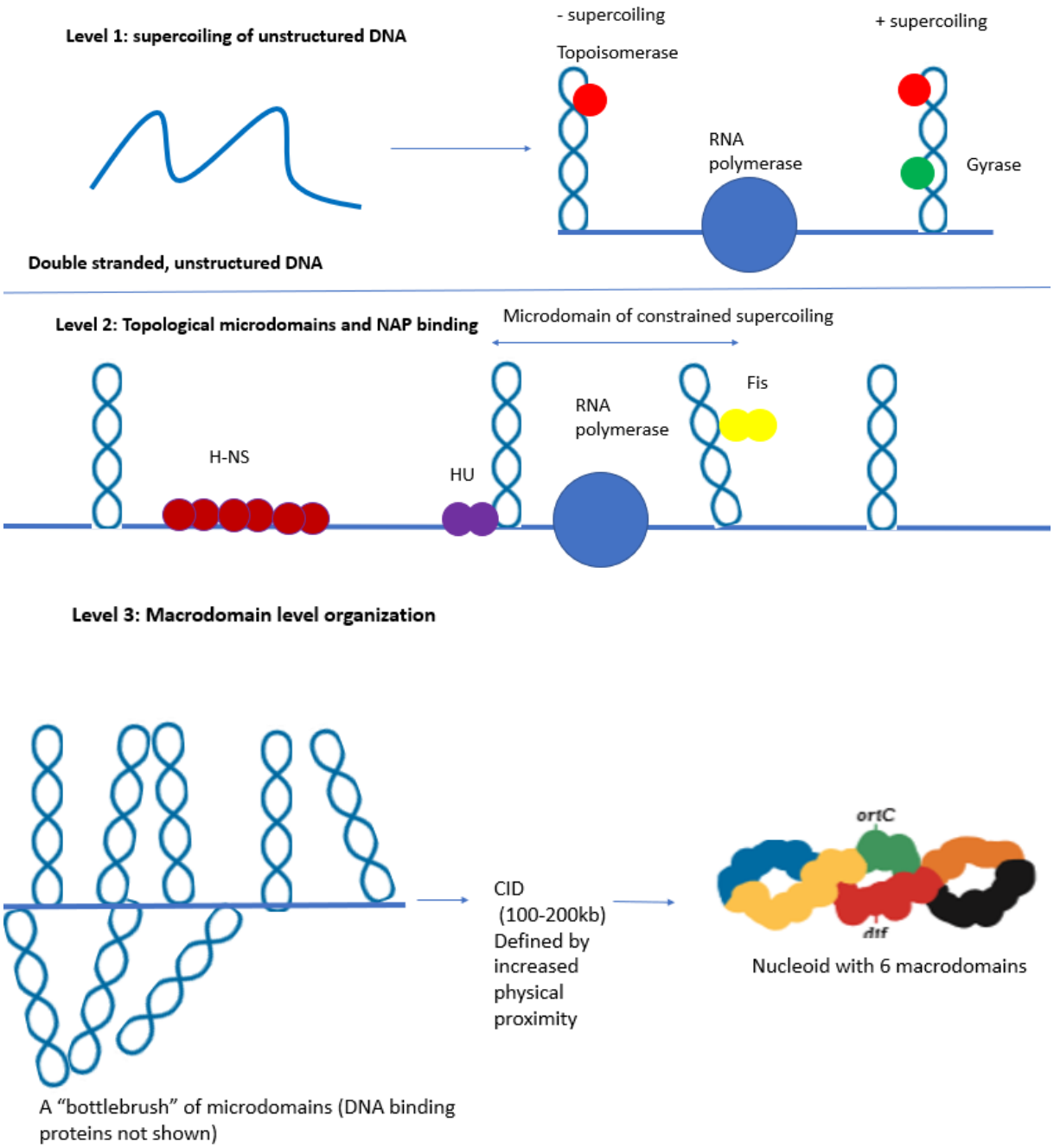


Figure 1 : Structural organization of the E.coli chromosome: (Level 1) The action of DNA processing enzymes such as RNA and DNA polymerases causes twisting of the DNA double helix. Their action causes overwinding in front (+ supercoiling) and underwinding behind (- supercoiling) These twisted structures are called plectonemes and can be relaxed using the dedicated enzymes topisomerases or introduced using gyrases. (Level 2) DNA is further compacted through the action of abundant DNA binding proteins called nucleoid associated proteins (NAP's). Their binding creates barriers for supercoiling diffusion and leads to the establishment of 8-12kb topologically insulated domains called microdomains. Additionally, their binding causes specific deformations depending on the NAP such as DNA bending, bridging, wrapping and coating. This leads to a 'bottlebrush' of microdomains. (Level 3) The microdomains organize into 100-200kb chromatin interaction domains (CID) in which physical interactions happen more frequently than outside. These CID's are then organized into Mb size macrodomains : (Green) Ori macrodomain (Black) Non-Structured Right macrodomain (Orange) Right macrodomain (Red) Ter macrodomain (Blue)Left macrodomain (Yellow) Non-structured left macrodomain. The figure was assembled from figures in [270,278–280]

Interestingly, several features of chromosomal structural organization have been shown to profoundly impact DNA transcription and replication. For example, supercoiling has been shown to directly affect the expression of many genes[281]. Moreover, constraining supercoiling to a localized domain on the chromosome is thought to cause genes within such microdomain to have correlated transcription and transcription bursting profiles, underlining the importance of these microdomains for gene expression control[282]. Macrodomains on the other hand seem to contribute to efficient replication [283] [284]and limit genome plasticity[285]. To date, there is no known evidence suggesting a direct link between macrodomain organization and gene expression.

The dependence of gene expression on its position in the chromosome sequence is a consequence of this link between chromosome structure and gene expression[286]. This implies that there is an evolutionary pressure to organize the positioning of genes in ways that optimize their expression.

Several *in silico* analyses have shown that most genes are not randomly scattered along the chromosome sequence, but follow distinct chromosome layouts that reflect structural features of the chromosome[287] [288] [289].

This leads to a three way relationship between gene position, gene expression and chromosome structure that evolves to optimize gene expression and function. This relationship is depicted in figure 2.

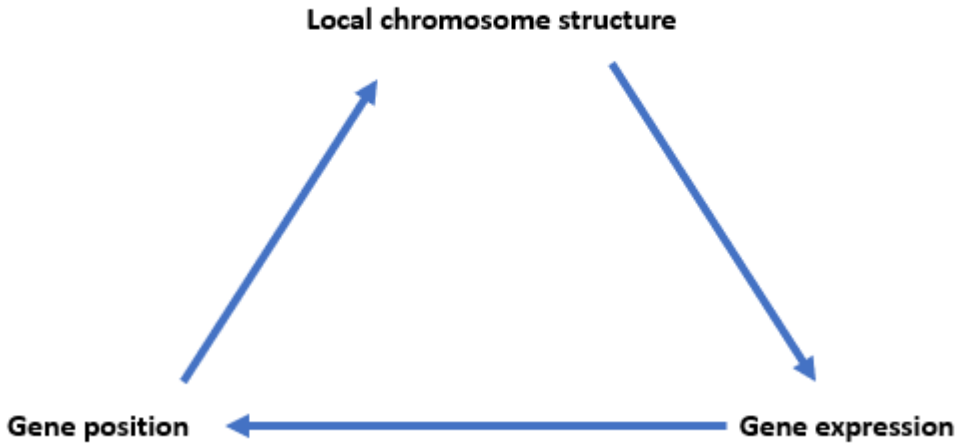


Figure 2 : A three way relation between chromosome structure, gene expression and gene position Gene position determines in which local chromosome structure is located. This in turn influences gene expression. Since gene expression is tightly related to gene function, the position of a gene is under evolutionary pressure to permit optimal expression.

In the next section, the transcription based solenoid model (TBS) is introduced. This model explains the observation of periodic arrangement of cofunctional genes on the *E.coli* chromosome using this relationship between gene position, local chromosome structure and gene expression and is the object of this study.

1.2 The Transcription Based Solenoidal model: Periodic organization of genes optimizes expression

Computational analysis of fully sequenced genomes showed that coregulated genes tended to be organized close together in regulons of 20 genes or less, whereas larger regulons tended to be organized periodically along the chromosome in many different organisms[287], most notably in *E. coli*[290] and *Saccharomyces cerevisiae*[291]. In order to explain this phenomenon, the Transcription based Solenoidal (TBS) framework of chromosome expression was elaborated by Kepes and Vaillant in 2003[292] and is depicted in figure 3. According to this model, gene expression regulation is optimized by confining the necessary transcription factors, RNA polymerases and coregulated genes into a small nuclear volume called a transcription focus. Like other chemical reactions, the confining of all necessary factors inside this small volume was thought to improve the speed and robustness of gene expression and regulation[292]. These transcription foci were thought to self-assemble through mediation of bivalent transcription factors, which would recruit their target genes and then even more transcription factors to the focus in a positive feedback loop. According to the model, the spatial clustering of coregulated genes is facilitated either by organizing them close together on the chromosome sequence (1D clustering) or by organizing them periodically along the sequence (interpreted as 3D clustering). In this scenario, the model states that the 3D folding of the chromosome will bring them together in space either by folding into solenoids (eukaryotes) or plectonemes (prokaryotes)[292].

When the TBS model was formulated, it explained a lot of observations at the time, particularly in yeast[292]. First, transcription factors are remarkably specific, despite the presence of a multitude of unused protein binding motifs[293–296]. The TBS explains this by making gene position a discriminator between used and unused sites[292]. Second, transcription factors find their targets incredibly fast[297,298] and the TBS framework explains this by suggesting that the distance between transcription factors and its targets is decreased under a periodic organization[292]. Finally, the TBS model suggests that by changing the chromosomal structure and bringing together different genes, organisms can rapidly alter their gene expression programs[299]. However, the TBS framework has never been directly tested.

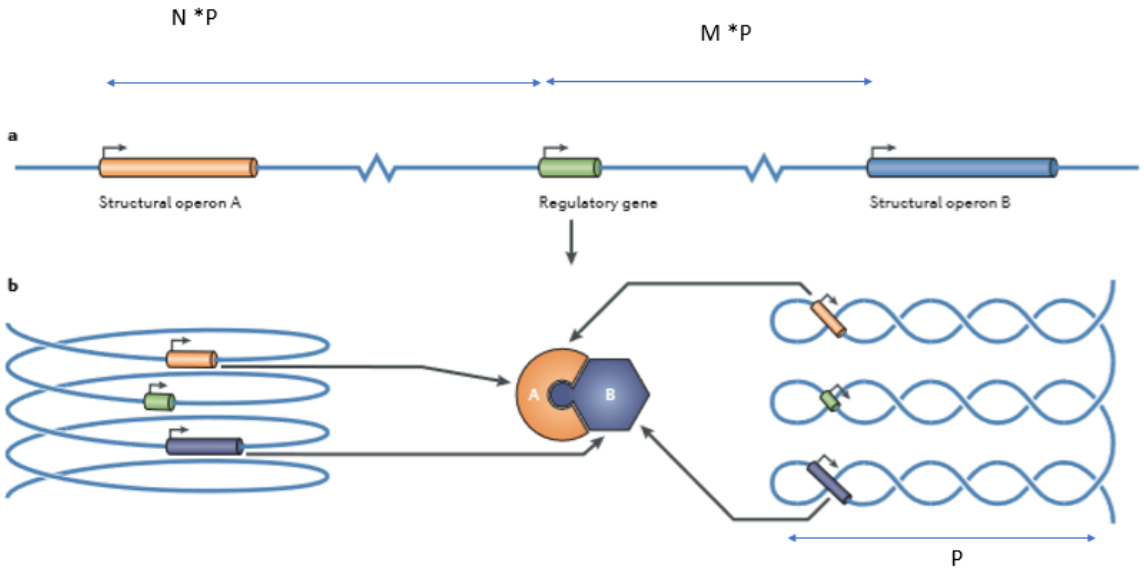


*Figure 3 : The Transcription Based Solenoid model : Imagine a regulon consisting of a regulatory gene (green), a structural operon A (orange) and a structural operon B (blue). Next, imagine that the length of a plectoneme is P and that the genes are separated by distances that are a multiple of P (N*P and M*P) a) When depicted in a 1D conformation, the genes of the regulon are separated by a long distance. b) When depicted in a 3D confirmation with plectonemes (bacteria) or solenoids (eukaryotes), the genes of the regulon are now in close proximity. The TBS model predicts that this confirmation facilitates the search of the regulatory gene for its targets and also the assembly of gene products A and B into a complex. Picture taken from* [1]

In this work, the TBS framework is tested by analyzing the periodic organization of stress response genes and by studying its implications for stress response regulation.

1.3 Stress response systems in *E.coli*

As most bacteria, *E. coli* has to cope with a rapidly changing and often stressful environment. To this end, *E. coli* has evolved mechanisms capable of sensing and coping with environmental stress. Firstly, *E. coli* has acquired a diverse set of transcription factors dedicated to sensing environmental stress and altering gene expression to increase stress resistance and/or decrease the stress threat. These transcription factors include OxyR (H_2O_2 stress)[300], SoxR and SoxS [300] (redox cycling compound stress), Fur (iron deprivation) [301], GadE (acid stress) [302] and OmpR (osmotic stress)[303]. Additionally, *E. coli* has also evolved a set of alternative σ factors that change the specificity of its unique RNA polymerase toward stress response genes. These include σ^{24} (Extreme heat shock), σ^{28} (Flagellar synthesis/chemotaxis), σ^{32} (Heat shock), σ^{38} (General starvation) and σ^{54} (Nitrogen regulated genes)[304]. Interestingly, *E. coli* has also been shown to drastically change the structure of its chromosome under stress conditions with global impacts on gene expression[281]. For example, the degree of DNA supercoiling has been found to respond to several stresses including oxidative stress[305], osmotic stress[306], heat shock[307], cold shock[308] and nutrient starvation[309]. As an extreme

example, the NAP Dps self-aggregates under starvation conditions while bound to DNA and thus condenses the chromosome and physically protects it from damage[310].

Despite recent advances in the elucidation of the *E. coli* chromosome structure under non-stressed conditions[311], the 3D structure of the chromosome under stress conditions, its impact on gene expression and the switching mechanisms between different chromosomal structures have all been far less understood. This knowledge is indispensable for increased understanding of the functioning of natural chromosomes under stress conditions.

1.4 Objectives of Annex I

The goal of this project was to test the Transcription Based Soleonoid framework. This was done in four steps: (i) defining periodicity evaluation criteria using control datasets (Section 3.1.1) (ii) analyzing gene periodicity in the OxyR, SoxS, Fur, GadE and OmpR regulons (Section 3.1.2) iii) using these results to choose a suitable candidate regulon for further investigation (Section 3.2.1) and iv) testing the effect of displacing the stress response regulator on the stress response (Section 3.2.2).

2. Materials and methods

2.1 Periodicity analysis

2.1.1 Periodicity analysis using GREAT :SCAN :patterns

In order to investigate the interplay between periodic chromosome organization, chromosome expression and chromosome structure, our group has developed GREAT : SCAN :patterns, the main computational tool used in this work. GREAT : SCAN :patterns is an online tool that exhaustively analyzes all periodic patterns that can be detected from the positions of coregulated genes in the genome[312]. The explanation here is largely based on [312] and [313].

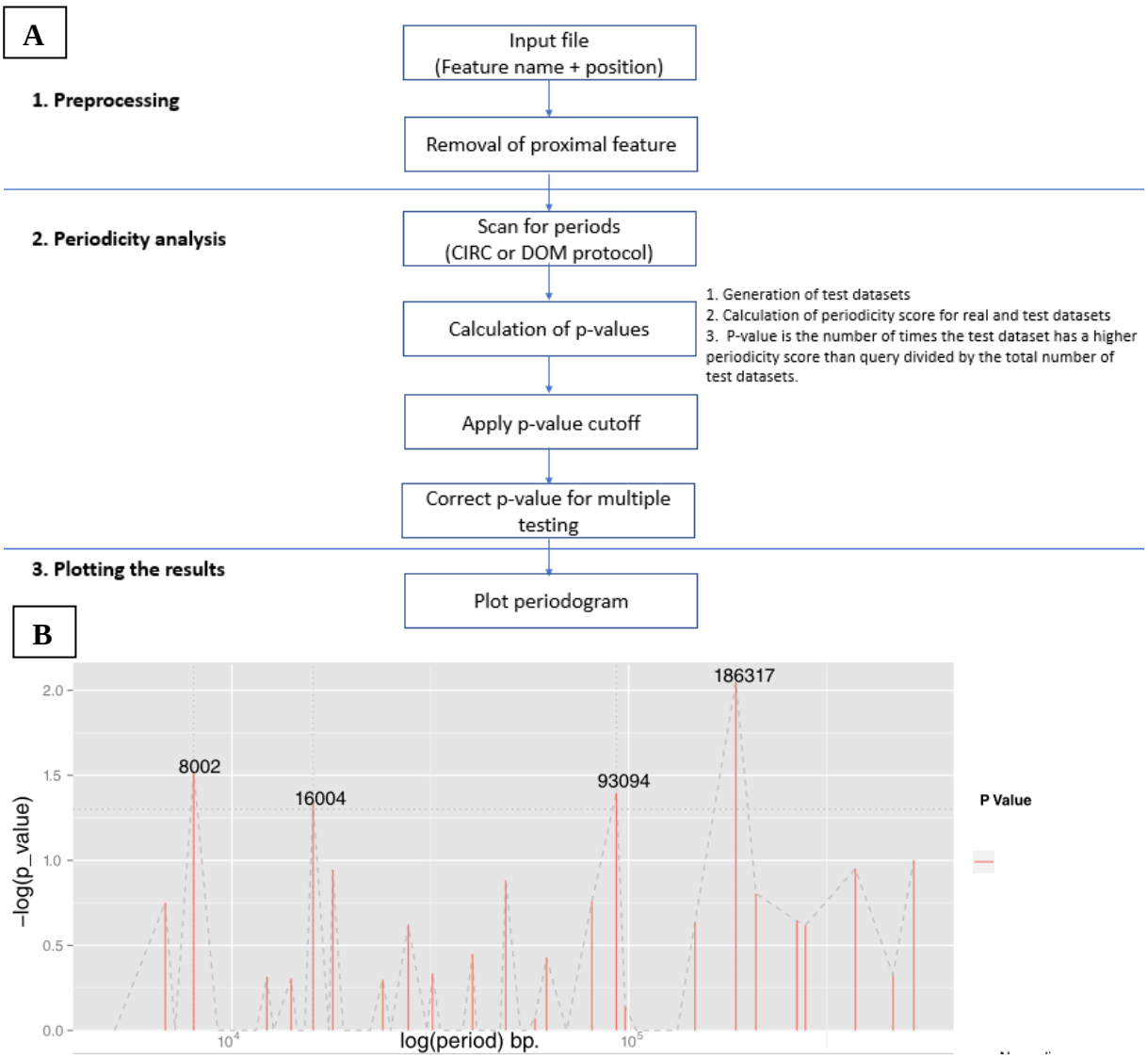
The input of GREAT : SCAN :patterns simply consists of a file with gene identifiers and their chromosome positions[312]. Depending on the nature of the data that was used to create the input file, chromosome positions may mean the 5’ end of the genes (defined by transcriptome data and 3C interaction data) or the middle of a transcription factor binding site (ChIP data).

Briefly, GREAT: SCAN: patterns works in two steps for our purposes, as shown in figure 4.

The first step involves the detection and visualization of all periods along the full length chromosome. Initially, the data is preprocessed to remove the effects of proximal gene identifiers[312]. This is necessary, because proximal gene identifiers can artificially increase the p-value of the periods when paired with another, sufficiently distant gene identifier[312]. This can lead to false positive periods.[313] To avoid this problem, the program replaces a set of proximal gene identifiers by its barycentre. Proximity is defined as two times the average gene distance (*E. coli*:1000 bp)[313]. Next, the p-value of each period is calculated. First, the program produces test datasets containing randomly generated

chromosome positions[312]. Positions are drawn from a uniform distribution[312], i.e. every position between 0 and the end of the chromosome has equal probability to be drawn. Next, periodicity scores are calculated as in [314]. The p-value is then defined as the number of times the test dataset had a higher periodicity score than the real dataset divided by the total number of test datasets[312]. Next, the p-values are corrected for multiple hypothesis testing using a period-length-dependent correction[312]. Finally, if at least one period meets the preprogrammed threshold (standard threshold: p-value < 0,05 or S=-log(p-value) > 1,3), all of the periods are plotted in a periodogram[312].

In a second step, the chromosome is scanned to determine the length of each region over which a periodicity occurs[312]. To do this, GREAT : SCAN: patterns uses a variable size sliding window that scans the chromosome for regions in which periodic organization is found, starting with a window size of 10 kb that geometrically increases with a constant factor and ends with a sliding window that encompasses 95 % of the chromosome[312]. Period p-values are subsequently calculated in the same way as for the whole chromosome[312].



C

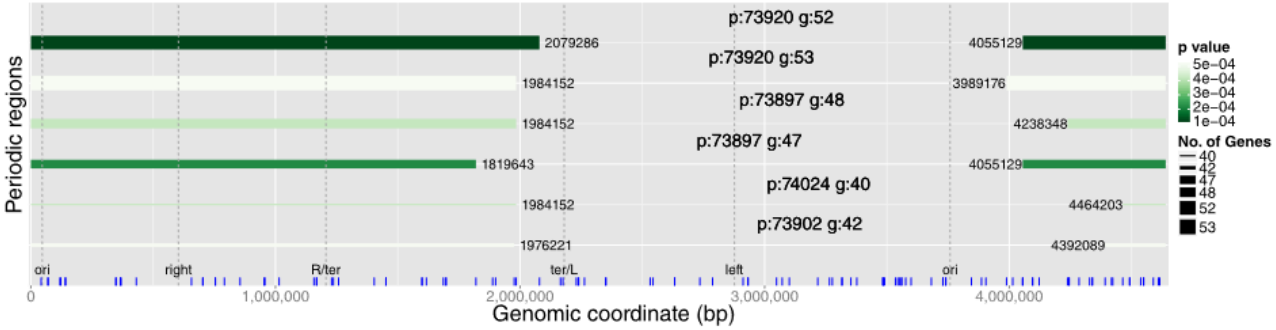


Figure 4: GREAT: SCAN: patterns (A) Flowchart of the algorithm : GREAT : SCAN : patterns takes a file with two columns as input. Next, the program scans for periods and calculates p-value. First, the program produces test datasets containing randomly generated chromosome positions. Positions are drawn from a uniform distribution, i.e. every position between 0 and the end of the chromosome has equal probability to be drawn. Next, periodicity scores are calculated as in [314]. Finally, a p-value is calculated. The p-value is defined as the number of times the test dataset had a higher periodicity score than the real dataset divided by the total number of test datasets. Next, the p-values are corrected for multiple hypothesis testing using a period-length-dependent correction. Finally, if at least one period meets the preprogrammed threshold (standard threshold : p-value < 0,05 or -log(p-value) > 1,3), all of the periods are plotted in a periodogram. The periodogram contains a significance score (-log(p-value)) on the y-axis and a log(period length) on the x-axis (B) Example output for a global periodogram for the CRP genes. Taken from [312] (C) Example output for a local periodogram for the CRP genes. Taken from [312]

2.1.2 Origin and preparation of datasets for periodicity analysis

Three different types of datasets were used as input for GREAT :SCAN :patterns, namely : synthetic datasets, ChIP-exo datasets and RegulonDB datasets.

2.1.2.1 Synthetic control datasets

Synthetic datasets are datasets created to test the signal-to-noise of the periodicity analysis program used in this work. These datasets fall into two categories: negative control datasets and positive control datasets. Figure 5 gives a visual overview of them. Both negative and positive control datasets were made using the open-source statistical tool R [315,316]. These datasets were both .csv files consisting of a column of unique identifiers (‘mock_position’) and a column of numbers representing chromosomal coordinates. In the case of negative control datasets, chromosomal coordinates were generated using the following formula :

$$NC = floor(runif(1, GL))$$

With :

NC : chromosomal coordinate for the negative control dataset

GL : The genome length of E.coli MG1655, i.e. 4641652

Runif : A randomly drawn number from a uniform distribution between two numbers i.e 1 and GL

Floor: The result is rounded down to the closest whole number.

Chromosomal coordinates for the positive control datasets on the other hand were generated using the following formula :

$$PC = floor(runif\left(1, \frac{GL}{P}\right))x\ P$$

With :

PC : chromosomal coordinate for the positive control dataset

GL : The genome length of *E. coli* MG1655, i.e. 4641652

P : The period that is programmed in the positive control dataset, i.e. 8000

Runif : A randomly drawn number from a uniform distribution between two numbers i.e 1 and GL/P

Floor: The result is rounded down to the closest whole number.

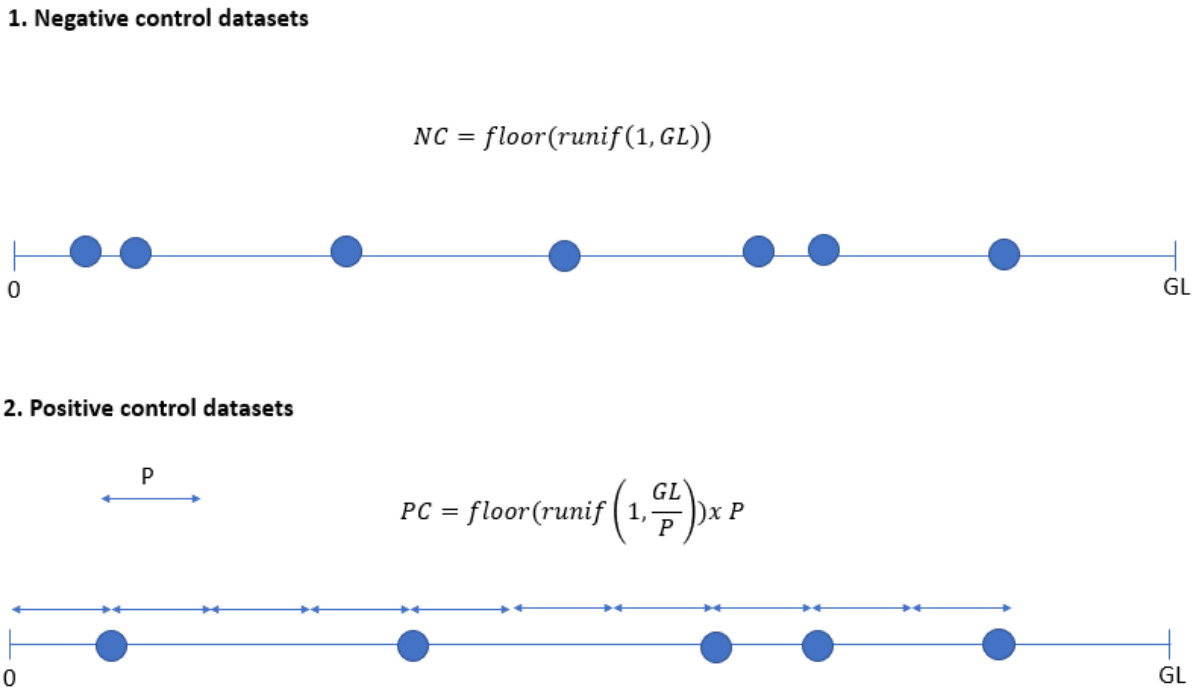


*Figure 5 : Synthethic datasets are generated to test signal to noise ratio of GREAT. Two datasets are generated : negative and positive control datasets. (1) Negative control datasets consist of points randomly distributed along the chromosome and should not display periodicity. (2) Positive control datasets consist of randomly generated points that are always n*P separated where n is a whole number. They should therefore display periodicity of P. These datasets are generated using the indicated formulas in the statistical program R*[315,316].

2.1.2.2 Chromation Immuno Precipitation – exonuclease (ChIP-exo) datasets

Chromatin Immuno Precipitation-exonuclease (ChIP-exo) is an experimental technique used to determine the exact binding position of a transcription factor on the chromosome[317]. Figure 6 gives a detailed overview of the approach[317]. Since this approach allows for the identification of direct regulatory targets of a transcription factor, these datasets would likely be free of contamination of indirect targets that may complicate the periodicity analysis. Hence, ChIP-exo datasets were picked as the datasets of choice and extracted from the following publications [300] [301] [302] [303]. The exact binding positions of OxyR, SoxS, GadE, Fur and OmpR were calculated by taking the average of the begin and end coordinates of the binding region. These coordinates were remapped on the newest version of the *E. coli* K12 MG1655 chromosome (NC_000913.3) and saved with a gene identifier (name of the closest gene) as a .csv file for further analysis.

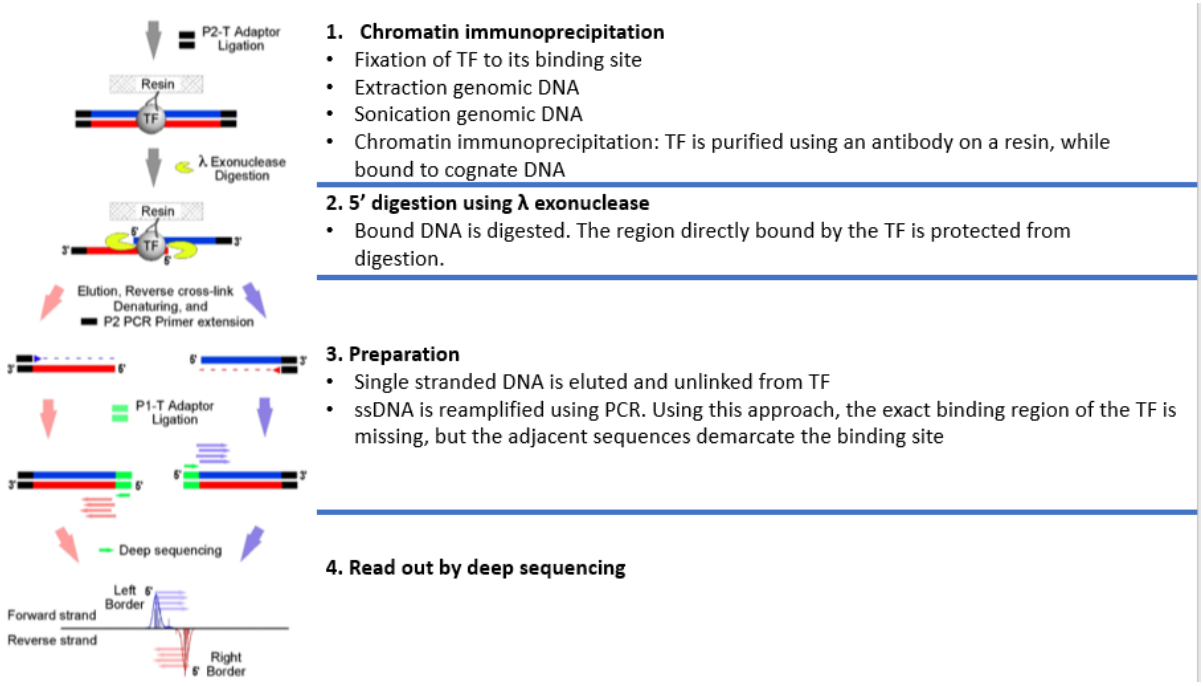


Figure 6 : ChIP-exo datasets : ChIP-exo is a technique for determining transcription factor (TF) binding sites with near nucleotide precision. Different steps of the procedure are shown.

2.1.2.3 RegulonDB datasets

RegulonDB is a database that collects and organizes all the experimental knowledge on regulation of transcription in *E.coli* [318]. Among other things, this database contains a dataset with all known or predicted transcription factor binding sites [319]. First, the dataset with transcription factor binding sites was downloaded. Next, the dataset was curated using a home-made R-script to only keep binding sites that have strong enough experimental support using the RegulonDB evidence classification code (introduced in [320]). Different criteria for the strength of evidence for a binding site were compared. As a first criterion, only binding sites that had 2 strong or 3 weak forms of evidence in favor of it were considered, as used in [312]. As a second criterion, only binding sites that had at least 1 or 2 strong pieces of evidence combined with evidence from at least one *in vivo* experiment were considered. Since the theory of periodicity predicts that well-positioned weak binding sites may outperform ill-positioned strong binding sites, evidence from *in vivo* experiments was given more weight for this analysis than RegulonDB generally does. After this curation step, the R-script was used to split the curated datasets by transcription factor and put them in a format amenable to GREAT : SCAN: patterns analysis.

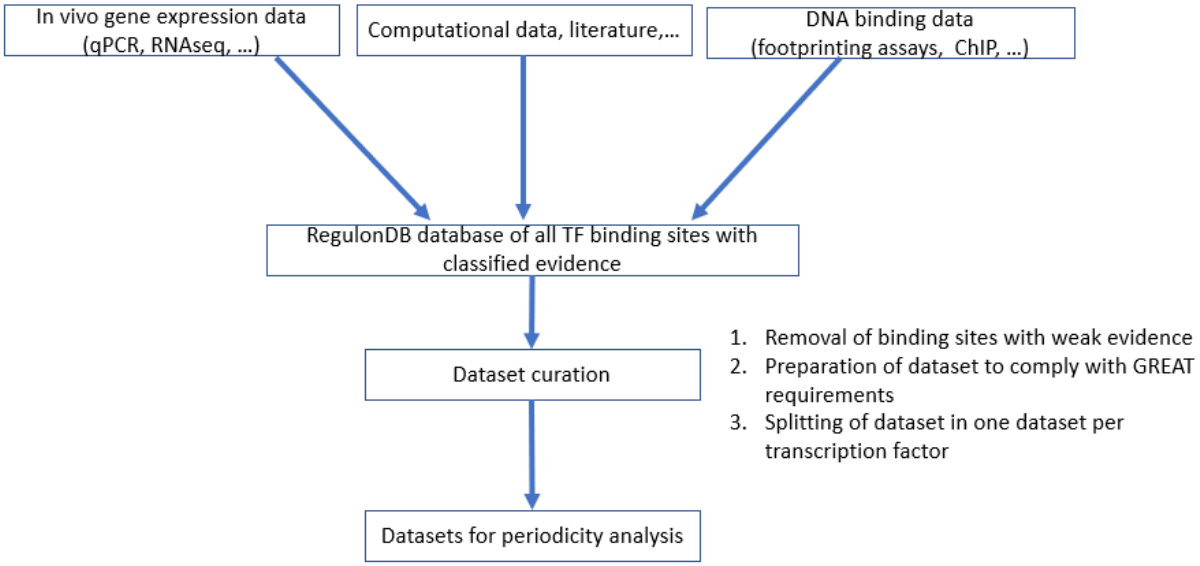


Figure 7 : RegulonDB datasets : RegulonDB is a database collecting all information on gene regulation in E.coli. It takes in data from gene expression analysis, DNA binding analysis, computational prediction etc and makes a list of TF binding sites with the evidence that supports it. This list was downloaded and curated. First, TF supported by weak evidence was removed (see text for criteria). Second, the list was reduced to two columns (feature identifier and chromosome position) in order to comply with GREAT. Finally, the list was split by TF in order to create datasets ready for periodicity analysis.

2.2 Bacterial strains and culture conditions

A list of the strains used in this work can be found in Annex I supplementary Table 1. A list of primers used can be found in Annex I supplementary table II. *E. coli* cells were grown at 37 ℃ in liquid LB under strong shaking (200-230 rpm). Antibiotics were used at the following concentration unless specified otherwise: spectinomycin (Sp, 60 µg/ml), kanamycin (Kan, 50 µg/ml), ampicillin (Amp, 100 µg/ml) and phleomycin (Ph , 10 µg/ml in low salt LB). Induction molecules were used at the following concentrations unless specified otherwise: IPTG (IPTG, 2 mM) and arabinose (ara, 1%)

Recombination competent cells were prepared by growing overnight cells containing pTKRED on LB + Sp at 30 ℃. The next day, the liquid cultures were diluted 1/1000 in LB + Sp + IPTG and grown at 30 ℃. At OD_{600nm} = 0.6, 20 mL of cell culture was chilled on ice for 10 min. The cell culture was spun down (4000 rpm, 4℃, 6 min) and supernatant was removed. Next, cells were washed in 10 ml ice cold $CaCl_2$ (1 M), spun down (4000 rpm, 4℃, 6 min) and resuspended in 5 ml ice cold $CaCl_2$ (1 M). Next, cells were spun down again and resuspended in 500 µl ice cold $CaCl_2$ (1 M) + glycerol (25%). The cell suspension were then aliquoted in 50 µl aliquots for electroporation and recombination.

ΔoxyRS was created by amplifying the excisable *PhleoR* cassette of plasmid 36 using primers with overhangs homologous to the flanking sequences of *oxyRS* (pSH148-149, Q5, NEB, standard reaction conditions). After the PCR reaction, the template plasmid 36 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37℃). The PCR product was then purified using the NEB DNA cleanup kit. The purified PCR product was then electroporated in recombination competent strain 74 (MG1655 + pTKRED, 1800 V) and grown overnight on LB. The liquid culture was plated on low salt LB + Ph and grown overnight on 37 ℃. The *PhleoR* marker was then removed by (i) λ RED induction to promote homologous recombination inside the cassette (2h, LB + Sp + IPTG) and (ii) excising the cassette with I-SceI (overnight, LB + Sp + IPTG + ara) using the strategy outlined in [321]. Removal of the *PhleoR* cassette was screened for using grid plating to check for loss of Ph resistance. pTKRED was cleared from this background by overnight growth on LB at 37℃ and a grid plate screen to check for loss of Sp resistance.

Plasmid 61 was created by first amplifying *oxyRS* (which contains the divergently transcribed *oxyR* and *oxyS* genes) using primers with overhangs homologous to plasmid 35 (excisable Kan cassette) (pSH158-159, Q5, NEB, standard reaction conditions). Template plasmid 35 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37℃) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Next, plasmid 35 was linearized using inverse PCR (iPCR) (pSH156-157, Q5, NEB, standard reaction conditions). Template plasmid 35 was again removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37℃) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Next, the purified *oxyRS* PCR product was mixed with the purified

iPCR product of plasmid 35 and assembled using HiFi assembly (NEB, 50°C, 15 min). The assembly product was transformed in *E. coli* strain DH5α and positive transformants were selected for KanR. Plasmid 62 was created by introducing terminators on both sides of the *oxyRS-Kan* cassette in plasmid 61 using two rounds of iPCR. The terminators were chosen from [322] to have >98% termination efficiency in the direction they were inserted. Briefly, iPCR was performed on plasmid 61 to introduce terminators on the left side of the cassette (pSH178-179, Q5, NEB, standard reaction conditions). Template plasmid 61 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37°C) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Next, the PCR product was phosphorylated using T4 polynucleotide kinase (NEB, 10 U, 30 min, 37°C). The T4 polynucleotide kinase was then heath inactivated (20 min, 65°C). The PCR product was chilled on ice for 2 min and T4 DNA ligase was added (NEB, 20 U, 2h, room temperature). The ligation products were then transformed in DH5α and positive transformants were selected for KanR. The intermediate plasmid (with left side terminators) was checked by sequencing and used as a template for introducing the right side terminators using the same procedure (pSH182-183).

Plasmid 92 was created by transferring the *oxyRS-Kan-terminator* cassette from plasmid 62 to the backbone of plasmid 91. First *oxyRS-Kan-terminator* was amplified using primers with overhangs homologous to plasmid 91 (pSH220-221, Q5, NEB, standard reaction conditions). Template plasmid 62 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37°C) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Next, plasmid 91 was linearized using inverse PCR (iPCR) (pSH222-223, Q5, NEB, standard reaction conditions). Template plasmid 91 was again removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37°C) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Next, the purified *oxyRS-Kan-terminator* PCR product was mixed with the purified iPCR product of plasmid 91 and assembled using HiFi assembly (NEB, 50°C, 15 min). The assembly product was transformed in *E. coli* strain DH5α and positive transformants were selected for KanR.

Since deletion of *oxyRS* leads to increased mutagenesis[323,324], *oxyRS* movant strains (ORS31, ORS32, ORSFT8, ORSFT9 and ORSFT13) were constructed without a *ΔoxyRS* intermediate. This was done through two rounds of recombination with PCR product and KanR cassette removal. During the first round, the *oxyRS-Kan-terminator* was amplified from plasmid 92 using primers with overhangs homologous to their insertion location (see supplementary table II, Q5, NEB, standard reaction conditions). Template plasmid 92 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37°C) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). The purified PCR product was then electroporated in recombination competent strain 74 (MG1655 + pTKRED, 1800 V) and grown overnight on LB. Each transformation was performed three times. The liquid culture was plated on LB + Kan and grown overnight on 37 °C. Correct integration was then screened

for by colony PCR using external primers to each integration location (see supplementary table II, Dreamtaq, Thermoscientific, standard reaction conditions). The *KanR* marker was then removed by (i) λ RED induction to promote homologous recombination inside the cassette (2h, LB + Sp + IPTG) and (ii) excising the cassette with I-SceI (overnight, LB + Sp + IPTG + ara) using the strategy outlined in [321]. Removal of the *KanR* cassette was screened for using grid plating to check for loss of Kan resistance.

During the second round, the *oxyRS* genes in the original location were removed using the same procedure as for the *ΔoxyRS* strain (although using plasmid 35 and Kan selection).

ORS_init is a strain where the *oxyRS* genes are replaced by *oxyRS-Kan-terminator* cassette. This was constructed through two rounds of recombination with PCR product and one round of KanR cassette removal.

During the first round, the *oxyRS-Kan-terminator* was amplified from plasmid 92 using primers with overhangs homologous to the *oxyRS* location (pSH148-149, Q5, NEB, standard reaction conditions). Template plasmid 92 was removed using DpnI digestion (NEB, 20 U, CutSmart, 2h, 37°C) and the PCR product was purified using the Monarch DNA cleanup kit (NEB). Recombination of the PCR product into a recombination competent 74 strain (MG1655 + pTKRED) lead to misrecombination with the native *oxyRS* genes and loss of the right side terminators (Data not shown). Therefore, the purified PCR product was then electroporated in recombination competent strain *ΔoxyRS* + pTKRED (1800 V) and grown overnight on LB. Each transformation was performed 3 times. The liquid culture was plated on LB + Kan and grown overnight on 37 °C. Correct integration was then screened for by colony PCR using external primers to each integration location (pSH03-04, Dreamtaq, Thermoscientific, standard reaction conditions) and one external primer combined with a primer inside the region with the right side terminator (pSH04-156, Dreamtaq, Thermoscientific, standard reaction conditions).

During the second round, the *oxyRS-Kan-terminator* was amplified from the strain obtained during the first round using external primers that created a 1kb overhang on the left and a 2 kb overhang on the right to increase the probability of a proper recombination event (pSH03-240, Q5, NEB, standard reaction conditions). The PCR product was purified using the Monarch DNA cleanup kit (NEB). The purified PCR product was then electroporated in recombination competent strain 74 (MG1655 + pTKRED, 1800 V) and grown overnight on LB. Each transformation was performed three times. The liquid culture was plated on LB + Kan and grown overnight on 37 °C. Correct integration was then screened for by colony PCR using (pSH03-04, Dreamtaq, Thermoscientific, standard reaction conditions) and one external primer combined with a primer inside the region with the right side terminator (pSH04-156, Dreamtaq, Thermoscientific, standard reaction conditions). About 1/3 of the clones carried the correct integration. The *KanR* marker was then removed by (i) λ RED induction

to promote homologous recombination inside the cassette (2h, LB + Sp + IPTG) and (ii) excising the cassette with I-SceI (overnight, LB + Sp + IPTG + ara) using the strategy outlined in [321]. Removal of the *KanR* cassette was screened for using grid plating to check for loss of Kan resistance. The genotype of final transformants was verified by PCR and subsequent DNA sequencing.

2.3 Zone of inhibition experiments for measuring oxidative stress

Oxidative stress sensitivity was measured using the Kirby-Bauer method[325]. The *E. coli* strains were incubated overnight from glycerol stocks in Luria Broth medium at 37°C shaking at 220 rpm. The next day, they were re diluted in LB to an OD of 0.040 and incubated in a water bath at 37 °C, shaking at 200 rpm. At an OD of around 0.2, the cells were diluted 1/100 in 5 mL of LB and spread evenly on 3 Petri dishes containing LB with agar. Excess medium was subsequently removed. Next, a pipette tip was used to punch a hole in the center of the LB + agar layer. H_2O_2 dilutions (0 mM, and 100 mM and 1M) were subsequently added to the holes and the plates were incubated overnight at 37 °C.
The next day, the zone of inhibition of the plates were photographed using G-box with a ruler as internal control and saved under a loss-less tiff format. Diameters of inhibition were measured in Fiji using the ruler to set the scale.

3. Results

3.1 Periodicity analysis of stress response genes in *E. coli*

3.1.1 Defining periodicity evaluation criteria based on analysis of synthetic datasets

Before analyzing TF binding site datasets, GREAT :SCAN: patterns was first tested on synthetic datasets in order to define periodicity evaluation criteria. The synthetic datasets fell into two categories : positive control and negative control datasets. Positive control datasets consisted of coordinates that were programmed to always have a multiple of a distance P between them and therefore should give rise to period P. Negative control datasets consisted of coordinates that were drawn randomly with equal probability between 1 and the total length of the genome. Therefore, they should not give rise to a periodic signal.

3.1.1 Analysis of a positive and a negative control dataset: qualitative insights

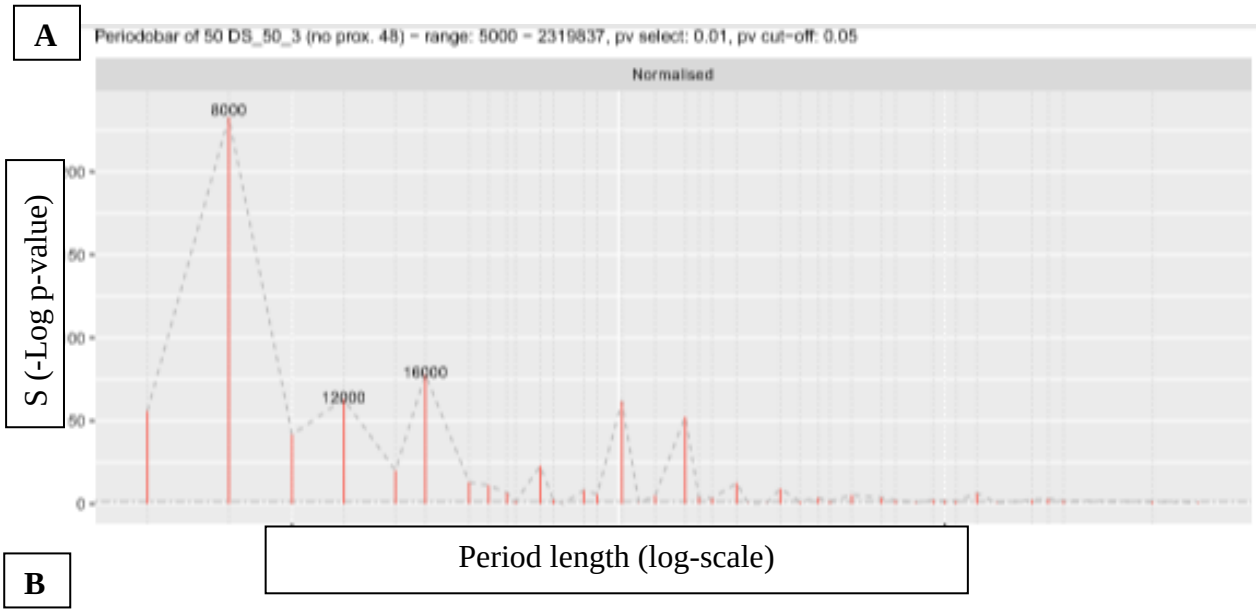
Figure 8 contains an example output for a positive control dataset of 50 coordinates with a preprogrammed period of 8kb (A+B) and an example output of a negative control dataset of equal size (C+D). The y-axis shows the significance score (-log(p-value)) and the period length (also depicted in log scale).

Fig 8 A+B shows that GREAT :SCAN :patterns can find the 8 kb period with a high degree of significance in the positive control dataset. The 8 kb period has a significance score S > 200 and local periods of 8 kb with an S > 50 are readily observed. This is way more significant than the threshold (S = 1,3 corresponding to a p-value = 0.05).

However, other periods that are not obviously multiples of 8kb (6 kb and 6.4 kb etc) are also observed with great significance (S > 30). This result suggests that dominant periodic signals can be identified by their p-value, but a true secondary periodic signal in a dataset will be hard to distinguish from one of these non-obvious multiples of the primary period.

By contrast, figure 8 (C+D) shows that GREAT :SCAN :patterns can still find significant periods in a random dataset albeit with a much lower S score. Nevertheless, S scores above 2-3 are readily observed in this dataset (compare to the threshold S=1,3).

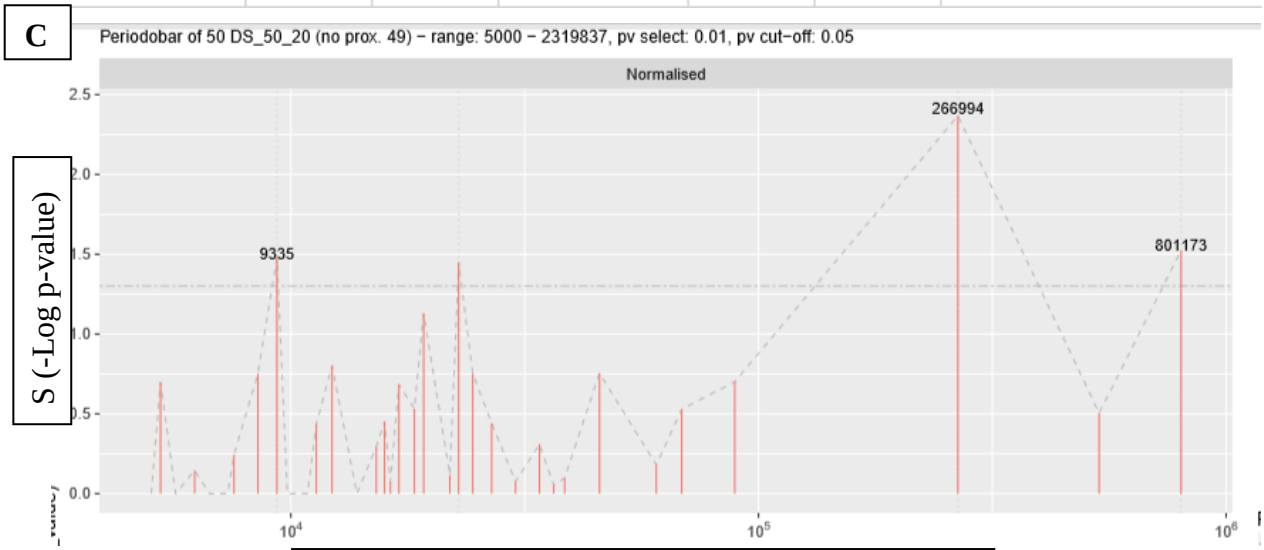
A



B

Period (bp)	begin	end	regionLength	Pvalue	S score	Coverage
8000	4240000	3120000	3819447	1E-58	58	0.78
16000	4240000	3120000	3819447	2E-52	51.69897	0.78
12000	4240000	3120000	3819447	3E-39	38.52288	0.78
6000	4240000	3120000	3819447	1E-35	35	0.78
6400	4240000	3120000	3819447	7E-33	32.1549	0.78
10000	4240000	3120000	3819447	1E-26	26	0.78
8000	4152000	3120000	3731447	1E-63	63	0.8
16000	4152000	3120000	3731447	5E-52	51.30103	0.8
12000	4152000	3120000	3731447	7E-37	36.1549	0.8

C



D

Period (bp)	begin	end	regionLength	Pvalue	S score	Coverage
280503	1089372	4558417	3472225	0.003	2.5228787	0.66
558722	1089372	4558417	3472225	0.005	2.30103	0.66
269760	2304296	4625386	2371579	0.001	3	0.5
269376	2304296	4420038	2155981	0.003	2.5228787	0.42
15143	4382551	1395503	1781803	0.005	2.30103	0.46
15142	4111858	1130721	1781803	0.003	2.5228787	0.46

Figure 8: Example output for synthetic datasets: (A+B) Output for a positive control dataset with 50 positions and a preprogrammed period of 8 kb (A) Periodogram of global periods. X-axis depicts period size and y-axis period significance as log p-value. (B) Local periods in tabular format with from left to right (i) period length (ii) Left coordinate of the periodic region (iii) Right coordinate of the periodic region (iv) Length of periodic region (v) p-value (vi) S-score of the period and (vii) Periodic region coverage: the fraction of the number of coordinates inside the periodic region vs the total number of coordinates (C+D) Output for a negative control dataset

3.1.2 Definition of periodicity evaluation criteria: global periods

To get an idea of the generality of these observations, these experiments with synthetic datasets were repeated 10-20 times for datasets of different sizes. The investigated size range was between 5 and 100 datapoints, which was a relevant range for the datasets investigated in this work (see next sections). Figure 9 shows the global properties of dataset controls. Figure 9A shows that regardless of dataset size, significant periods with S > 1,3 are often observed in negative control datasets. Figure 9B shows the same analysis for positive control datasets. Here, we see that the average S-score of the programmed period observed in these datasets reaches a minimum for datasets with 20 coordinates and then increases again as dataset size increases. Nevertheless, the average S score in positive control datasets of different sizes ranges between 150 and 300, demonstrating that GREAT SCAN :patterns has a good signal to noise ratio. Figure 9C shows that with increasing dataset size, the number of datasets that don't reach the significance threshold steadily increases for negative control datasets. Still, around 60% of the datasets with 100 coordinates produces a period with a S-score > 1.3.

These results suggested that the preprogrammed threshold of S > 1.3 was not sufficient to exclude false positives. This lead to the following evaluation criterium which was also used in [287] :

A global period is considered significant if it has an S-score > 4.

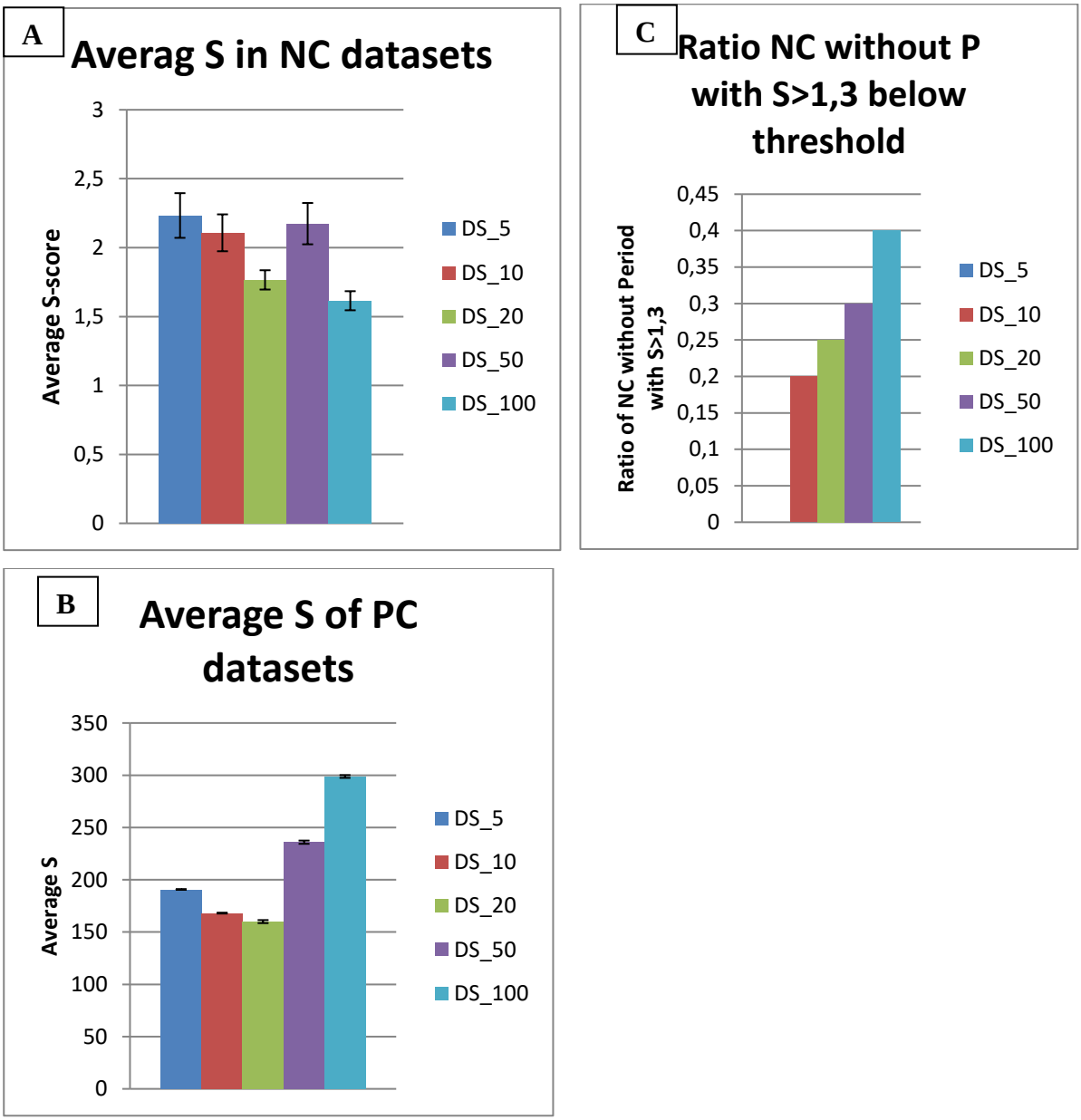


Figure 9 : Definition of global periodicity criteria by analyzing positive and negative control datasets of different sizes : (A) Average S-score (negative logarithm of p-value) of negative control datasets of 5 to 100 positions (B) Average S-score of the preprogrammed periods in positive control datasets between 5 and 100 positions (c) Ratio of datasets between 5 and 100 positions that failed to produce periods with S > 1,3 (p-value < 0,05).

3.1.3 Definition of periodicity evaluation criteria: local periods

Next, the same analysis was performed to define criteria for evaluating local periods. Local periods can be evaluated both on their p-value and their coverage (Number of periodic points within the region). Figure 10 shows the results of this. Figure 10A and B show the average S score and coverage respectively, when top periods are evaluated on S score alone. It shows that the average p-value increases and the average coverage decreases with increasing dataset size.

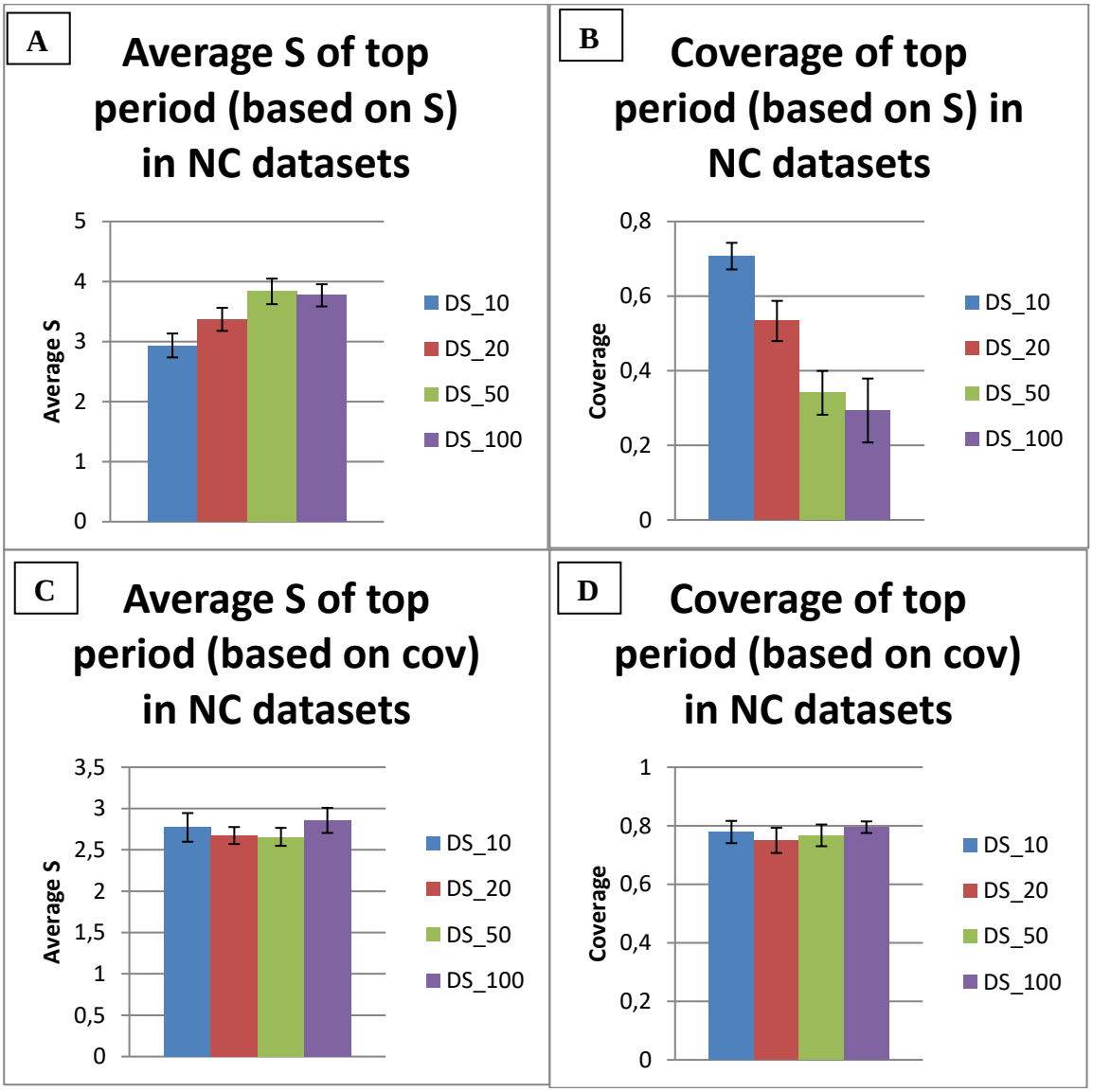


Figure 10 : Definition of local periodicity criteria by analyzing negative control datasets of different sizes. (A) Average S score of top period observed in negative control datasets between 5 and 100 points when the top period is evaluated based on S-score (B) Coverage (defined as the ratio of genes within the periodic region to total number of genes) of top period when the top period is evaluated based on S-score. (C) Average S-score of top period observed in negative control datasets between 5 and 100 points when the top period is evaluated based on coverage. (D) Coverage (defined as the ratio of genes within the periodic region to total number of genes) of top period when the top period is evaluated based on coverage.

Conversely, figure 10C and D show the average p-value and coverage of periods selected for their coverage. Together, these pictures show that local periods with either a decent S-score (>4) or a good coverage occur in negative control datasets.

Local periods must therefore be evaluated on both criteria at the same time. This leads us to the local period evaluation criterium :

A local period is considered significant if it has an S-score >4 and a coverage > 50%.

In this section, evaluation criteria for periods were defined using insights gained from the analysis of synthetic datasets. These will be applied to query datasets in the next sections.

3.1.2 Periodicity analysis of stress response regulons

3.1.2.1 Periodicity analysis of ChIP-exo datasets

The first goal of this work was to study the periodic organization of the binding sites of several stress response regulators on the chromosome. To this end, the exact binding positions of OxyR, SoxS, GadE, Fur and OmpR, were extracted from Chromatin Immuno Precipitation (ChIP) experiments[300] [301] [302] [303], remapped on the newest version of the *E. coli* K12 MG1655 chromosome (NC_000913.3) and analyzed with GREAT:SCAN:patterns. Neither global nor local periodicity analysis produced periods that would be considered significant using the evaluation criteria mentioned above although some local periods come close. This includes the 8.9 kb local period of Fur (Fe replete), the 5.8 kb local period of Fur (Fe starv) and the 28.5 kb local period of GadE. Nevertheless, these results were not considered enough to support the conclusion that these regulons are periodically organized.

Table 1 shows a summary of the periods observed in each dataset. Table 1 : Periodicity analysis of key stress response regulators from ChIP-exo binding data. (Left) Length and S-scores of the best global periods that were observed. (Right) Length, S-scores and coverage within the best local periodic regions that were observed. The color codes denote the regulon in which the period was observed.

	Global periods		Local periods		
Regulon	Period	S-score	Period	S-score	Coverage
SoxS (24)	15	1.50	9.3	2.40	0.63
	40.5	1.34	40.5	2.52	0.83
	54.4	1.82	81	2.52	0.63
	61.7	1.60	121.4	2.30	0.79
	81	2.05			
	202.5	1.63			
Fur (Fe replete) (110)	8.6	1.65	8.9	3.40	0.29
	227.3	1.39	20.3	3.00	0.75
			228.6	2.52	0.68
Fur (Fe starv) (59)	148.1	1.38	5.8	3.22	0.88
	237.1	1.62	227.8	2.40	0.64
OmpR (24)	13.2	2.17	10.5	2.70	0.38
	26.5	1.56	13.2	2.52	0.92
			15.2	2.40	0.63
GadE (14)	325,8	1.48	5.6	3.00	0.79
			14.2	2.40	0.71
			28.5	4.00	0.71
			149.9	3.00	0.79
			342.5	2.52	0.86
OxyR (21)	16.2	1.94	29.4	2.70	0.52
			30.8	2.52	0.52
			68.1	2.70	0.52

Therefore, two *in silico* approaches were undertaken to analyze further this study. In the first, new datasets for these same regulons were obtained from RegulonDB to see if significant periods could be obtained from them or if the same periods would reoccur in them (Section 3.1.2.2). In the second approach, harmonic analysis was used to see if different stress response systems that are activated under the same conditions also have a common periodic organization (Section 3.1.2.3).

3.1.2.2 Periodicity analysis of RegulonDB datasets

Periodicity was analyzed in RegulonDB datasets of the same stress response regulons as in section 3.1.2 and compared. Only binding sites with a strong degree of evidence were analyzed. This includes binding sites with 2 strong or 3 weak evidences[312], 1 strong +1 *in vivo* evidence and 1 *in vivo* evidence. Table 2 shows the results of this for global periodicity analysis and table 3 shows the results for local periodicity analysis. These results show that the same periods are rarely found in more than one dataset of the same stress response regulon. Nevertheless, we do observe that the 40.5 kb SoxS global period, the 7.9 kb OxyR global period, the 25.7 kb OxyR local period and the 8.6 kb Fur global period (which is 1/10th of the 86-87 kb periods observed) are found in more than 1 dataset, providing some extra support for these periods.

Table 2 : Comparison of the global periods observed for key stress response regulators in (Left to Right) ChIP-exo results (Section 3.1.2), RegulonDB datapoints with either 2 strong or 3 weak interactions, RegulonDB with 1 strong and 1 in vivo interaction and RegulonDB with 1 in vivo interaction. Data is depicted as in table 1. Matches between periods within different datasets of the same regulon are marked in red.

	ChIP-exo		2 Strong or 3 weak		1 Strong and 1 In Vivo		1 In Vivo	
Regulon	Period	S-score	Period	S-score	Period	S-score	Period	S-score
SoxS (24)	15	1.50	16.5	1.77	40.5	2.78	103.6	1.92
	40.5	1.34	19.7	1.80	60.5	1.52		
	54.4	1.82			89.1	1.35		
	61.7	1.60			121	1.69		
	81	2.05			181.6	1.40		
	202.5	1.63			242	1.43		
Fur (Fe replete) (110)	8.6	1.65	87.3	2.17	86.5	2.20	None	
	227.3	1.39	174.5	2.62	124.2	1.79		
			232.9	1.49				
Fur (Fe starv) (59)	148.1	1.38	349	1.35				
	237.1	1.62						
OmpR (24)	13.2	2.17	None				7.8	1.83
	26.5	1.56					10.5	2.21
GadE (14)	325,8	1.48	None		148.8	1.56	7	1.56
							10	2.06
							31.1	1.64
OxyR (21)	16.2	1.94	None		7.9	2.61	7.9	3.27
					25.7	1.31	15.7	1.44

Table 3: Comparison of the local periods observed for key stress response regulators in (Left to Right) ChIP-exo results (Section 3.1.2), RegulonDB datapoints with either 2 strong or 3 weak interactions, RegulonDB with 1 strong and 1 in vivo interaction and RegulonDB with 1 in vivo interaction. Data is depicted as in table 1. Matches between periods within different datasets of the same regulon are highlighted in red.

	ChIP-exo			2 strong or 3 weak			1 strong and 1 in vivo			1 in vivo		
Regulon	Period	S-score	Coverage	Period	S-score	Coverage	Period	S-score	Coverage	Period	S-score	Coverage
SoxS (24)	9.3	2.40	0.63	None			283	3.10	0.83	14.9	3.30	0.54
	40.5	2.52	0.83									
	81	2.52	0.63									
	121.4	2.30	0.79									
e replete)	8.9	3.40	0.29	None			None			None		
	20.3	3.00	0.75									
	228.6	2.52	0.68									
(Fe starv)	5.8	3.22	0.88	None			None			None		
	227.8	2.40	0.64									
OmpR (24)	10.5	2.70	0.38	None			None			None		
	13.2	2.52	0.92									
	15.2	2.40	0.63									
GadE (14)	5.6	3.00	0.79	None			None			None		
	14.2	2.40	0.71									
	28.5	4.00	0.71									
	149.9	3.00	0.79									
	342.5	2.52	0.86									
OxyR (21)	29.4	2.70	0.52	None			25.7	3.15	0.56	25.7	3.70	0.53
	30.8	2.52	0.52									
	68.1	2.70	0.52									

3.1.2.3 Harmonic analysis of ChIP-exo data

To better understand the relationships between the found periods, a harmonic analysis was performed to see which periods were multiples of other periods and therefore shared common chromosomal organizations. Harmonic periods were allowed to differ 1% maximally from anyone of its multiples. Periods were considered to be part of the same harmonic family if they were a harmonic of at least one period in the family. Related harmonic periods and their corresponding regulons are shown for each period in tables 4 and 5. Cis-harmonic periods are periods that were harmonic within a dataset and trans-harmonic periods were periods that were harmonic between datasets and could point to common periodic organizations. Using this analysis, we found a family of 8,1kb periods connecting the OxyR, SoxS and Fur (Fe replete) regulons. Other observed families are the 5,8 kb family in SoxS and Fur (Fe starv). These common periodic themes mirror known crosstalk between these regulons. Fe^{2+} is known to exacerbate oxidative stress, making its regulation a key part of the oxidative stress defense (See section 3.2.1 for more information on the oxidative stress response).

Table 4: Harmonic analysis of global periods obtained from analysis of ChIP-exo data of key stress regulators. Data is depicted as in table 1, with the addition of cis-harmonic periods (Period multiples observed within a dataset) and trans-harmonic periods (Multiples of periods observed in other datasets).

	Global periods			
Regulon	Period	S-score	Cis-harmonic	Trans harmonic
SoxS (24)	15	1.50		
	40.5	1.34	81, 121, 202.5	5.8 (Fur Fe starv)/20.3 (Fur Fe replete)
	54.4	1.82		
	61.7	1.60		
	81	2.05	40.5	5.8 (Fur Fe starv)/20.3 (Fur Fe replete)/ 16.2 OxyR
	202.5	1.63	40.5	
Fur (Fe replete) (110)	8.6	1.65		
	227.3	1.39		
Fur (Fe starv) (59)	148.1	1.38		
	237.1	1.62		
OmpR (24)	13.2	2.17	26.5	
	26.5	1.56	13.2	
GadE (14)	325,8	1.48		
OxyR (21)	16.2	1.94		81 (SoxS)

Table 5: Harmonic analysis of local periods obtained from analysis of ChIP-exo data of key stress regulators. Data is depicted as in table 1, with the addition of cis-harmonic periods (Period multiples observed within a dataset) and trans-harmonic periods (Multiples of periods observed in other datasets).

	Local periods				
Regulon	Period	S-score	Coverage	Cis-harmonic periods	Trans harmonic periods
SoxS (24)	9.3	2.40	0.63		
	40.5	2.52	0.83	81, 121, 202.5	5.8 (Fur Fe starv)/20.3 (Fur Fe replete)
	81	2.52	0.63	40.5	5.8 (Fur Fe starv)/20.3 (Fur Fe replete)/ 16.2 kb OxyR
	121.4	2.30	0.79	40.5	5.8 (Fur Fe starv)/20.3 (Fur Fe replete)
Fur (Fe replete) (110)	8.9	3.40	0.29		
	20.3	3.00	0.75		
	228.6	2.52	0.68	227.3	
Fur (Fe starv) (59)	5.8	3.22	0.88		40.5/81/121.4/202.5 (SoxS)
	227.8	2.40	0.64		227.3/228.5 (Fur Fe replete)
OmpR (24)	10.5	2.70	0.38		
	13.2	2.52	0.92	26.5	
	15.2	2.40	0.63		
GadE (14)	5.6	3.00	0.79		
	14.2	2.40	0.71	28.5	
	28.5	4.00	0.71	14.2	
	149.9	3.00	0.79		
	342.5	2.52	0.86		
OxyR (21)	29.4	2.70	0.52		
	30.8	2.52	0.52		
	68.1	2.70	0.52		

3.1.3 Summary

To sum up, GREAT found periods for the stress response systems from ChIP-exo data, but the S-scores were deemed insignificant based on our insights from the analysis of synthetic control datasets.

Analysis of datasets for these same regulons extracted from RegulonDB provided some additional support for OxyR, SoxS and Fur periods. On top of that, harmonic analysis also found common periodic layouts that might be reflective of known crosstalk between stress response systems, including OxyR-SoxS-Fur. It is therefore possible that GREAT:SCAN:patterns is not sensitive enough to detect periods in biological datasets with reasonable certainty, but that the periods found in the ChIP-exo datasets are still biologically relevant.

To test this possibility, one of these transcription factors was moved to different positions on the chromosome to test if the position of the gene impacts its regulation of the stress response.

3.2 Effect of transcription factor positioning on the oxidative stress response in *E.coli*

3.2.1 Choosing a suitable test regulon : OxyR-SoxS-Fur and the oxidative stress response

The goal of this section is to decide which transcription factor (OxyR, SoxS or Fur) is best suited for *in vivo* testing. First, the biology of oxidative stress and the oxidative stress response is discussed. Second, experimental and theoretical criteria are proposed to decide which transcription factor should be moved.

3.2.1.1 Introduction to oxidative stress and the oxidative stress response

Oxidative stress is caused by oxidants. Oxidants are substances that accept electrons from other molecules during oxido-reduction reactions. The most commonly encountered oxidant in biological systems since the advent of photosynthesis is oxygen (O_2). While O_2 itself is not very reactive with biological molecules, it can give rise to highly reactive partially reduced Reactive Oxygen Species (ROS)[327].

These damaging ROS molecules are the superoxide ion (O_2^-), hydrogen peroxide (H_2O_2) and the hydroxyl radical ($OH^\cdot$)[328] These molecules can react with (i) DNA, causing DNA damage and mutagenesis[329], (ii) amino acids [330] and lipids and most importantly (iii) Fe^{2+}, causing its oxidation into Fe^{3+} and the inactivation of Fe-S dehydratases[331] and mononuclear iron proteins[332], which fulfill critical roles in central carbon metabolism (CCM)[328].

Due to their damaging effects, bacteria often encounter exogeneous ROS molecules as part of an effective immune response in bacteria, plants and animals[328]. H_2O_2 is uncharged and can diffuse past the membrane of the bacterium[328]. O_2^-, on the other hand, is produced by redox-cycling molecules that can diffuse past the membrane and create O_2^- inside the targets cytoplasm[328].

Oxidative stress is then defined as an imbalance between the amount of ROS molecules in the cell versus the cell's ability to detoxify them. For this reason, all living aerobic cells have evolved defense mechanisms against oxidative stress and these defense mechanisms are collectively called the oxidative stress response[328].

In *E. coli*, the response to oxidative stress is controlled by the transcriptional regulators OxyR, OxyS, SoxR and SoxS[328]. OxyR can sense H_2O_2 through oxidation of two sensory cysteine residues[333,334], which leads to a conformational change and the activation of transcription of about 20-25 operons[300]. SoxR on the other hand, can sense the presence of redox cycling compounds, which leads to expression of SoxS and the activation of another 20 operons.[300] The effect of the members of the OxyR and SoxRS regulons are summarized in figure 11. From this figure, it is clear that the maintenance of iron metabolism is essential for an effective oxidative stress response. Consequently, the master regulator of iron metabolism Fur is both a target of OxyR and SoxS[300].

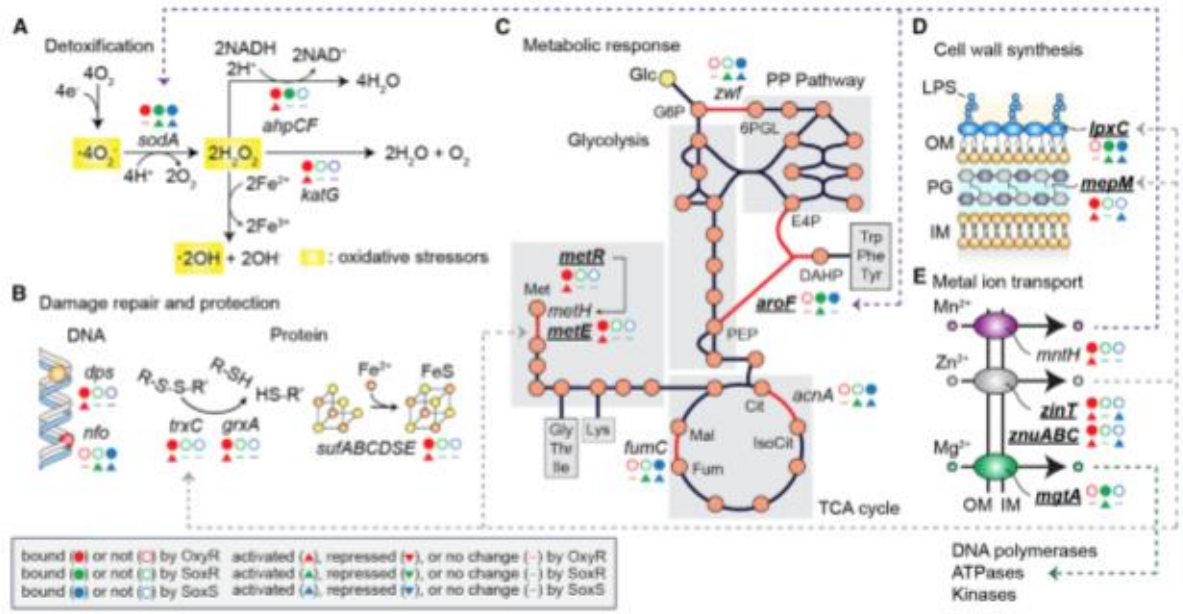


Figure 11: Biological processes affected by OxyRS and SoxRS targets. (A) Detoxification of ROS molecules. (B) Protection and repair of DNA and protein after oxidative stress. (C) Rerouting central metabolism to replace or overexpress enzymes sensitive to ROS or to produce reducing agents. (D) Cell wall synthesis. (E) Metal ion transport to replace Fe^{2+} by other divalent cations Picture was taken from[335]

3.2.1.2 Selection of regulon for testing the TBS framework

According to the TBS model factor oligomerization seems to be a necessary requirement for

colocalization. For experimental purposes, the best suited test regulon maximally impacts the oxidative stress response and minimally impacts cell function under non-experimental conditions.

SoxS is a member of the AraC family of regulators, but lacks its dimerization domain[336,337]. It induces gene expression by binding to the α subunit of the RNA polymerase binding as a monomer and scanning the DNA for a 20 bp degenerate sequence called the soxbox that is at the correct distance from the promoter[338–341]. Given this mechanism, it is unlikely that SoxS is vital for recruiting periodic genes of the regulon and thereby establishing a theoretical transcription factory and it was ruled out as an experimental test system.

Fur is part of its own family of transcriptional regulators, called the FUR family[342]. In general, Fur represses gene expression by binding to a palindromic AT rich sequence as a dimer. The binding of Fe^{2+} stabilizes the interaction of the Fur dimer with its cognate binding sequence[343]. This would make the Fur regulon a suitable candidate for testing. However, Fur is a global regulator that directly controls not only iron transport and metabolism, but also the expression of key metabolic genes that use Fe^{2+} involved in the TCA cycle and deoxyribonucleotide synthesis[301,342]. Disruption of Fur has been shown to profoundly change the transcriptional program of the cell under non-oxidative conditions[301]. These features make Fur a suitable, but not ideal regulon for testing the periodicity framework.

OxyR is a member of the LysR family of transcriptional regulators[323,344,345]. OxyR forms a homotetramer in *E.coli* [346] that undergoes a major conformational change when its sensory cysteine residues are oxidized[323,333]. In its oxidized form, OxyR recognizes four adjacent ATAG elements with 10 bp spacing[347], whereas it only binds to two in its reduced form[323,348]. It then cooperatively binds with RNA polymerase to stimulate gene expression of its target genes[323,349]. The OxyR core regulon consists of 20-25 operons and is extended to about 40 operons when the targets of its divergently transcribed *oxyS* RNA molecule is included[323,350]. OxyR mutants are hypersensitive to oxidative stress[351] and unfortunately prone to mutagenesis as a result[323,324]. Nevertheless, thanks to (i) the ability of OxyR to oligomerize, (ii) its profound importance for oxidative stress defense, (iii) its relatively small regulon that minimally impacts cell metabolism in non-experimental conditions apart from increased mutagenesis and (iv) the connections between oxidative stress, iron maintenance that were reflected in the periodicities observed in this work, the OxyR regulon was chosen as our test regulon.

3.2.2 Zone of inhibition experiments in *oxyRS* repositioned mutants

The goal of this experiment was to test if the repositioning of *oxyRS* impacts the oxidative stress response. In this first experiment, 3 strains with *oxyRS* near the origin (*ORS_init*, *ORS31* and *ORS32*)

are compared to 3 strains with *oxyRS* near the terminus (*ORS_FT8*, *ORS_FT9* and *ORS_FT13*) when grown in LB. This is an extreme repositioning which should not only disrupt periodic patterns, but also drastically alter copy number of *oxyRS* [352–354].

The effect of repositioning on the OxyRS stress response was tested using an initial set of zone of inhibition experiments with 0, 100 and 1000 mM H_2O_2. The results of these experiments are found in figure 10. Figure 10A and B show examples of plates without (A) and with (B) 1M H_2O_2 added the central hole for the wild type MG1655. This shows that 1M H_2O_2 can impede MG1655 growth in a ring around it, creating a zone of inhibition (ZOI).

Next, oxidative stress sensitivity is quantified by measuring the ZOI of MG1655, *ΔoxyRS* and the o*xyRS* movant strains at 100 mM H_2O_2. Figure 10C shows these results. *ΔoxyRS* has a much higher ZOI diameter, indicating an increased sensitivity to H_2O_2, as expected [355]. However, the ZOI diameter of the *oxyRS* movants is similar to the ZOI diameter of MG1655, showing no observable difference in oxidative stress sensitivity.

In order to improve the sensitivity of this test, the ZOI experiment was repeated when 1M was added to the central hole. These results are shown in figure 10D. Again, the differences observed are in the range of 1 to 2 mm, which is more likely explained by measurement inaccuracy than a real difference in oxidative stress sensitivity.

In conclusion, *oxyRS* repositioning does not seem to strongly alter sensitivity to oxidative stress, not even when the TF is moved to a location near the terminus.

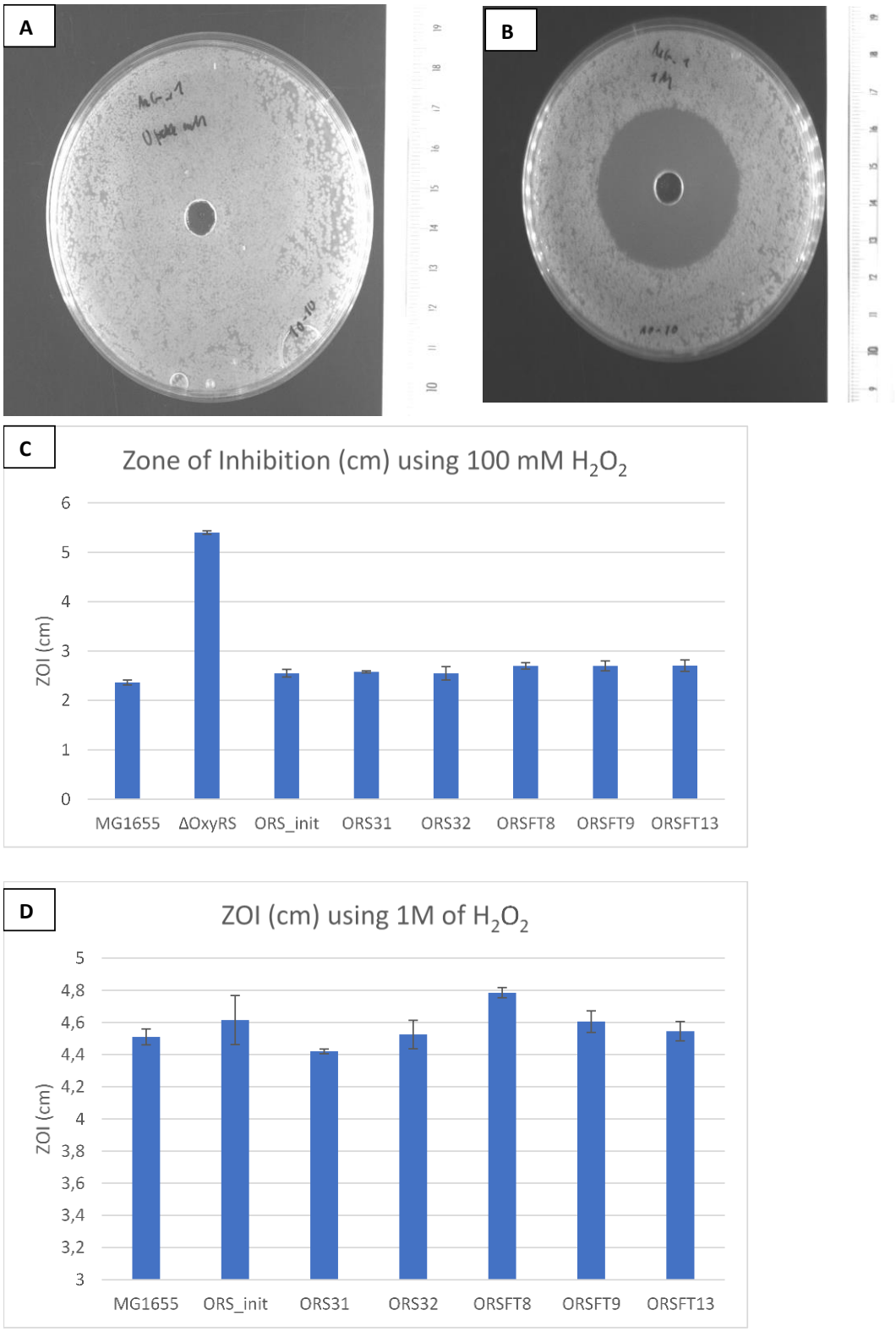


Figure 12 : Oxidative stress sensitivity of oxyRS movant strains using the Kirby-Bauer zone of inhibition test. (A-B) Comparison of the zone of inhibition of MG1655 when challenged with (A) 0 M or (B) 1 M of H_2O_2. (C) Diameter of the Zone of Inhibition (cm) of MG1655, ΔoxyRS and the oxyRS movants when challenged with 100 mM H_2O_2. (D) Diameter of the Zone of Inhibition (cm) of MG1655 and the oxyRS movants when challenged with 1M H_2O_2.

4. Discussion

The aim of this project was to test the Transcription Based Solenoidal (TBS) framework by assessing the impact of periodic gene organization on the stress response in *E. coli*. To this end, the presence of periodic gene organisations of various stress response regulons was analysed using our in-house software GREAT : SCAN :Patterns. These results were then used to pick the OxyRS regulon as a suitable experimental system for *in vivo* testing. Finally, the effect of displacing *oxyRS* on the stress response was assessed.

4.1 Periodic organization of (stress response) genes

Before analysing periodicity in stress response regulons, a set of synthetic datasets were analysed using GREAT : SCAN: Patterns to define periodicity evaluation criteria.

First, these experiments demonstrated that GREAT : SCAN: patterns has an excellent signal to noise, as the S-score (negative logarithm of the p-value of the period) for periods from the positive control dataset were at least 50 times higher than the ones from the negative control dataset. However, the results from the negative control datasets demonstrated that datasets with random positions still give rise to periods with an S-score of max 4. These last results are similar to the periodicity scores obtained by *Junier et al*, who used the same core algorithm as the one in GREAT : SCAN :Patterns[287]. For this reason, the same periodicity evaluation criterion of not considering periods with S score < 4 for global and local periods was used in this work. In the case of local periods, an extra coverage criterion of > 50% of all genes was added as it was observed that local periods over short stretches of the chromosome can give rise to highly significant periods in negative control datasets.

Armed with the evaluation criteria, ChIP-exo datasets of key stress regulons OxyR, SoxS, GadE, Fur and OmpR were analyzed and compared to the periodicity evaluation criteria. The ChIP-exo technique gives the direct regulatory targets of a transcription factor with near-nucleotide precision[317]. Hence, there is no need for applying any data clean-up procedures to prepare the data for periodicity analysis, other than the technical cut-offs performed in the publications. Unfortunately, while many periodic layouts were observed, the p-value of any of them did not meet the threshold set by the evaluation criteria. In fact, applying the periodicity evaluation criteria to previous works, suggests that none of the previously observed periods would be considered significant[290,291,312,313,356], except for the ones from Junier et al[287]. In addition, other methods have been applied to periodicity analysis of gene position as well[357–359]. It would be interesting to see if the interpretation of the results of those papers would change when tested using the synthetic control dataset approach. Regardless, the results from this work suggest more caution in the interpretation of periodic patterns.

On the other hand, a non-significant p-value does not rule out that the periodic pattern is relevant.

Query biological datasets might just be too noisy to give statistically significant results, while still providing the correct period. This could be revealed by analyzing different kinds of datasets and demonstrating that they point to the same conclusion. For instance, common periodic layouts between 3 different kinds of ChIP experiments for each chromosome arm were found in yeast[291]. In one case, the periodicities found on the different chromosome arms even corresponded well with the periodicity observed for the distribution of replication origins[291]. These results suggest that the lack of significant periods observed in the ChIP-exo datasets may be due to a lack of accuracy of the analysis program, rather than the absence of a periodic signal. It therefore suggests that additional evidence is needed to support these periods.

Additional support came from two sources in this work: (i) comparison of the ChIP-exo results with periodicity results from datasets obtained from RegulonDB and (ii) harmonic analysis to find common periodic organizations between different stress response organizations.

The RegulonDB datasets contain manually curated binding sites of the stress response regulators that is the collection of all known literature on this topic. This means that these datasets contain both strongly and weakly supported binding site, hence data-cleaning is required. Since no objective criterion exists to clean the datasets, one published method[360] was used as well as 2 other ones that seemed relevant to the topic. These results showed that most periods could not be found again in other datasets. However, the 40,5 kb SoxS period and the 8,6 kb Fur period were found again, giving them some extra weight.

Harmonic analysis further strengthened this point by finding a connection between the ChIP-exo periods of OxyR-SoxS-Fur regulons. Interestingly, these common periodic organizations reflected observed *in vivo* crosstalk between these stress response systems[350].

In conclusion, the results presented in this work suggest more caution is needed when interpreting periodic signals and provides a blueprint for future work in the field.

4.2 The relation between gene position, chromosome structure and gene expression

In order to further support the periodicity data from the previous section, a transcription factor was displaced to observe the effect on gene expression of its targets. The OxyR-SoxS-Fur regulon was considered for further analysis based on the periodicity results. The OxyR regulon was chosen as test regulon, because (i) OxyR has the ability to oligomerize, which is a necessary condition for colocalization[346] (ii) OxyR mutants are hypersensitive to oxidative stress[351], maximally impacting oxidative stress response under experimental conditions and (iii) OxyR controls a relatively small regulon dedicated to oxidative stress response and will minimally perturb cellular metabolism under non-experimental conditions[323,350].

Next, *oxyRS* was moved to six locations in the chromosome in order to assess the effect of gene position on the stress response. As an initial experiment, three locations near the origin were chosen

and three near the terminus to test the maximal effect of *oxyRS* repositioning on the stress response before testing periodicity [352–354].

The oxidative stress response of the *oxyRS* movants was tested using the Kirby Bauer test[325]. While *ΔoxyRS* showed drastically increased sensitivity to oxidative stress, the *oxyRS* movants were equally sensitive to it as wild type. These results suggest that the position of a transcriptional regulator does not impact gene regulation of its targets.

In this work, negative autoregulation of OxyR expression might have played a role in buffering the effects of gene translocation[345]. Nevertheless, 5 out of 6 insertion locations were out of phase with the OxyR targets, which should have disturbed gene regulation according to the TBS framework. Moreover, while gene displacement has been shown to impact the expression level of a moved gene in bacteria[286,353,354,361–370], displacement of a transcription factor gene does not always measurably affect gene regulation[354,366] (see [371] for counterexample).

Conflicting results on the role of gene colocalization in gene expression in bacteria(another key prediction of the TBS model) have emerged in recent years. Recent results from single molecule imaging partially support the TBS model[372]. The NAP H-NS was shown to form dense foci that colocalized with their regulated genes in *E.coli*[373], despite no observed effect of its displacement[366]. The transcription factor GalR also formed dense foci in stationary phase[374]. Additionally, 3C studies detected physical interactions between distant members of the GalR regulon that were abolished in a strain deleted for GalR[374]. Also, dimerization mutants of GalR and H-NS did not cluster anymore and yielded reduced gene regulation[373,374]. By contrast, other NAP's such as Fis and HU are uniformly distributed on the *E. coli* chromosome[373] and other transcription factors like LacI cluster near their target genes, but do not come together in space[375]. Finally, genes from the arabinose operon do not change cellular localization upon induction of gene expression[376]. These results suggest that oligomerization of the transcription factor is a necessary, but not sufficient condition for spatial clustering and an effect on gene expression[372].

Moreover, the more recent studies seem to suggest that chromosome structure and its effect on gene expression in bacteria is shaped by only a handful of global regulators. In a very recent study, the effect of gene position on transcriptional propensity (RNA expression/DNA) in *E. coli* was studied for the first time at a very large scale (144000 insertion sites)[370]. These results revealed that locations of high propensity corresponded to Fis and HU binding sites, two global regulators[370]. Interestingly, Fis is the only known transcriptional regulator for which target transcription is affected when its gene is displaced[371]. Additionally, deletion of the *hu* genes impacts the location-to-location variability in gene expression when a gene is displaced[367].

In conclusion, the results of this work suggest that gene position of *oxyRS* plays no major role in the regulation of its targets. This is in line with conflicting results on the importance of gene position for gene regulation. Moreover, recent studies suggest that only a handful, abundantly expressed proteins shape chromosome architecture in *E. coli*.

4.3 Conclusion and future directions

In the end, this work was unable to demonstrate an effect of gene periodicity on oxidative stress response in *E. coli*. Periods of key stress response regulators were deemed insignificant using our evaluation criteria and displacement of the *oxyRS* stress response regulator failed to produce significant increases in stress sensitivity, consistent with some recent results in literature. Given these results and the untimely retirement of the supervisor that had started this project, the project was stopped in accordance with the thesis committee.

Future work on a possible role of gene periodicity in its regulation would benefit from several aspects of the approach taken in this work, namely (i) establish periodicity evaluation criteria using control datasets (ii) support periodicity results through the analysis of multiple datasets per regulon and through cross-referencing common periodic themes between regulons with known regulatory cross-talk and finally (iii) direct testing of the hypothesis by displacement of a reporter in and out of phase. Additionally, the experiment would best be done in eukaryotes (e.g. yeast). As stated above, periodicity results in yeast, while not significant on their own, are supported by the identification of common periodic themes in strongly different datasets[291]. Additionally, gene position effects are generally stronger in eukaryotes and their existence has not been controversial for many years[377–380]. Moreover, the existence of transcription neighborhoods (colocalization of active cofunctional genes inside a transcription focus) is not controversial in eukaryotes either (for numerous examples: see [377]). Therefore, a similar study to the one presented here in eukaryotes might shed an interesting new light on the three-way relationship between gene position, 3D structure and gene expression in eukaryotes.

1. Vermeulen, K., Van Bockstaele, D. R. & Berneman, Z. N. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif.* **36**, 131–149 (2003).
2. Lodish, H., Berk, A. & Zipursky, S. Overview of the Cell cycle and its control. in *Molecular Cell Biology* vol. 13.1 (W.H.Freeman, 2000).
3. Brehm, A. *et al.* Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597–601 (1998).
4. Voitenleitner, C., Fanning, E. & Nasheuer, H.-P. Phosphorylation of DNA polymerase a-primase by Cyclin A-dependent kinases regulates initiation of DNA replication in vitro. *Oncogene* **14**, 1611–1615 (1997).
5. Siliciano, J. D. *et al.* DNA damage induces phosphorylation of the amino terminus of p53. *Genes Dev.* **11**, 3471–3481 (1997).
6. Adams, D. W. & Errington, J. Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat. Rev. Microbiol.* **7**, 642–653 (2009).
7. Badrinarayanan, A., Le, T. B. K. & Laub, M. T. Bacterial Chromosome Organization and Segregation. *Annu. Rev. Cell Dev. Biol.* **31**, 171–199 (2015).
8. Reyes-Lamothe, R. & Sherratt, D. J. The bacterial cell cycle, chromosome inheritance and cell growth. *Nat. Rev. Microbiol.* **17**, 467–478 (2019).
9. Robert, L. Size sensors in bacteria, cell cycle control, and size control. *Front. Microbiol.* **6**, (2015).
10. Boye, E. & Nordström, K. Coupling the cell cycle to cell growth: A look at the parameters that regulate cell-cycle events. *EMBO Rep.* **4**, 757–760 (2003).
11. Kreuzer, K. N. DNA Damage Responses in Prokaryotes: Regulating Gene Expression, Modulating Growth Patterns, and Manipulating Replication Forks. *Cold Spring Harb. Perspect. Biol.* **5**, a012674–a012674 (2013).
12. Lindås, A.-C. & Bernander, R. The cell cycle of archaea. *Nat. Rev. Microbiol.* **11**, 627–638

(2013).

13. Grogan, D. W., Bernander, R. & Poplawski, A. Altered patterns of cellular growth, morphology, replication and division in conditional-lethal mutants of the thermophilic archaeon Sulfolobus acidocaldarius. *Microbiology* **146**, 749–757 (2000).
14. Hjort, K. & Bernander, R. Cell cycle regulation in the hyperthermophilic crenarchaeon Sulfolobus acidocaldarius: S. acidocaldarius cell cycle. *Mol. Microbiol.* **40**, 225–234 (2001).
15. Gabrielli, B., Brooks, K. & Pavey, S. Defective Cell Cycle Checkpoints as Targets for Anti-Cancer Therapies. *Front. Pharmacol.* **3**, (2012).
16. Jameson, K. & Wilkinson, A. Control of Initiation of DNA Replication in Bacillus subtilis and Escherichia coli. *Genes* **8**, 22 (2017).
17. Fuller, R. S., Funnell, B. E. & Kornberg, A. The dnaA Protein Complex with the E. coli Chromosomal Replication Origin (OK) and Other DNA Sites. *Cell* **38**, 889–900 (1984).
18. Zorman, S., Seitz, H., Sclavi, B. & Strick, T. R. Topological characterization of the DnaA–oriC complex using single-molecule nanomanipuation. *Nucleic Acids Res.* **40**, 7375–7383 (2012).
19. Erzberger, J. P., Mott, M. L. & Berger, J. M. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* **13**, 676–683 (2006).
20. Kowalski, D. & Eddy, M. J. The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. *EMBO J.* **8**, 4335–4344 (1989).
21. Bramhill, D. & Kornberg, A. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome. *Cell* **52**, 743–755 (1988).
22. Bruand, C., Ehrlich, S. D. & Jannière, L. Primosome assembly site in Bacillus subtilis. *EMBO J.* **14**, 2642–2650 (1995).
23. Zhang, W. *et al.* The Bacillus subtilis DnaD and DnaB Proteins Exhibit Different DNA Remodelling Activities. *J. Mol. Biol.* **351**, 66–75 (2005).
24. Smits, W. K., Goranov, A. I. & Grossman, A. D. Ordered association of helicase loader

proteins with the *Bacillus subtilis* origin of replication *in vivo*. *Mol. Microbiol.* **75**, 452–461 (2010).

25. Fang, L., Davey, M. J. & O'Donnell, M. Replisome Assembly at oriC, the Replication Origin of E. coli, Reveals an Explanation for Initiation Sites outside an Origin. *Mol. Cell* **4**, 541–553 (1999).
26. Reyes-Lamothe, R., Sherratt, D. J. & Leake, M. C. Stoichiometry and Architecture of Active DNA Replication Machinery in Escherichia coli. *Science* **328**, 498–501 (2010).
27. Robinson, A., J. Causer, R. & E. Dixon, N. Architecture and Conservation of the Bacterial DNA Replication Machinery, an Underexploited Drug Target. *Curr. Drug Targets* **13**, 352–372 (2012).
28. Beattie, T. R. & Reyes-Lamothe, R. A Replisomeâ€TMs journey through the bacterial chromosome. *Front. Microbiol.* **6**, (2015).
29. Voet, D. & Voet, J. DNA Replication, Repair, and Recombination. in *Biochemistry* 1171–1259 (Wiley, 2011).
30. Corn, J. E., Pelton, J. G. & Berger, J. M. Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. *Nat. Struct. Mol. Biol.* **15**, 163–169 (2008).
31. Corn, J. E. & Berger, J. M. Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions. *Nucleic Acids Res.* **34**, 4082–4088 (2006).
32. Sanders, G. M., Dallmann, H. G. & McHenry, C. S. Reconstitution of the B. subtilis Replisome with 13 Proteins Including Two Distinct Replicases. *Mol. Cell* **37**, 273–281 (2010).
33. Hill, T. M. Arrest of Bacterial DNA Replication. *Annu. Rev. Microbiol.* **46**, 603–633 (1992).
34. Lewis, P. J., Ralston, G. B., Christopherson, R. I. & Wake, R. G. Identification of the Replication Terminator .Protein Binding Sites in the Terminus Region of the Bacillus subtiks Chromosome and Stoichiometry of the Binding. *J. Mol. Biol.* **214**, 73–84 (1990).
35. Vivian, J. P., Porter, C. J., Wilce, J. A. & Wilce, M. C. J. An Asymmetric Structure of the Bacillus subtilis Replication Terminator Protein in Complex with DNA. *J. Mol. Biol.* **370**,

481–491 (2007).

36. Langley, D. B., Smith, M. T., Lewis, P. J. & Wake, R. G. Protein?nucleoside contacts in the interaction between the replication terminator protein of Bacillus subtilis and the DNA terminator. *Mol. Microbiol.* **10**, 771–779 (1993).
37. Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H. & Ogasawara, N. Cloning of an autonomously replicating sequence (*ars*) from the *Bacillus subtilis* chromosome. *Mol. Microbiol.* **6**, 309–315 (1992).
38. Dervyn, E. Two Essential DNA Polymerases at the Bacterial Replication Fork. *Science* **294**, 1716–1719 (2001).
39. Hidaka, M., Kobayashi, T., Takenaka, S., Takeya, H. & Horiuchi, T. Purification of a DNA Replication Terminus (ter) Site-binding Protein in Escherichia coli and Identification of the Structural Gene. *J. Biol. Chem.* **267**, 21031–21037 (1989).
40. Hill, T. M., Henson, J. M. & Kuempel, P. L. The terminus region of the Escherichia coli chromosome contains two separate loci that exhibit polar inhibition of replication. *Proc. Natl. Acad. Sci.* **84**, 1754–1758 (1987).
41. Neylon, C., Kralicek, A. V., Hill, T. M. & Dixon, N. E. Replication Termination in Escherichia coli: Structure and Antihelicase Activity of the Tus-Ter Complex. *Microbiol. Mol. Biol. Rev.* **69**, 501–526 (2005).
42. Briggs, G. S., Smits, W. K. & Soultanas, P. Chromosomal Replication Initiation Machinery of Low-G+C-Content Firmicutes. *J. Bacteriol.* **194**, 5162–5170 (2012).
43. Hayashi, M., Ogura, Y., Harry, E. J., Ogasawara, N. & Moriya, S. *Bacillus subtilis* YabA is involved in determining the timing and synchrony of replication initiation. *FEMS Microbiol. Lett.* **247**, 73–79 (2005).
44. Felicori, L. *et al.* Tetramerization and interdomain flexibility of the replication initiation controller YabA enables simultaneous binding to multiple partners. *Nucleic Acids Res.* **44**, 449–463 (2016).

45. Cho, E., Ogasawara, N. & Ishikawa, S. The functional analysis of YabA, which interacts with DnaA and regulates initiation of chromosome replication in Bacillus subtils. *Genes Genet. Syst.* **83**, 111–125 (2008).

46. Merrikh, H. & Grossman, A. D. Control of the replication initiator DnaA by an anti-cooperativity factor: Control of DnaA by an anti-cooperativity factor. *Mol. Microbiol.* **82**, 434–446 (2011).

47. Scholefield, G., Errington, J. & Murray, H. Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA: Soj inhibits DnaA helix formation. *EMBO J.* **31**, 1542–1555 (2012).

48. Scholefield, G., Whiting, R., Errington, J. & Murray, H. Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation: Oligomeric status switches Soj regulatory activity. *Mol. Microbiol.* **79**, 1089–1100 (2011).

49. Murray, H. & Errington, J. Dynamic Control of the DNA Replication Initiation Protein DnaA by Soj/ParA. *Cell* **135**, 74–84 (2008).

50. Leonard, T. A., Butler, P. J. & L**?**we, J. Bacterial chromosome segregation: structure and DNA binding of the Soj dimer ? a conserved biological switch. *EMBO J.* **24**, 270–282 (2005).

51. Okumura, H. *et al.* Regulation of chromosomal replication initiation by oriC-proximal DnaA-box clusters in Bacillus subtilis. *Nucleic Acids Res.* **40**, 220–234 (2012).

52. Ishida, T. *et al.* DiaA, a Novel DnaA-binding Protein, Ensures the Timely Initiation of *Escherichia coli* Chromosome Replication. *J. Biol. Chem.* **279**, 45546–45555 (2004).

53. Terradot, L. & Zawilak-Pawlik, A. Structural insight into *Helicobacter pylori* DNA replication initiation. *Gut Microbes* **1**, 330–334 (2010).

54. Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E. & Leonard, A. C. Escherichia coli prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA: Regulation of bacterial pre-RC assembly. *Mol. Microbiol.* **51**, 1347–1359 (2004).

55. Cassler, M. R., Grimwade, J. E. & Leonard, A. C. Cell cycle-specific changes in nucleoprotein

complexes at a chromosomal replication origin. *EMBO J.* **14**, 5833–5841 (1995).

56. Grimwade, J. E., Ryan, V. T. & Leonard, A. C. IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of Escherichia coli. *Mol. Microbiol.* **35**, 835–844 (2000).
57. Rasmussen, K. V. & Schaechter, M. SeqA limits DnaA activity in replication from oriC in Escherichia coli. *Mol. Microbiol.* **14**, 763–772 (1994).
58. Lu, M. SeqA: A negative modulator of replication initiation in E. coli. *Cell* **77**, 413–426 (1994).
59. Campbell, J. L. & Kleckner, N. E. coli oriC and the dnaA Gene Promoter Are Sequestered from dam Methyltransfferase Following the Passage of the Chromosomal Replication Fork. 13.
60. Han, J. S., Kang, S., Kim, S. H., Ko, M. J. & Hwang, D. S. Binding of SeqA Protein to Hemi-methylated GATC Sequences Enhances Their Interaction and Aggregation Properties. *J. Biol. Chem.* **279**, 30236–30243 (2004).
61. Waldminghaus, T. & Skarstad, K. The Escherichia coli SeqA protein. *Plasmid* **61**, 141–150 (2009).
62. Kurokawa, K. *et al.* Rapid Exchange of Bound ADP on the *Staphylococcus aureus* Replication Initiation Protein DnaA. *J. Biol. Chem.* **284**, 34201–34210 (2009).
63. Su'etsugu, M., Nakamura, K., Keyamura, K., Kudo, Y. & Katayama, T. Hda Monomerization by ADP Binding Promotes Replicase Clamp-mediated DnaA-ATP Hydrolysis. *J. Biol. Chem.* **283**, 36118–36131 (2008).
64. Su'etsugu, M. *et al.* The DnaA N-terminal domain interacts with Hda to facilitate replicase clamp-mediated inactivation of DnaA: DnaA domain I binding to Hda. *Environ. Microbiol.* **15**, 3183–3195 (2013).
65. Nozaki, S., Yamada, Y. & Ogawa, T. Initiator titration complex formed at *datA* with the aid of IHF regulates replication timing in *Escherichia coli*. *Genes Cells* **14**, 329–341 (2009).
66. Kasho, K. & Katayama, T. DnaA binding locus datA promotes DnaA-ATP hydrolysis to

enable cell cycle-coordinated replication initiation. *Proc. Natl. Acad. Sci.* **110**, 936–941 (2013).

67. Fujimitsu, K., Senriuchi, T. & Katayama, T. Specific genomic sequences of E. coli promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev.* **23**, 1221–1233 (2009).
68. Kasho, K., Fujimitsu, K., Matoba, T., Oshima, T. & Katayama, T. Timely binding of IHF and Fis to DARS2 regulates ATP–DnaA production and replication initiation. *Nucleic Acids Res.* **42**, 13134–13149 (2014).
69. Zhang, Q. *et al.* Reversible lysine acetylation is involved in DNA replication initiation by regulating activities of initiator DnaA in Escherichia coli. *Sci. Rep.* **6**, (2016).
70. Wang, J. D. & Levin, P. A. Metabolism, cell growth and the bacterial cell cycle. *Nat. Rev. Microbiol.* **7**, 822–827 (2009).
71. Cooper, S. & Helmstetter, C. Chromosome Replication and the Division Cycle of Escherichia coli B/r. *J Mol Biol* 619–644 (1967).
72. Michelsen, O. Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. *Microbiology* **149**, 1001–1010 (2003).
73. Yoshikawa, H., O'Sullivan, A. & Sueoka, N. SEQUENTIAL REPLICATION OF THE BACILLUS SUBTILIS CHROMOSOME, III. REGULATION OF INITIATION. *Genetics* 973–980 (1964).
74. Donachie, W. D. Relationship between Cell size and Time of Initiation of DNA replication. *Nature* **219**, 1077–1079 (1968).
75. Schaechter, M., MaalOe, O. & Kjeldgaard, N. O. Dependency on Medium and Temperature of Cel Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *J.gen.Microbiol.* **19**, 592–606 (1958).
76. Campos, M. *et al.* A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell* **159**, 1433–1446 (2014).
77. Taheri-Araghi, S. *et al.* Cell-Size Control and Homeostasis in Bacteria. *Curr. Biol.* **25**, 385–

391 (2015).

78. Wold, S., Skarstad, K., Steen, H. B., Stokke, T. & Boye, E. The initiation mass for DNA replication in Escherichia coli K-12 is dependent on growth rate. *EMBO J.* **13**, 2097–2102 (1994).

79. Boye, E., Stokke, T., Kleckner, N. & Skarstad, K. Coordinating DNA replication initiation with cell growth: differential roles for DnaA and SeqA proteins. *Proc. Natl. Acad. Sci.* **93**, 12206–12211 (1996).

80. Bates, D. & Kleckner, N. Chromosome and Replisome Dynamics in E. coli: Loss of Sister Cohesion Triggers Global Chromosome Movement and Mediates Chromosome Segregation. *Cell* **121**, 899–911 (2005).

81. Chien, A.-C., Hill, N. S. & Levin, P. A. Cell Size Control in Bacteria. *Curr. Biol.* **22**, R340–R349 (2012).

82. Hill, N. S., Kadoya, R., Chattoraj, D. K. & Levin, P. A. Cell Size and the Initiation of DNA Replication in Bacteria. *PLoS Genet.* **8**, e1002549 (2012).

83. Weart, R. B. *et al.* A Metabolic Sensor Governing Cell Size in Bacteria. *Cell* **130**, 335–347 (2007).

84. Bipatnath, M., Dennis, P. P. & Bremer, H. Initiation and Velocity of Chromosome Replication in Escherichia coli B/r and K-12. *J BACTERIOL* **180**, 9 (1998).

85. Churchward, G. Growth Rate-Dependent Control of Chromosome Replication Initiation in Escherichia coli. 7.

86. Nouri, H. *et al.* Multiple links connect central carbon metabolism to DNA replication initiation and elongation in *Bacillus subtilis*. *DNA Res.* **25**, 641–653 (2018).

87. Barańska, S. *et al.* Replicating DNA by cell factories: roles of central carbon metabolism and transcription in the control of DNA replication in microbes, and implications for understanding this process in human cells. *Microb. Cell Factories* **12**, 55 (2013).

88. Zyskind, J. W. & Smith, D. W. DNA replication, the bacterial cell cycle, and cell growth. *Cell*

69, 5–8 (1992).

89. Chance, B., Estabrook, R. & Ghosh, A. DAMPED SINUSOIDAL OSCILLATIONS OF CYTOPLASMIC REDUCED PYRIDINE NUCLEOTIDE IN YEAST CELLS. *Proc Natl Acad Sci U A* **51**, 1244–1251 (1964).
90. Klevecz, R. R., Bolen, J., Forrest, G. & Murray, D. B. A genomewide oscillation in transcription gates DNA replication and cell cycle. *Proc. Natl. Acad. Sci.* **101**, 1200–1205 (2004).
91. Tu, B. P. Logic of the Yeast Metabolic Cycle: Temporal Compartmentalization of Cellular Processes. *Science* **310**, 1152–1158 (2005).
92. Burnetti, A. J., Aydin, M. & Buchler, N. E. Cell cycle Start is coupled to entry into the yeast metabolic cycle across diverse strains and growth rates. *Mol. Biol. Cell* **27**, 64–74 (2016).
93. Chiaramello, A. E. & Zyskind, J. W. Expression of Escherichia coli dnaA and mioC genes as a function of growth rate. *J. Bacteriol.* **171**, 4272–4280 (1989).
94. Hanawalt, P. C., Maaløe, O., Cummings, D. J. & Schaechter, M. The normal DNA replication cycle. II. *J. Mol. Biol.* **3**, 156–165 (1961).
95. Ogura, Y., Imai, Y., Ogasawara, N. & Moriya, S. Autoregulation of the dnaA-dnaN Operon and Effects of DnaA Protein Levels on Replication Initiation in Bacillus subtilis. *J. Bacteriol.* **183**, 3833–3841 (2001).
96. Skarstad, K., Lobner-Olesen, A., Atlung, T., von Meyenburg, K. & Boye, E. Initiation of DNA replication in Escherichia coil after overproduction of the DnaA protein. *Mol Gen Genet* 50–56 (1989).
97. Xu, J. C. & Bremer, H. Chromosome replication in Escherichia coli induced by oversupply of DnaA. *Mol Gen Genet* 138–142 (1988).
98. Schaus, N., O'Day, K., Peters, W. & Wright, A. Isolation and Characterization of Amber Mutations in Gene dnaA of Escherichia coli K-12. *J BACTERIOL* **145**, 10 (1981).
99. Barker, M. M., Gaal, T., Josaitis, C. A. & Gourse, R. L. Mechanism of regulation of

transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* **305**, 673–688 (2001).

100. Cashel, M., Gentry, D. R., Hernandez, V. H. & Vinella, D. The Stringent Response. In: Escherichia coli and Salmonella. in *Cellular and Molecular Biology* (ASM, 1996).
101. Chiaramello, A. E. & Zyskind, J. W. Coupling of DNA replication to growth rate in Escherichia coli: a possible role for guanosine tetraphosphate. *J. Bacteriol.* **172**, 2013–2019 (1990).
102. Samadpour, A. N. & Merrikh, H. DNA gyrase activity regulates DnaA-dependent replication initiation in *Bacillus subtilis*: DNA gyrase regulates DnaA-dependent initiation. *Mol. Microbiol.* **108**, 115–127 (2018).
103. Reece, R. J. & Maxwell, A. DNA Gyrase: Structure and Function. *Crit. Rev. Biochem. Mol. Biol.* **26**, 335–375 (1991).
104. Churchward, G. & Bremer, H. Determination of Deoxyribonucleic Acid Replication Time in Exponentially Growing Escherichia coli B/r. **130**, 8 (1977).
105. Herrick, J. & Sclavi, B. Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. *Mol. Microbiol.* **63**, 22–34 (2007).
106. Torrents, E. *et al.* NrdR Controls Differential Expression of the Escherichia coli Ribonucleotide Reductase Genes. *J. Bacteriol.* **189**, 5012–5021 (2007).
107. McKethan, B. L. & Spiro, S. Cooperative and allosterically controlled nucleotide binding regulates the DNA binding activity of NrdR: Characterization of NrdR in *Escherichia coli*. *Mol. Microbiol.* n/a-n/a (2013) doi:10.1111/mmi.12364.
108. DeNapoli, J., Tehranchi, A. K. & Wang, J. D. Dose-dependent reduction of replication elongation rate by (p)ppGpp in *Escherichia coli* and *Bacillus subtilis*: (p)ppGpp modulates replication elongation rates. *Mol. Microbiol.* **88**, 93–104 (2013).
109. Wang, J. D., Sanders, G. M. & Grossman, A. D. Nutritional Control of Elongation of DNA Replication by (p)ppGpp. *Cell* **128**, 865–875 (2007).

110. Murray, H. & Koh, A. Multiple Regulatory Systems Coordinate DNA Replication with Cell Growth in Bacillus subtilis. *PLoS Genet.* **10**, e1004731 (2014).

111. Flåtten, I., Fossum-Raunehaug, S., Taipale, R., Martinsen, S. & Skarstad, K. The DnaA Protein Is Not the Limiting Factor for Initiation of Replication in Escherichia coli. *PLOS Genet.* **11**, e1005276 (2015).

112. Hernandez, J. & Bremer, H. Characterization of RNA and DNA Synthesis in Escherichia coli Strains Devoid of ppGpp. 12.

113. Saxena, R., Fingland, N., Patil, D., Sharma, A. & Crooke, E. Crosstalk between DnaA Protein, the Initiator of Escherichia coli Chromosomal Replication, and Acidic Phospholipids Present in Bacterial Membranes. *Int. J. Mol. Sci.* **14**, 8517–8537 (2013).

114. SekimizuS, K. & Kornberg, A. Cardiolipin Activation of dnaA Protein, the Initiation Protein of Replication in Escherichia coZi. *J.Biol.Chem* 7131–7135 (1988).

115. Yung, B. Y. & Kornberg, A. Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in Escherichia coli. *Proc. Natl. Acad. Sci.* **85**, 7202–7205 (1988).

116. CastumaS, C. E., Crooke, E. & Kornberg, A. Fluid Membranes with Acidic Domains Activate DnaA, the Initiator Protein of Replication in Escherichia coZi* (. *J. Biol. Chem.* **268**, 24665–24668 (1993).

117. Makise, M., Mima, S., Katsu, T., Tsuchiya, T. & Mizushima, T. Acidic phospholipids inhibit the DNA-binding activity of DnaA protein, the initiator of chromosomal DNA replication in Escherichia coli: Inhibition of DnaA binding to oriC by phospholipids. *Mol. Microbiol.* **46**, 245–256 (2002).

118. Heacock, P. N. & Dowhan, W. Alteration of the Phospholipid Composition of Escherichia coli through Genetic Manipulation. *J. Biol. Chem.* **264**, 14972–14977 (1989).

119. Heacock, P. N. & Dowhan, W. Construction of a Lethal Mutation in the Synthesis of the Major Acidic Phospholipids of Escherichia coli*. *J. Biol. Chem.* **262**, 13044–13049 (1987).

120. Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA Induces Cell Growth and Proliferation by Promoting the Acetylation of Histones at Growth Genes. *Mol. Cell* **42**, 426–437 (2011).

121. Huang, T.-S. & Nagy, P. D. Direct Inhibition of Tombusvirus Plus-Strand RNA Synthesis by a Dominant Negative Mutant of a Host Metabolic Enzyme, Glyceraldehyde-3-Phosphate Dehydrogenase, in Yeast and Plants. *J. Virol.* **85**, 9090–9102 (2011).

122. Prasanth, K. R. *et al.* Glyceraldehyde 3-Phosphate Dehydrogenase Negatively Regulates the Replication of Bamboo Mosaic Virus and Its Associated Satellite RNA. *J. Virol.* **85**, 8829–8840 (2011).

123. Kim, B. H. & Gadd, G. M. *Bacterial Physiology and Metabolism*. (Cambridge University Press., 2008).

124. Lunt, S. Y. & Vander Heiden, M. G. Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. *Annu. Rev. Cell Dev. Biol.* **27**, 441–464 (2011).

125. Brasen, C., Esser, D., Rauch, B. & Siebers, B. Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation. *Microbiol. Mol. Biol. Rev.* **78**, 89–175 (2014).

126. Eymann, C. *et al.* A comprehensive proteome map of growingBacillus subtilis cells. *PROTEOMICS* **4**, 2849–2876 (2004).

127. Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci.* **110**, 10039–10044 (2013).

128. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**, 613–624 (2008).

129. Deutscher, J. *et al.* The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiol. Mol. Biol. Rev.* **78**, 231–256 (2014).

130. Alpert, C. A., Frank, R., Stueber, K., Deutscher, J. & Hengstenberg, W. Phosphoenolpyruvate-

dependent protein kinase enzyme I of Streptococcus faecalis: purification and properties of the enzyme and characterization of its active center. *Biochemistry* **24**, 959–964 (1985).

131. Gassner, M. *et al.* The Phosphoenolpyruvate-Dependent Phosphotransferase System of Staphylococcus aureus. 2. 1H and 31P Nuclear-Magnetic-Resonance Studies on the Phosphocarrier Protein HPr, Phosphohistidines and Phosphorylated HPr. *Eur. J. Biochem.* **75**, 287–296 (1977).
132. Dorschug, M., Frank, R., Kalbitzer, H. R., Hengstenberg, W. & Deutscher, J. Phosphoenolpyruvate-dependent phosphorylation site in enzyme IIIglc of the Escherichia coli phosphotransferase system. *Eur. J. Biochem.* **144**, 113–119 (1984).
133. Deutscher, J., Beyreuther, K., Sobek, H. M. & Stuber, K. Phosphoenolpyruvate-Dependent Phosphotransferase System of Staphylococcus aureus: Factor IIIlaC, a Trimeric Phospho-Carrier Protein That Also Acts as a Phase Transfer Catalystt. *Biochemistry* **21**, 4867–4873 (1982).
134. Pas, H. H. & Robillard, G. T. S-Phosphocysteine and Phosphohistidine Are Intermediates in the Phosphoenolpyruvate-Dependent Mannitol Transport Catalyzed by Escherichia coli HIMt1. *Biochemistry* **27**, 5835–5839 (1988).
135. Cao, Y. *et al.* Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature* **473**, 50–54 (2011).
136. Bettenbrock, K. *et al.* A Quantitative Approach to Catabolite Repression in *Escherichia coli*. *J. Biol. Chem.* **281**, 2578–2584 (2006).
137. Hogema, B. M. *et al.* Inducer exclusion in Escherichia coli by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Mol. Microbiol.* **30**, 487–498 (1998).
138. Vadeboncoeur, C., Brochu, D. & Reizer, J. Quantitative Determination of the Intracellular Concentration of the Various Forms of HPr, a Phosphocarrier Protein of the Phosphoenolpyruvate: Sugar Phosphotransferase System in Growing Cells of Oral

Streptococci. *Anal. Biochem.* **196**, 24–30 (1991).

139. Warner, J. B. & Lolkema, J. S. CcpA-Dependent Carbon Catabolite Repression in Bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 475–490 (2003).
140. Nicholson, W. L. *et al.* Catabolite repression-resistant mutations of the Bacillus subtilis alpha-amylase promoter affect transcription levels and are in an operator-like sequence. *J. Mol. Biol.* **198**, 609–618 (1987).
141. Nessler, S. *et al.* HPr Kinase/Phosphorylase, the Sensor Enzyme of Catabolite Repression in Gram-Positive Bacteria: Structural Aspects of the Enzyme and the Complex with Its Protein Substrate. *J. Bacteriol.* **185**, 4003–4010 (2003).
142. Galinier, A. *et al.* New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proc. Natl. Acad. Sci.* **95**, 1823–1828 (1998).
143. Jault, J.-M. *et al.* The HPr Kinase from *Bacillus subtilis* Is a Homo-oligomeric Enzyme Which Exhibits Strong Positive Cooperativity for Nucleotide and Fructose 1,6-Bisphosphate Binding. *J. Biol. Chem.* **275**, 1773–1780 (2000).
144. Reizer, J. *et al.* A novel protein kinase that controls carbon catabolite repression in bacteria. *Mol. Microbiol.* **27**, 1157–1169 (1998).
145. Mijakovic, I. *et al.* Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *Proc. Natl. Acad. Sci.* **99**, 13442–13447 (2002).
146. Deutscher, J., Küster, E., Bergstedt, U., Charrier, V. & Hillen, W. Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in Gram-positive bacteria. *Mol. Microbiol.* **15**, 1049–1053 (1995).
147. Jones, B. E. *et al.* Binding of the Catabolite Repressor Protein CcpA to Its DNA Target Is Regulated by Phosphorylation of its Corepressor HPr. *J. Biol. Chem.* **272**, 26530–26535 (1997).
148. Seidel, G., Diel, M., Fuchsbauer, N. & Hillen, W. Quantitative interdependence of coeffectors, CcpA and cre in carbon catabolite regulation of Bacillus subtilis: Regulatory differences of

HPrSerP and CrhP. *FEBS J.* **272**, 2566–2577 (2005).

149. Schumacher, M. A., Seidel, G., Hillen, W. & Brennan, R. G. Structural Mechanism for the Fine-tuning of CcpA Function by The Small Molecule Effectors Glucose 6-Phosphate and Fructose 1,6-Bisphosphate. *J. Mol. Biol.* **368**, 1042–1050 (2007).
150. Djordjevic, G. M., Tchieu, J. H. & Saier, M. H. Genes Involved in Control of Galactose Uptake in Lactobacillus brevis and Reconstitution of the Regulatory System in Bacillus subtilis. *J. Bacteriol.* **183**, 3224–3236 (2001).
151. Poolman, B., Knol, J., Mollet, B., Nieuwenhuis, B. & Sulter, G. Regulation of bacterial sugar-H+ symport by phosphoenolpyruvate-dependent enzyme I/HPr-mediated phosphorylation. *Proc. Natl. Acad. Sci.* **92**, 778–782 (1995).
152. Gunnewijk, M. G. W. & Poolman, B. Phosphorylation State of HPr Determines the Level of Expression and the Extent of Phosphorylation of the Lactose Transport Protein of *Streptococcus thermophilus*. *J. Biol. Chem.* **275**, 34073–34079 (2000).
153. Stulke, J., Arnaud, M., Rapoport, G. & Martin-Verstraete, I. PRD ? a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol. Microbiol.* **28**, 865–874 (1998).
154. Martin-Verstraete, I. *et al.* Antagonistic effects of dual PTS-catalysed phosphorylation on the Bacillus subtilis transcriptional activator LevR. *Mol. Microbiol.* **28**, 293–303 (1998).
155. Tortosa, P. *et al.* Sites of positive and negative regulation in the Bacillus subtilis antiterminators LicT and SacY. *Mol. Microbiol.* **41**, 1381–1393 (2001).
156. Krüger, S., Gertz, S. & Hecker, M. Transcriptional analysis of bglPH expression in Bacillus subtilis: evidence for two distinct pathways mediating carbon catabolite repression. *J. Bacteriol.* **178**, 2637–2644 (1996).
157. Lindner, C., Galinier, A., Hecker, M. & Deutscher, J. Regulation of the activity of the Bacillus subtilis antiterminator LicT by multiple PEP-dependent, enzyme I- and HPr-catalysed phosphorylation. *Mol. Microbiol.* **31**, 995–1006 (1999).

158. Lindner, C., Hecker, M., Le Coq, D. & Deutscher, J. Bacillus subtilis Mutant LicT Antiterminators Exhibiting Enzyme I- and HPr-Independent Antitermination Affect Catabolite Repression of the bglPH Operon. *J. Bacteriol.* **184**, 4819–4828 (2002).

159. Bouraoui, H., Ventroux, M., Noirot-Gros, M.-F., Deutscher, J. & Joyet, P. Membrane sequestration by the EIIB domain of the mannitol permease MtlA activates the *Bacillus subtilis mtl* operon regulator MtlR: *B. subtilis* MtlR sequestration by the EIIB ^{Mtl} domain. *Mol. Microbiol.* **87**, 789–801 (2013).

160. Krin, E., Sismeiro, O., Danchin, A. & Bertin, P. N. The regulation of Enzyme IIAGlc expression controls adenylate cyclase activity in Escherichia coli. *Microbiology* **148**, 1553–1559 (2002).

161. Rothe, F. M., Wrede, C., Lehnik-Habrink, M., Gorke, B. & Stulke, J. Dynamic Localization of a Transcription Factor in Bacillus subtilis: the LicT Antiterminator Relocalizes in Response to Inducer Availability. *J. Bacteriol.* **195**, 2146–2154 (2013).

162. Jannière, L. *et al.* Genetic Evidence for a Link Between Glycolysis and DNA Replication. *PLoS ONE* **2**, e447 (2007).

163. Maciąg, M., Nowicki, D., Janniere, L., Szalewska-Pałasz, A. & Węgrzyn, G. Genetic response to metabolic fluctuations: correlation between central carbon metabolism and DNA replication in Escherichia coli. *Microb. Cell Factories* **10**, 19 (2011).

164. Maciąg-Dorszyńska, M., Ignatowska, M., Jannière, L., Węgrzyn, G. & Szalewska-Pałasz, A. Mutations in central carbon metabolism genes suppress defects in nucleoid position and cell division of replication mutants in Escherichia coli. *Gene* **503**, 31–35 (2012).

165. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci.* **96**, 2896–2901 (1999).

166. Chen, Y. & Tye, B. K. The yeast Mcm1 protein is regulated posttranscriptionally by the flux of glycolysis. *Mol. Cell. Biol.* **15**, 4631–4639 (1995).

167. Chang, V. K. *et al.* Mcm1 Binds Replication Origins. *J. Biol. Chem.* **278**, 6093–6100 (2003).

168. Chang, V. K., Donato, J. J., Chan, C. S. & Tye, B. K. Mcm1 Promotes Replication Initiation by Binding Specific Elements at Replication Origins. *Mol. Cell. Biol.* **24**, 6514–6524 (2004).

169. Sprague, G. F. Isolation and Characterization of a Saccharomyces cerevisiae Mutant Deficient in Pyruvate Kinase Activity. *J BACTERIOL* **130**, 10 (1977).

170. Dickinson, J. R. & Williams, A. S. The cdc30 Mutation in Saccharomyces cerevisiae Results in a Temperature-sensitive Isoenzyme of Phosphoglucose Isomerase. *Microbiology* **133**, 135–140 (1987).

171. Hartwell, L. H. GENETIC CONTROL OF THE CELL DIVISION CYCLE I N YEAST: *Genetics* **74**, 267–286 (1973).

172. Kaplan, Y. & Kupiec, M. A role for the yeast cell cycle/splicing factor Cdc40 in the G1/S transition. *Curr. Genet.* **51**, 123–140 (2007).

173. Shor, E. *et al.* The Origin Recognition Complex Interacts with a Subset of Metabolic Genes Tightly Linked to Origins of Replication. *PLoS Genet.* **5**, e1000755 (2009).

174. Konieczna, A., Szczepańska, A., Sawiuk, K., Węgrzyn, G. & Łyżeń, R. Effects of partial silencing of genes coding for enzymes involved in glycolysis and tricarboxylic acid cycle on the enterance of human fibroblasts to the S phase. *BMC Cell Biol.* **16**, (2015).

175. Fornalewicz, K., Wieczorek, A., Węgrzyn, G. & Łyżeń, R. Silencing of the pentose phosphate pathway genes influences DNA replication in human fibroblasts. *Gene* **635**, 33–38 (2017).

176. Buckland, R. J. *et al.* Increased and Imbalanced dNTP Pools Symmetrically Promote Both Leading and Lagging Strand Replication Infidelity. *PLoS Genet.* **10**, e1004846 (2014).

177. Tymecka-Mulik, J. *et al.* Suppression of the Escherichia coli dnaA46 mutation by changes in the activities of the pyruvate-acetate node links DNA replication regulation to central carbon metabolism. *PLOS ONE* **12**, e0176050 (2017).

178. Rannou, O. *et al.* Functional interplay of DnaE polymerase, DnaG primase and DnaC helicase within a ternary complex, and primase to polymerase hand-off during lagging strand DNA replication in Bacillus subtilis. *Nucleic Acids Res.* **41**, 5303–5320 (2013).

179. Paschalis, V. *et al.* Interactions of the *Bacillus subtilis* DnaE polymerase with replisomal proteins modulate its activity and fidelity. *Open Biol.* **7**, 170146 (2017).

180. Kilkenny, M. L. *et al.* The human CTF4-orthologue AND-1 interacts with DNA polymerase α/primase via its unique C-terminal HMG box. *Open Biol.* **7**, 170217 (2017).

181. Noirot-Gros, M.-F. *et al.* An expanded view of bacterial DNA replication. *Proc. Natl. Acad. Sci.* **99**, 8342–8347 (2002).

182. Stein, A. & Firshein, W. Probable Identification of a Membrane-Associated Repressor of Bacillus subtilis DNA Replication as the E2 Subunit of the Pyruvate Dehydrogenase Complex. *J. Bacteriol.* **182**, 2119–2124 (2000).

183. Ronai, Z. Glycolytic enzymes as DNA binding proteins. *Int. J. Biochem.* **25**, 1073–1076 (1993).

184. Boukouris, A. E., Zervopoulos, S. D. & Michelakis, E. D. Metabolic Enzymes Moonlighting in the Nucleus: Metabolic Regulation of Gene Transcription. *Trends Biochem. Sci.* **41**, 712–730 (2016).

185. Baxi, M. D. & Vishwanatha, J. K. Uracil DNA-glycosylase/glyceraldehyde-3-phosphate dehydrogenase is an Ap4A binding protein. *Biochemistry* **34**, 9700–9707 (1995).

186. Grosse, F., Nasheuer, H.-P., Scholtissek, S. & Schomburg, U. Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-alpha-primase complex. *Eur. J. Biochem.* **160**, 459–467 (1986).

187. Popanda, O., Fox, G. & Thielmann, H. W. Modulation of DNA polymerases a , d and ´ by lactate dehydrogenase and 3-phosphoglycerate kinase. *Biochim. Biophys. Acta* **1397**, 102–117 (1998).

188. Jindal, H. K. & Vishwanatha, K. Functional Identity of a Primer Recognition Protein as Phosphoglycerate Kinase. *J. Biol. Chem.* **265**, 6540–6543 (1990).

189. Kumble, K. D., Iversen, P. L. & Vishwanatha, J. K. The role of primer recognition proteins in DNA replication: inhibition of cellular proliferation by antisense oligodeoxyribonucleotides. *J.*

Cell Sci. **101**, 35–41 (1992).

190. Li, X. *et al.* Nuclear PGK1 Alleviates ADP-Dependent Inhibition of CDC7 to Promote DNA Replication. *Mol. Cell* **72**, 650-660.e8 (2018).
191. Tian, M. *et al.* Salicylic Acid Inhibits the Replication of *Tomato bushy stunt virus* by Directly Targeting a Host Component in the Replication Complex. *Mol. Plant. Microbe Interact.* **28**, 379–386 (2015).
192. Prasanth, K. R., Chuang, C. & Nagy, P. D. Co-opting ATP-generating glycolytic enzyme PGK1 phosphoglycerate kinase facilitates the assembly of viral replicase complexes. *PLOS Pathog.* **13**, e1006689 (2017).
193. van Noort, V. *et al.* Cross-talk between phosphorylation and lysine acetylation in a genome-reduced bacterium. *Mol. Syst. Biol.* **8**, (2012).
194. Hentchel, K. L. & Escalante-Semerena, J. C. Acylation of Biomolecules in Prokaryotes: a Widespread Strategy for the Control of Biological Function and Metabolic Stress. *Microbiol. Mol. Biol. Rev.* **79**, 321–346 (2015).
195. Garcia-Garcia, T. *et al.* Role of Protein Phosphorylation in the Regulation of Cell Cycle and DNA-Related Processes in Bacteria. *Front. Microbiol.* **7**, (2016).
196. Shi, L. *et al.* Protein-tyrosine phosphorylation interaction network in Bacillus subtilis reveals new substrates, kinase activators and kinase cross-talk. *Front. Microbiol.* **5**, (2014).
197. Petranovic, D. *et al.* Bacillus subtilis strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication: Bacillus subtilis ΔptkA deficient in DNA replication. *Mol. Microbiol.* **63**, 1797–1805 (2007).
198. Mijakovic, I. Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* **34**, 1588–1596 (2006).
199. Ma, Y., Kanakousaki, K. & Buttitta, L. How the cell cycle impacts chromatin architecture and influences cell fate. *Front. Genet.* **6**, (2015).
200. Shi, L. & Tu, B. P. Acetyl-CoA induces transcription of the key G1 cyclin CLN3 to promote

entry into the cell division cycle in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **110**, 7318–7323 (2013).

201. Zheng, L., Roeder, R. G. & Luo, Y. S Phase Activation of the Histone H2B Promoter by OCA-S, a Coactivator Complex that Contains GAPDH as a Key Component. *Cell* **114**, 255–266 (2003).
202. Sutendra, G. *et al.* A Nuclear Pyruvate Dehydrogenase Complex Is Important for the Generation of Acetyl-CoA and Histone Acetylation. *Cell* **158**, 84–97 (2014).
203. Yang, W. *et al.* PKM2 Phosphorylates Histone H3 and Promotes Gene Transcription and Tumorigenesis. *Cell* **150**, 685–696 (2012).
204. Sauer, U. & Eikmanns, B. J. The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria: We dedicate this paper to Rudolf K. Thauer, Director of the Max-Planck-Institute for Terrestrial Microbiology in Marburg, Germany, on the occasion of his 65th birthday. *FEMS Microbiol. Rev.* **29**, 765–794 (2005).
205. Servant, P., Le Coq, D. & Aymerich, S. CcpN (YqzB), a novel regulator for CcpA-independent catabolite repression of Bacillus subtilis gluconeogenic genes: Catabolite repression of gluconeogenesis genes by CcpN. *Mol. Microbiol.* **55**, 1435–1451 (2005).
206. Yoshida, K. -i. Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in Bacillus subtilis. *Nucleic Acids Res.* **29**, 683–692 (2001).
207. Blencke, H.-M. *et al.* Transcriptional profiling of gene expression in response to glucose in Bacillus subtilis: regulation of the central metabolic pathways. *Metab. Eng.* **5**, 133–149 (2003).
208. Diesterhaft, M. & Freese, E. Pyruvate kinase of Bacillus subtilis. *Biochim. Biophys. Acta BBA - Enzymol.* **268**, 373–380 (1972).
209. Diesterhaft, M. & Freese, E. Role of Pyruvate Carboxylase, Phosphoenolpyruvate Carboxykinase, and Malic Enzyme during Growth and Sporulation of Bacillus subtilis. *J. Biol.*

Chem. **248**, 6062–6070 (1973).

210. Monahan, L. G., Hajduk, I. V., Blaber, S. P., Charles, I. G. & Harry, E. J. Coordinating Bacterial Cell Division with Nutrient Availability: a Role for Glycolysis. *mBio* **5**, (2014).
211. Dombrauckas, J. D., Santarsiero, B. D. & Mesecar, A. D. Structural Basis for Tumor Pyruvate Kinase M2 Allosteric Regulation and Catalysis †,‡. *Biochemistry* **44**, 9417–9429 (2005).
212. Mattevi, A. *et al.* Crystal structure of Escherichia coli pyruvate kinase type I: molecular basis of the allosteric transition. *Structure* **3**, 729–741 (1995).
213. Zhong, W. *et al.* Allosteric pyruvate kinase-based "logic gate" synergistically senses energy and sugar levels in Mycobacterium tuberculosis. *Nat. Commun.* **8**, (2017).
214. Suzuki, K., Ito, S., Shimizu-Ibuka, A. & Sakai, H. Crystal Structure of Pyruvate Kinase from Geobacillus stearothermophilus. *J. Biochem. (Tokyo)* **144**, 305–312 (2008).
215. Cheng, X., Friesen, R. H. E. & Lee, J. C. Effects of Conserved Residues on the Regulation of Rabbit Muscle Pyruvate Kinase. *J. Biol. Chem.* **271**, 6313–6321 (1996).
216. Muirhead, H. *et al.* The structure of cat muscle pyruvate kinase. *EMBO J.* **5**, 475–481 (1986).
217. Larsen, T. M., Benning, M. M., Rayment, I. & Reed, G. H. Structure of the Bis(Mg $^{2+}$)−ATP−Oxalate Complex of the Rabbit Muscle Pyruvate Kinase at 2.1 Å Resolution: ATP Binding over a Barrel †,‡. *Biochemistry* **37**, 6247–6255 (1998).
218. Rigden, D. J., Phillips, S. E. V., Michels, P. A. M. & Fothergill-Gilmore, L. A. The structure of pyruvate kinase from Leishmania mexicana reveals details of the allosteric transition and unusual effector specificity. *J. Mol. Biol.* **293**, 745–749 (1999).
219. Schormann, N., Hayden, K. L., Lee, P., Banerjee, S. & Chattopadhyay, D. An overview of structure, function, and regulation of pyruvate kinases. *Protein Sci.* **28**, 1771–1784 (2019).
220. Sakai, H. Possible Structure and Function of the Extra C-Terminal Sequence of Pyruvate Kinase from Bacillus stearothermophilus. *J. Biochem. (Tokyo)* **136**, 471–476 (2004).
221. Sakai, H. Mutagenesis of the Active Site Lysine 221 of the Pyruvate Kinase from Bacillus stearothermophilus. *J. Biochem. (Tokyo)* **137**, 141–145 (2005).

222. Bollenbach, T. J., Mesecar, A. D. & Nowak, T. Role of Lysine 240 in the Mechanism of Yeast Pyruvate Kinase Catalysis †. *Biochemistry* **38**, 9137–9145 (1999).

223. Teplyakov, A. *et al.* Structure of phosphorylated enzyme I, the phosphoenolpyruvate:sugar phosphotransferase system sugar translocation signal protein. *Proc. Natl. Acad. Sci.* **103**, 16218–16223 (2006).

224. Herzberg, O. *et al.* Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci.* **93**, 2652–2657 (1996).

225. Burnell, J. N. & Hatch, M. D. Regulation of C4 photosynthesis: Identification of a catalytically important histidine residue and its role in the regulation of pyruvate,Pi dikinase. *Arch. Biochem. Biophys.* **231**, 175–182 (1984).

226. Tolentino, R., Chastain, C. & Burnell, J. Identification of the amino acid involved in the regulation of bacterial pyruvate, orthophosphate dikinase and phosphoenolpyruvate synthetase. *Adv. Biol. Chem.* **03**, 12–21 (2013).

227. Real-time quantitative PCR: Theory and Applications. *Encycl. Mol. Cell Biol. Mol. Med.* **11**, 483–521 (2005).

228. Skarstad, K., Boye, E. & Steen, H. B. Timing of initiation of chromosome replication in individual Escherichia coli cells. *EMBO J.* **5**, 1711–1717 (1986).

229. Adan, A., Alizada, G., Kiraz, Y., Baran, Y. & Nalbant, A. Flow cytometry: basic principles and applications. *Crit. Rev. Biotechnol.* **37**, 163–176 (2017).

230. Magill, N. G. & Setlow, P. Properties of purified sporlets produced by spoII mutants of Bacillus subtilis. *J. Bacteriol.* **174**, 8148–8151 (1992).

231. Zheng, H. *et al.* Interrogating the *Escherichia coli* cell cycle by cell dimension perturbations. *Proc. Natl. Acad. Sci.* **113**, 15000–15005 (2016).

232. Séror, S. J. *et al.* A mutant cysteinyl-tRNA synthetase affecting timing of chromosomal replication initiation in B. subtilis and conferring resistance to a protein kinase C inhibitor. *EMBO J.* **13**, 2472–2480 (1994).

233. Stokke, C., Flåtten, I. & Skarstad, K. An Easy-To-Use Simulation Program Demonstrates Variations in Bacterial Cell Cycle Parameters Depending on Medium and Temperature. *PLoS ONE* **7**, e30981 (2012).

234. Snaebjornsson, M. T. & Schulze, A. Non-canonical functions of enzymes facilitate cross-talk between cell metabolic and regulatory pathways. *Exp. Mol. Med.* **50**, 34 (2018).

235. Commichau, F. M. & Stülke, J. Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression: Trigger enzymes in transcription regulation. *Mol. Microbiol.* **67**, 692–702 (2007).

236. Hsu, M.-C. & Hung, W.-C. Pyruvate kinase M2 fuels multiple aspects of cancer cells: from cellular metabolism, transcriptional regulation to extracellular signaling. *Mol. Cancer* **17**, (2018).

237. Jiang, Y. *et al.* PKM2 phosphorylates MLC2 and regulates cytokinesis of tumour cells. *Nat. Commun.* **5**, (2014).

238. Jiang, Y. *et al.* PKM2 Regulates Chromosome Segregation and Mitosis Progression of Tumor Cells. *Mol. Cell* **53**, 75–87 (2014).

239. Sizemore, S. T. *et al.* Pyruvate kinase M2 regulates homologous recombination-mediated DNA double-strand break repair. *Cell Res.* **28**, 1090–1102 (2018).

240. Azoitei, N. *et al.* PKM2 promotes tumor angiogenesis by regulating HIF-1α through NF-κB activation. *Mol. Cancer* **15**, (2016).

241. Liang, J. *et al.* Mitochondrial PKM2 regulates oxidative stress-induced apoptosis by stabilizing Bcl2. *Cell Res.* **27**, 329–351 (2017).

242. Buschow, S. I. *et al.* MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis. *Immunol. Cell Biol.* **88**, 851–856 (2010).

243. Mazurek, S., Drexler, H. C. A., Troppmair, J., Eigenbrodt, E. & Rapp, U. R. Regulation of Pyruvate Kinase Type M2 by A-Raf: A Possible Glycolytic Stop or Go Mechanism. *ANTICANCER Res.* 9 (2007).

244. Li, L., Zhang, Y., Qiao, J., Yang, J. J. & Liu, Z.-R. Pyruvate Kinase M2 in Blood Circulation Facilitates Tumor Growth by Promoting Angiogenesis. *J. Biol. Chem.* **289**, 25812–25821 (2014).

245. Yang, P. *et al.* Secreted pyruvate kinase M2 facilitates cell migration via PI3K/Akt and Wnt/β-catenin pathway in colon cancer cells. *Biochem. Biophys. Res. Commun.* **459**, 327–332 (2015).

246. Zhang, Y., Li, L., Liu, Y. & Liu, Z.-R. PKM2 released by neutrophils at wound site facilitates early wound healing by promoting angiogenesis: PKM2 facilitates wound healing. *Wound Repair Regen.* **24**, 328–336 (2016).

247. Hitosugi, T. *et al.* Tyrosine Phosphorylation Inhibits PKM2 to Promote the Warburg Effect and Tumor Growth. *Sci. Signal.* **2**, ra73–ra73 (2009).

248. Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M. & Cantley, L. C. Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature* **452**, 181–186 (2008).

249. Mukherjee, J. *et al.* PKM2 uses control of HuR localization to regulate p27 and cell cycle progression in human glioblastoma cells: PKM2 Regulates HuR-mediated p27 Expression. *Int. J. Cancer* **139**, 99–111 (2016).

250. Huang, L. *et al.* Interaction with Pyruvate Kinase M2 Destabilizes Tristetraprolin by Proteasome Degradation and Regulates Cell Proliferation in Breast Cancer. *Sci. Rep.* **6**, (2016).

251. Gowda, G. A. N. & Djukovic, D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. in *Mass Spectrometry in Metabolomics* (ed. Raftery, D.) vol. 1198 3–12 (Springer New York, 2014).

252. Zahiri, J., Bozorgmehr, J. & Masoudi-Nejad, A. Computational Prediction of Protein–Protein Interaction Networks: Algorithms and Resources. *Curr. Genomics* **14**, 397–414 (2013).

253. Zhou, M., Li, Q. & Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* **11**, 738–756 (2016).

254. Dephoure, N., Gould, K. L., Gygi, S. P. & Kellogg, D. R. Mapping and analysis of

phosphorylation sites: a quick guide for cell biologists. *Mol. Biol. Cell* **24**, 535–542 (2013).

255. Yang, W. *et al.* Nuclear PKM2 regulates β-catenin transactivation upon EGFR activation. *Nature* **480**, 118–122 (2011).
256. Gao, X., Wang, H., Yang, J. J., Liu, X. & Liu, Z.-R. Pyruvate Kinase M2 Regulates Gene Transcription by Acting as a Protein Kinase. *Mol. Cell* **45**, 598–609 (2012).
257. Keller, K. E., Doctor, Z. M., Dwyer, Z. W. & Lee, Y.-S. SAICAR Induces Protein Kinase Activity of PKM2 that Is Necessary for Sustained Proliferative Signaling of Cancer Cells. *Mol. Cell* **53**, 700–709 (2014).
258. Hosios, A. M., Fiske, B. P., Gui, D. Y. & Vander Heiden, M. G. Lack of Evidence for PKM2 Protein Kinase Activity. *Mol. Cell* **59**, 850–857 (2015).
259. Yang, W. Structural basis of PKM2 regulation. *Protein Cell* **6**, 238–240 (2015).
260. Warburg, O. On the Origin of Cancer Cells. *Science* **123**, 309–314 (1956).
261. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).
262. Xu, X. D. *et al.* Warburg Effect or Reverse Warburg Effect? A Review of Cancer Metabolism. *Oncol. Res. Treat.* **38**, 117–122 (2015).
263. Bertram, J. S. The molecular biology of cancer. *Mol. Aspects Med.* 57 (2001).
264. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
265. Loeb, L. A., Loeb, K. R. & Anderson, J. P. Multiple mutations and cancer. *Proc. Natl. Acad. Sci.* **100**, 776–781 (2003).
266. Macheret, M. & Halazonetis, T. D. DNA Replication Stress as a Hallmark of Cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 425–448 (2015).
267. Anastasiou, D. *et al.* Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.* **8**, 839–847 (2012).
268. Buck, M. D., Sowell, R. T., Kaech, S. M. & Pearce, E. L. Metabolic Instruction of Immunity.

Cell **169**, 570–586 (2017).

269. Vivarelli, S. *et al.* Gut Microbiota and Cancer: From Pathogenesis to Therapy. *Cancers* **11**, 38 (2019).
270. Dorman, C. J. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? *Nat Rev Microbiol* **11**, 349–355 (2013).
271. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci.* **84**, 7024–7027 (1987).
272. Postow, L. Topological domain structure of the Escherichia coli chromosome. *Genes Dev.* **18**, 1766–1779 (2004).
273. Le, T. B. K., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. *Science* **342**, 731–734 (2013).
274. Valens, M., Penaud, S., Rossignol, M., Cornet, F. & Boccard, F. Macrodomain organization of the Escherichia coli chromosome. *EMBO J.* **23**, 4330–4341 (2004).
275. Thiel, A., Valens, M., Vallet-Gely, I., Espéli, O. & Boccard, F. Long-Range Chromosome Organization in E. coli: A Site-Specific System Isolates the Ter Macrodomain. *PLoS Genet.* **8**, e1002672 (2012).
276. Valens, M., Thiel, A. & Boccard, F. The MaoP/maoS Site-Specific System Organizes the Ori Region of the E. coli Chromosome into a Macrodomain. *PLOS Genet.* **12**, e1006309 (2016).
277. Duigou, S. & Boccard, F. Long range chromosome organization in Escherichia coli: The position of the replication origin defines the non-structured regions and the Right and Left macrodomains. *PLOS Genet.* **13**, e1006758 (2017).
278. Wang, X., Llopis, P. M. & Rudner, D. Z. Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.* **14**, 191–203 (2013).
279. Ma, J. & Wang, M. D. DNA supercoiling during transcription. *Biophys. Rev.* **8**, 75–87 (2016).
280. Forth, S., Sheinin, M. Y., Inman, J. & Wang, M. D. Torque Measurement at the Single-Molecule Level. *Annu. Rev. Biophys.* **42**, 583–604 (2013).

281. Peter, B. J. *et al.* Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli. *Genome Biol.* **5**, R87.1-15 (2004).

282. Bohrer, C. H. & Roberts, E. A biophysical model of supercoiling dependent transcription predicts a structural aspect to gene regulation. *BMC Biophys.* **9**, 2 (2015).

283. Lesterlin, C., Mercier, R., Boccard, F., Barre, F. X. & Cornet, F. Roles for replichores and macrodomains in segregation of the Escherichia coli chromosome. *EMBO Rep.* **6**, 557–562 (2005).

284. Espéli, O. *et al.* A MatP-divisome interaction coordinates chromosome segregation with cell division in E. coli. *EMBO J.* **31**, 3198–3211 (2012).

285. Esnault, E., Valens, M., Espéli, O. & Boccard, F. Chromosome Structuring Limits Genome Plasticity in Escherichia coli. *Plos One* **9**, 1–8 (2007).

286. Bryant, J. A., Sellars, L. E., Busby, S. J. W. & Lee, D. J. Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic Acids Res.* **42**, 11383–11392 (2014).

287. Junier, I., Hérisson, J. & Képès, F. Genomic organization of evolutionarily correlated genes in bacteria: Limits and strategies. *J. Mol. Biol.* **419**, 369–386 (2012).

288. Sobetzko, P., Travers, A. & Muskhelishvili, G. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl. Acad. Sci.* **109**, E42–E50 (2012).

289. Sobetzko, P., Glinkowska, M., Travers, A. & Muskhelishvili, G. DNA thermodynamic stability and supercoil dynamics determine the gene expression program during the bacterial growth cycle. *Mol. Biosyst.* **9**, 1643 (2013).

290. Képès, F. Periodic transcriptional organization of the E. coli genome. *J. Mol. Biol.* **340**, 957–964 (2004).

291. Képès, F. Periodic epi-organization of the yeast genome revealed by the distribution of promoter sites. *J. Mol. Biol.* **329**, 859–865 (2003).

292. Képès, F. & Vaillant, C. Transcription-Based Solenoidal Model of Chromosomes. *Complexus*

1, 171–180 (2004).

293. Wyrick, J. J. Genome-Wide Distribution of ORC and MCM Proteins in S. cerevisiae: High-Resolution Mapping of Replication Origins. *Science* **294**, 2357–2360 (2001).
294. Lieb, J. D., Liu, X., Botstein, D. & Brown, P. O. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein–DNA association. *Nat. Genet.* **28**, 327–334 (2001).
295. Simon, I. *et al.* Serial Regulation of Transcriptional Regulators in the Yeast Cell Cycle. *Cell* **106**, 697–708 (2001).
296. Wingender, E. TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* **24**, 238–241 (1996).
297. Dröge, P. & Müller-Hill, B. High local protein concentrations at promoters: Strategies in prokaryotic and eukaryotic cells. *BioEssays* **23**, 179–183 (2001).
298. Müller-Hill, B. The function of auxiliary operators. *Mol. Microbiol.* **29**, 13–18 (1998).
299. Norris, V. Modelling Escherichia coli. The concept of competitive coherence. *Life Sci.* **321**, 777–787 (1998).
300. Seo, S. W., Kim, D., Szubin, R. & Palsson, B. O. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. *Cell Rep.* **12**, 1289–1299 (2015).
301. Seo, S. W. *et al.* Deciphering fur transcriptional regulatory network highlights its complex role beyond iron metabolism in Escherichia coli. *Nat. Commun.* **5**, 1–10 (2014).
302. Seo, S. W., Kim, D., O'Brien, E. J., Szubin, R. & Palsson, B. O. Decoding genome-wide GadEWX-transcriptional regulatory networks reveals multifaceted cellular responses to acid stress in Escherichia coli. *Nat. Commun.* **6**, 1–8 (2015).
303. Seo, S. W. *et al.* Revealing genome-scale transcriptional regulatory landscape of OmpR highlights its expanded regulatory roles under osmotic stress in Escherichia coli K-12 MG1655. *Sci. Rep.* **7**, 1–10 (2017).
304. Chung, H. J., Bang, W. & Drake, M. a. Comprehensive Stress Response of Reviews

Escherichia coli in Food Science and Food Safety. *Compr. Rev. InFood Sci. Food Saf.* **5**, 52–64 (2006).

305. Horiuchi, H., Takagi, M. & Yano, K. Relaxation of supercoiled plasmid DNA by oxidative stresses in Escherichia coli. *J. Bacteriol.* **160**, 1017–1021 (1984).
306. McClellan, J. A., Boublíková, P., Palecek, E. & Lilley, D. M. Superhelical torsion in cellular DNA responds directly to environmental and genetic factors. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8373–8377 (1990).
307. Mizushima, T., Natori, S. & Sekimizu, K. Relaxation of supercoiled DNA associated with induction of heat shock proteins in Escherichia coli. *Mol Gen Genet* **238**, 1–5 (1993).
308. Mizushima, T., Kataoka, K., Ogata, Y., Inoue, R. & Sekimizu, K. Increase in negative supercoiling of plasmid DNA in Escherichia coli exposed to cold shock. *Mol. Microbiol.* **23**, 381–386 (1997).
309. Balke, V. L. & Gralla, J. D. Changes in the linking number of supercoiled DNA accompany growth transitions in Escherichia coli. *J. Bacteriol.* **169**, 4499–4506 (1987).
310. Ceci, P. *et al.* DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Res.* **32**, 5935–5944 (2004).
311. Hacker, W. C., Li, S. & Elcock, A. H. Features of genomic organization in a nucleotide-resolution molecular model of the Escherichia coli chromosome. *Nucleic Acids Res.* **45**, 7541–7554 (2017).
312. Bouyioukos, C., Elati, M. & Képès, F. Analysis tools for the interplay between genome layout and regulation. *BMC Bioinformatics* **17**, 191 (2016).
313. Bouyioukos, C., Elati, M. & Képès, F. Protocols for probing genome architecture of regulatory networks in hydrocarbon and lipid microorganisms. *Hydrocarb. Lipid Microbiol. Protoc. - Springer Protoc. Handb.* 1–29 (2015) doi:10.1007/8623.
314. Junier, I., Hérisson, J. & Képès, F. Periodic pattern detection in sparse boolean sequences.

Algorithms Mol. Biol. AMB **5**, 31 (2010).

315. R studio Team. RStudio: Integrated Development for R. *R Studio Inc Boston* (2015).
316. R Core Team. R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna Austria* (2019).
317. Rhee, H. S. & Pugh, B. F. ChIP-exo Method for Identifying Genomic Location of DNA-Binding Proteins with Near-Single-Nucleotide Accuracy. in *Current Protocols in Molecular Biology* (eds. Ausubel, F. M. et al.) (John Wiley & Sons, Inc., 2012). doi:10.1002/0471142727.mb2124s100.
318. Santos-Zavaleta, A. *et al.* RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res.* **47**, D212–D220 (2019).
319. Gama-Castro, S. *et al.* RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res.* **44**, D133–D143 (2016).
320. Salgado, H. *et al.* RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res.* **41**, D203–D213 (2013).
321. Tas, H., Nguyen, C. T., Patel, R., Kim, N. H. & Kuhlman, T. E. An Integrated System for Precise Genome Modification in Escherichia coli. *PLOS ONE* **10**, e0136963 (2015).
322. Cambray, G. *et al.* Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res.* **41**, 5139–5148 (2013).
323. Chiang, S. M. & Schellhorn, H. E. Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Arch. Biochem. Biophys.* **525**, 161–169 (2012).
324. Storz, G., Christman, M. F., Sies, H. & Ames, B. N. Spontaneous mutagenesis and oxidative damage to DNA in Salmonella typhimurium. *Proc. Natl. Acad. Sci.* **84**, 8917–8921 (1987).
325. Bauer, A. W., Kirby, W. M. M., Sherris, J. C. & Turck, M. Antibiotic Susceptibility Testing by a Standardized Single Disk Method. *Am. J. Clin. Pathol.* **45**, 493–496 (1966).

326. Cho, B.-K., Kim, D., Knight, E. M., Zengler, K. & Palsson, B. O. Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. *BMC Biol.* **12**, 4 (2014).

327. Gerschman, R., Gilbert, D. L., Nye, S. W. & Dwyer, P. Oxygen Poisoning and X-irradiation: A Mechanism in Common. *Science* **119**, 623–626 (1954).

328. Imlay, J. A. The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* **11**, 443–454 (2013).

329. Park, S., You, X. & Imlay, J. A. Substantial DNA damage from submicromolar intracellular hydrogen peroxide detected in Hpx- mutants of Escherichia coli. *Proc. Natl. Acad. Sci.* **102**, 9317–9322 (2005).

330. Boehme, D. E., Vincent, K. & Brown, O. R. Oxygen and toxicity inhibition of amino acid synthesis. *Nature* **262**, 418–420 (1976).

331. Kuo, C. F., Mashino, T. & Fridovich, I. alpha&beta Dihydroxyisovalerate Dehydratase A SUPEROXIDE-SENSITIVE ENZYME. *J. Biol. Chem.* **262**, 4724–4727 (1987).

332. Anjem, A. & Imlay, J. A. Mononuclear Iron Enzymes Are Primary Targets of Hydrogen Peroxide Stress. *J. Biol. Chem.* **287**, 15544–15556 (2012).

333. Choi, H.-J. *et al.* Structural Basis of the Redox Switch in the OxyR Transcription Factor. *Cell* **105**, 103–113 (2001).

334. Aslund, F., Zheng, M., Beckwith, J. & Storz, G. Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol--disulfide status. *Proc. Natl. Acad. Sci.* **96**, 6161–6165 (1999).

335. Seo, S. W., Kim, D., Szubin, R. & Palsson, B. O. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. *Cell Rep.* **12**, 1289–1299 (2015).

336. Gallegos, M.-T., Schleif, R., Bairoch, A., Hofmann, K. & Ramos, J. L. AraC/XylS Family of Transcriptional Regulators. *MICROBIOL MOL BIOL REV* **61**, 18 (1997).

337. Martin, R. G. & Rosner, J. L. The AraC transcriptional activators. *Curr. Opin. Microbiol.* **4**, 132–137 (2001).

338. Duval. MarA, SoxS and Rob of Escherichia coli – Global Regulators of Multidrug Resistance, Virulence and Stress Response. *Int. J. Biotechnol. Wellness Ind.* **2**, 101–124 (2013).

339. Li, Z. & Demple, B. SoxS, an Activator of Superoxide Stress Genes PURIFICATION AND INTERACTION WITH DNA*. *J. Biol. Chem.* **269**, 18371–18377 (1994).

340. Fawcett, W. P. & Wolf, R. E. Purification of a MaIE-SoxS fusion protein and identification of the control sites of Escherichia coli superoxide-inducible genes. *Mol. Microbiol.* **14**, 669–679 (1994).

341. Jair, K.-W., Fawcett, W. P., Fujita, N., Ishihama, A. & Wolf, Jr., R. E. Ambidextrous transcriptional activation by SoxS: requirement for the C-terminal domain of the RNA polymerase alpha subunit in a subset of Escherichia coli superoxide-inducible genes. *Mol. Microbiol.* **19**, 307–317 (1996).

342. Fillat, M. F. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Arch. Biochem. Biophys.* **546**, 41–52 (2014).

343. Baggf, A. & Neilands, J. B. Ferric Uptake Regulation Protein Acts as a Repressor, Employing Iron(I1) as a Cofactor To Bind the Operator of an Iron Transport Operon in Escherichia coli? 7.

344. Tao, K., Makino, K., Yonei, S. & Nakata, A. Molecular cloning and nucleotide sequencing of oxyR, the positive regulatory gene of a regulon for an adaptive response to oxidative stress in Escherichia cob':Homologies between OxyR protein and a family of bacterial activator proteins. *Mol Gen Genet* **218**, 371–376 (1989).

345. Christman, M. F., Storz, G. & Ames, B. N. OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. *Proc. Natl. Acad. Sci.* **86**, 3484–3488 (1989).

346. Kullik, I., Stevens, J., Toledano, M. B. & Storz, G. Mutational analysis of the redox-sensitive

transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J. Bacteriol.* **177**, 1285–1291 (1995).

347. Storz, G., Tartaglia, L. & Ames, B. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science* **248**, 189–194 (1990).
348. Toledano, M. B. *et al.* Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: A mechanism for differential promoter selection. *Cell* **78**, 897–909 (1994).
349. Tao, K., Fujita, N. & Ishihama, A. Involvement of the RNA polymerase ? subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. *Mol. Microbiol.* **7**, 859–864 (1993).
350. Seo, S. W., Kim, D., Szubin, R. & Palsson, B. O. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in Escherichia coli K-12 MG1655. *Cell Rep.* **12**, 1289–1299 (2015).
351. Christman, M. F., Morgan, R. W., Jacobson, F. S. & Ames, B. N. Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in Salmonella typhimurium. *Cell* **41**, 753–762 (1985).
352. Fossum, S., Crooke, E. & Skarstad, K. Organization of sister origins and replisomes during multifork DNA replication in Escherichia coli. *EMBO J.* **26**, 4514–4522 (2007).
353. Sousa, C., de Lorenzo, V. & Cebolla, A. Modulation of gene expression through chromosomal positioning in Escherichia coli. *Microbiology* **143**, 2071–2078 (1997).
354. Block, D. H. S., Hussein, R., Liang, L. W. & Lim, H. N. Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Res.* **40**, 8979–8992 (2012).
355. Johnson, J. R., Clabots, C. & Rosen, H. Effect of Inactivation of the Global Oxidative Stress Regulator oxyR on the Colonization Ability of Escherichia coli O1:K1:H7 in a Mouse Model of Ascending Urinary Tract Infection. *Infect. Immun.* **74**, 461–468 (2006).
356. Mercier, G. A haploid-specific transcriptional response to irradiation in Saccharomyces cerevisiae. *Nucleic Acids Res.* **33**, 6635–6643 (2005).

357. Mathelier, A. & Carbone, A. Chromosomal periodicity and positional networks of genes in *Escherichia coli*. *Mol. Syst. Biol.* **6**, 366 (2010).

358. Jeong, K. S., Ahn, J. & Khodursky, A. B. Spatial patterns of transcriptional activity in the chromosome of. *Genome Biol.* 10 (2004).

359. Wright, M. A., Kharchenko, P., Church, G. M. & Segre, D. Chromosomal periodicity of evolutionarily conserved gene pairs. *Proc. Natl. Acad. Sci.* **104**, 10559–10564 (2007).

360. Bouyioukos, C., Elati, M. & Képès, F. Analysis tools for the interplay between genome layout and regulation. *BMC Bioinformatics* **17**, (2016).

361. Pavitt†, G. D. & Higgins, C. F. Chromosomal domains of supercoiling in Salmonella typhimurium. *Mol. Microbiol.* **10**, 685–696 (1993).

362. Bilyk, B., Horbal, L. & Luzhetskyy, A. Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in Streptomyces albus J1074. *Microb. Cell Factories* **16**, (2017).

363. Miller†, W. G. & Simons, R. W. Chromosomal supercoiling in Escherichia coli. *Mol. Microbiol.* **10**, 675–684 (1993).

364. Sauer, C. *et al.* Effect of Genome Position on Heterologous Gene Expression in *Bacillus subtilis* : An Unbiased Analysis. *ACS Synth. Biol.* **5**, 942–947 (2016).

365. Schmid, M. B. & Roth, J. R. Gene location affects expression level in Salmonella typhimurium. *J. Bacteriol.* **169**, 2872–2875 (1987).

366. Brambilla, E. & Sclavi, B. Gene Regulation by H-NS as a Function of Growth Conditions Depends on Chromosomal Position in *Escherichia coli*. *G3amp58 GenesGenomesGenetics* **5**, 605–614 (2015).

367. Berger, M. *et al.* Genes on a Wire: The Nucleoid-Associated Protein HU Insulates Transcription Units in Escherichia coli. *Sci. Rep.* **6**, (2016).

368. Soler-Bistué, A. *et al.* Genomic Location of the Major Ribosomal Protein Gene Locus Determines Vibrio cholerae Global Growth and Infectivity. *PLOS Genet.* **11**, e1005156

(2015).

369. Chandler, M. G. & Pritchard, R. H. The effect of gene concentration and relative gene dosage on gene output inEscherichia coli. *MGG Mol. Gen. Genet.* **138**, 127–141 (1975).
370. Scholz, S. A. *et al.* High-Resolution Mapping of the Escherichia coli Chromosome Reveals Positions of High and Low Transcription. *Cell Syst.* **8**, 212-225.e9 (2019).
371. Gerganova, V. *et al.* Chromosomal position shift of a regulatory gene alters the bacterial phenotype. *Nucleic Acids Res.* **43**, 8215–8226 (2015).
372. Weng, X. & Xiao, J. Spatial organization of transcription in bacterial cells. *Trends Genet.* **30**, 287–297 (2014).
373. Wang, W., Li, G.-W., Chen, C., Xie, X. S. & Zhuang, X. Chromosome Organization by a Nucleoid-Associated Protein in Live Bacteria. *Science* **333**, 1445–1449 (2011).
374. Qian, Z., Dimitriadis, E. K., Edgar, R., Eswaramoorthy, P. & Adhya, S. Galactose repressor mediated intersegmental chromosomal connections in Escherichia coli. *Proc. Natl. Acad. Sci.* **109**, 11336–11341 (2012).
375. Garza de Leon, F., Sellars, L., Stracy, M., Busby, S. J. W. & Kapanidis, A. N. Tracking Low-Copy Transcription Factors in Living Bacteria: The Case of the lac Repressor. *Biophys. J.* **112**, 1316–1327 (2017).
376. Sellars, L. E. *et al.* Development of a new fluorescent reporter:operator system: location of AraC regulated genes in Escherichia coli K-12. *BMC Microbiol.* **17**, (2017).
377. Nguyen, H. Q. & Bosco, G. Gene Positioning Effects on Expression in Eukaryotes. *Annu. Rev. Genet.* **49**, 627–646 (2015).
378. Chen, X. & Zhang, J. The Genomic Landscape of Position Effects on Protein Expression Level and Noise in Yeast. *Cell Syst.* **2**, 347–354 (2016).
379. Wu, X.-L. *et al.* Genome-wide landscape of position effects on heterogeneous gene expression in Saccharomyces cerevisiae. *Biotechnol. Biofuels* **10**, (2017).
380. Sun, M. & Zhang, J. Chromosome-wide co-fluctuation of stochastic gene expression in

mammalian cells. *PLOS Genet.* **15**, e1008389 (2019).

Supplementary material annex I

Table S1 : Genetic constructs used in this work

Context	Genotype	Main phenotypes	Source of construction
pJet	*pJet vector, KanR flanked by homologous sequences (25 bp) and I-sceI sites*	KanR, ApR	Laboratory stock, (Tas et al, 2015)
pJet	*pJet vector, PhleoR flanked by homologous sequences (25 bp) and I-sceI sites*	PhleoR, ApR	Laboratory stock, (Tas et al, 2015)
pJet	*pJet vector, KanR+oxyRS, flanked by homologous sequences (25 bp) and I-sceI sites*	KanR, ApR	This work
pJet	*pJet vector, KanR+oxyRS, flanked by homologous sequences (25 bp) and I-sceI sites and insulated by terminator terminators his(min), T500, Bba 1006 and Trp(min)*	KanR, ApR	This work
pTKRED in MG1655	*pTKRED*	Ts sensitive pSC101, Inducible expression of λ RED from pLac, Inducible expression of I-SceI from pAra, SpR, Ts plasmid replication	(Kuhlman and Cox, 2010)
pTKRED in DH5α	*pTKRED*	Ts sensitive pSC101, Inducible expression of λ RED from pLac, Inducible expression of I-SceI from pAra, SpR, Ts plasmid replication	(Kuhlman and Cox, 2010)
pTKIP	*PMB1 origin (15-20 copies per cell), mCherry, ApR, PhleoR, LP1 and LP2*	Red colonies, ApR, PhleoR	Laboratory stock, (Kuhlman and Cox, 2010)
pTKIP	*PMB1 origin, mCherry, ApR, , LP1 and LP2, KanR+oxyRS, flanked by homologous sequences (25 bp) and I-sceI sites and insulated by terminator terminators his(min), T500, Bba 1006 and Trp(min)*	Red colonies, ApR, KanR	This work
MG1655	*ΔoxyRS*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500 , Bba 1006 and Trp(min) integrated at original location,*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500, Bba 1006 and Trp(min) integrated at position 4110708*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500, Bba 1006 and Trp(min) integrated at position 4113722*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500, Bba 1006 and Trp(min) integrated at position 2196401*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500, Bba 1006 and Trp(min) integrated at position 2200273*	Sensitivity to oxidative stress	This work
MG1655	*oxyRS insulated by terminators his(min), T500, Bba 1006 and Trp(min) integrated at position 2219560*	Sensitivity to oxidative stress	This work

Table S2 : Primers used for genetic construction

Primer name	Purpose	Orientation	Sequence
pSH156	Construction plasmid 61	for	CGTCTATCTCTTTAGCTACCGGTCA
pSH157	Construction plasmid 61	rev	CGAAAAGCCCGCTTAATAACTGTT
pSH158	Construction plasmid 61	for	GTTATTAAGCGGGCTTTTCGAGCGGATCCTGGAGATC
pSH159	Construction plasmid 61	rev	GGTAGCTAAAGAGATAGACGTTAAACCGCCTG TTTTAAAACTTTATCGAAATG
pSH178	Construction plasmid 62	for	CAAATCCCGCCGAAAGGCGGGATTTTCTGGCTCGAG TTTTTCAGCAAGATATCTTG
pSH179	Construction plasmid 62	rev	TCCCGCAGCTTGCCAGCACTGCCCGCCTAATGAGCG GGCTTTTTTTTCATCCGGAAGATCTGGCG
pSH182	Construction plasmid 62	for	GCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTC CTACAATATTCTCAGCTGCCATGG
pSH183	Construction plasmid 62	rev	CCAAAAAATGGGAGCAGGAGAAAAAAAAACCCCGC C CCTGACAGGGCGGGGTTTTTTTTAGATCTTCTAGAAA GATATCTGGTCTGACCG
pSH220	Construction Plasmid 92	for	ccccaaggtccaaacggtgaTAATACGACTCA CTATAGGG
pSH221	Construction Plasmid 92	rev	ttcagggatgaggcgccatcAAGAACATCGAT TTTCCATGGCAG
pSH222	Construction Plasmid 92	for	GATGGCGCCTCATCCCTGAA
pSH223	Construction Plasmid 92	rev	TCACCGTTTGGACCTTGGG
pSH148	onstruction *ORS_init, ΔoxyR*	for	GCAGGCGATTGCTTTTGCGCAGGCTCGGTTAGGGTA AGAACATTTATATGTAATACGACTCACTATAGGG
pSH149	onstruction *ORS_init, ΔoxyR*	rev	CGGAAGCCTATCGGGTAGCTGCGTTAAACGGTTTAA ACCGCCTGTTTTAAAAGAACATCGATTTTCCATGGCA G
pSH240	onstruction *ORS_init, ΔoxyR*	rev	GCCAAAAGTACTGCTGGCAC
pSH120	Construction *ORS31*	for	GATCATCAGTGTGAGATTTTGCGTTAAGGCGAAGAG TTAAGGAAAGTAAGTAATACGACTCACTATAGGG
pSH121	Construction *ORS31*	rev	GTAATCGTTAAAGCTGCAGAAGCGGCTAAACAGGCT TAAGTCTGACAGGTGCCGGATTTCATATCCGGCAAA GAACATCGATTTTCCATGGCAG
pSH122	Construction *ORS32*	for	CATCATCACTCATTTATCAACCGTTTGATGCCGGCAT A TCGCGGGATGATCAATAAGATGAGTGCGGATTTGCT CATCGTGCCGTTTATCGATAAGTAACTAATACGACTC ACTATAGGG
pSH123	Construction *ORS32*	rev	ATGACGAAACATTTACGTCGCCGACCGGGCCATATG ACAGCGGAGCATTACTGGTAAGCGAAGAACATCGAT TTTCCATGGCAG
153F	Construction *ORS_FT8*	for	CATCGACAGCGCCTTTTCTTTATAAATTCCTAAAGTTG TTTTCTTGCGATTACGGCCCCAAGGTCCAAAC
153R	Construction *ORS_FT8*	rev	CAGTTGAATGCAGATGCTACCAGTATTTATGCGGGTT AGAGAGAGACAAATTGGCTTCAGGGATGAGGCG
154F	Construction *ORS_FT9*	for	CACCGCCAGCGAGCTAATTAACGATATTGAAGCCCT G TTTGAATAAGGAATACGGCCCCAAGGTCCAAAC
154R	Construction *ORS_FT9*	rev	TTCCAGTTGGGCGTTATCCGCCAGCCACGGTAATTC CTTGTCCATGCTCTTTGGCTTCAGGGATGAGGCG
158F	*Construction ORS_FT13*	for	TTCAGCGAATTAAGAGAAAAGCATAGTTGATAATGG CGGGGTTAGCTTGATACGGCCCCAAGGTCCAAAC
158R	Construction *ORS_FT13*	rev	CGAATATGCCTCTTAAGTATTTATCTGATTCTGCAACT GAATCCTTCCGCTTGGCTTCAGGGATGAGGCG
pSH03	Verification *ORS_init, ΔoxyRS*	for	CGATTCTGTCGCTGCAATCG
pSH04	erification *ORS_init, ΔoxyR*	rev	TCGCTGGCGTTACAGAACAT
pSH198	Verification *ORS31 Insertion location*	for	GTATTGCTGGCGACACAACG
pSH199	Verification *ORS31 Insertion location*	rev	CAGATCGACGCGGTACTGAA
pSH200	Verification *ORS32 Insertion location*	for	ATGGCCGAAGACAAAACTGC
pSH201	Verification *ORS32 Insertion location*	rev	ACGTTATGCCGCCGACTTAA
153 s1	Verification *ORS_FT8 Insertion location*	for	ATAATCCCCCTTCAAGGCGC
153 s2	Verification *ORS_FT8 Insertion location*	rev	GCAACTCGTTTCCCTGTTGC
154 s1	Verification *ORS_FT9 Insertion location*	for	GCAATGCCAGTGAACTGACC
154 s2	Verification *ORS_FT9 Insertion location*	rev	TCAGTCGCATATTCCAGCCC
158 s1	Verification *ORS_FT13 Insertion location*	for	TGCCCAGAGCTTTAAGAGTGG
158 s2	Verification *ORS_FT13 Insertion location*	rev	AAATATGACGCCGAGCCTGG

Titre : Découverte d'un nouveau type de régulateur de réplication d'ADN : l'enzyme glycolytique pyruvate kinase en *Bacillus subtilis*

Mots clés : *B. subtilis*, pyruvate kinase, réplication d'ADN

Résumé : On sait depuis longtemps que le fenêtrage de la réplication de l'ADN dans le cycle cellulaire chez les bactéries varie avec la richesse en nutriments. Cependant, le mécanisme sous-jacent reste pour l'instant inconnu. Le but de ce travail est de déterminer si PykA, une enzyme du métabolisme central carboné (MCC), est impliquée dans ce contrôle temporel de la réplication chez la bactérie Bacillus subtilis et si, pour assurer cette fonction, PykA utilise une activité non-métabolique cryptique qui lui permet d'agir en tant qu'effecteur temporel de la réplication en transmettant des informations sur l'état métabolique de la cellule à l'appareil de réplication.Pour tester ces hypothèses, l'effet de mutations dans le gène pykA sur la réplication a été évalué dans un milieu de culture où son activité métabolique n'est pas nécessaire. Les paramètres de réplication dans ces mutants ont alors été mesurés par des approches de cytométrie en flux et de PCR quantitative. Nos résultats ont confirmé que la délétion de pykA n'a pas d'effet sur la croissance des cellules dans le milieu utilisé mais que cette mutation affecte le contrôle temporel de la réplication en modifiant l'âge d'initiation de la réplication et la vitesse de réplication. La délétion du domaine catalytique ou du domaine « PEP utilizer » de PykA a montré que chacun de ces domaines joue un rôle dans le contrôle temporel de la réplication, assignant pour la première fois une fonction de réplication à ces domaines. Une mutagenèse ciblée du domaine catalytique a montré que seulement certains acides aminés impliqués dans l'activité métabolique de PykA sont importants pour la réplication. La mutagenèse des acides aminés conservés dans le motif Thr-Ser-His (TSH) du PEP utilizer suggère un rôle important de la thréonine dans le contrôle temporel de la réplication. L'expression du PEP utilizer seul (non fusionné au domaine catalytique), a permis de découvrir plusieurs principes de réciprocité entre les deux domains. Enfin, des tests in vitro ont démontré que PykA peut directement stimuler l'activité de la polymérase DnaE et inhiber l'activité de l'hélicase DnaC. Ensemble, ces résultats indiquent que PykA a une activité cryptique de réplication.
Pour démontrer que PykA transmet l'état métabolique de la cellule à la machinerie de réplication, l'effet de mutations dans pykA sur les paramètres de réplication de cellules subissant un changement métabolique a été évalué à l'aide des techniques décrites ci-dessus. Nous avons observé que l'adaptation des paramètres de réplication au cours du changement métabolique dans les cellules sauvages se déroule en trois phases: (i) une diminution progressive de la vitesse de la fourche de réplication de 15 à 45 minutes après le changement métabolique, (ii) une stimulation de l'initiation à partir du point 30 minutes et (iii) finalement une adaptation du taux de croissance après 45 minutes. Les mutations dans le domaine catalytique ne semblent pas avoir d'effet sur l'évolution de ces paramètres. En revanche, le PEP utilizer, fusionné ou non au domaine catalytique, affecte l'adaptation des paramètres de réplication par un processus nouveau dépendant de l'acide aminé His et non Thr du motif TSH. Ainsi donc, le PEP utilizer et plus généralement PykA, agit sur la réplication selon des voies particulières en fonction du métabolisme.Nous concluons que PykA est un nouveau type de régulateur de la réplication de l'ADN qui facilite le positionnement temporel de la réplication par le biais de processus variant avec la richesse des sources de carbone fournies dans l'environnement. Ces travaux, combinés à des travaux antérieurs dans ce domaine, suggèrent que le contrôle temporel de la réplication dans un large éventail de conditions de croissance métabolique est réalisé au moins en partie via un réseau d'enzymes MCC ayant des activités cryptiques. Cette conclusion a des implications importantes dans notre compréhension de la biologie cellulaire.

Title : Discovery of a new type of regulator of DNA replication: the glycolytic enzyme pyruvate kinase in *Bacillus subtilis*

Keywords : *B. subtilis*, pyruvate kinase, DNA replication

Abstract : It has been known for decades that the timing of DNA replication within the cell cycle in bacteria is coupled to nutrient richness. However, the mechanisms that enable this temporal gating have remained elusive. The goal of this work was to test whether the CCM enzyme PykA in the bacterium B. subtilis helps achieve this gating by moonlighting as a temporal effector of replication that communicates information about the metabolic state of the cell to the replication machinery.

To demonstrate that PykA moonlights as a replication protein, the effect of pykA mutations on DNA replication parameters was assessed using run-out flow cytometry and marker frequency analysis by qPCR in growth conditions where its metabolic activity is dispensable.

Deletion of pykA did not affect the growth rate during these conditions, but it affected the replication speed and timing of initiation, altering therefore the temporal control of replication. The deletion of either the catalytic domain or the PEP utilizer domain of PykA showed that both domains play a role in the temporal control of replication, identifying for the first time a replication function to these domains. Mutagenesis of specific amino acids inside the catalytic domain showed that some, but not all key catalytic amino acids are involved in replication. Mutagenesis of conserved amino acids in the TSH motif of the PEP utilizer on the other hand suggests an important role for threonine phosphorylation. Expression of the PEP utilizer detached from the catalytic domain in trans unexpectedly affected replication control at medium copy numbers. This phenotype depended strictly on H of the TSH motif, establishing a role for H in replication. Furthermore, the effect of expression of free PEP utilizer may depend of an unidentified inhibitor effector and is fully suppressed by a mutation in the catalytic domain. Moreover, expression of the catalytic domain in cis and the PEP utilizer in trans did not restore proper replication control. Finally, in vitro assays demonstrated that PykA can directly stimulate DnaE polymerase activity and inhibit DnaC helicase activity. Together, these results imply that PykA moonlights in replication.

To demonstrate that PykA communicates information about the metabolic state of the cell to the replication machinery, the effect of pykA mutations on DNA replication parameters during a metabolic shift was assessed using the techniques described above.

The adaptation of replication parameters during the metabolic shift in wild type cells unexpectedly consisted of three phases: (i) a decrease of replication fork speed between 15 and 45 min after the shift, (ii) an advancement of initiation timing and increase in initiation frequency from 30 min onwards and (iii) an adaptation of the growth rate sometime after 45 min after the shift. Mutations in the catalytic domain did not clearly demonstrate a change in the replication response to the shift. By contrast, the effect of the PEP utilizer domain (both in cis and in trans) on replication parameters did change during the shift. Moreover, the H amino acid, not T, was important for replication control during later time points of the shift. This suggests that the regulatory behaviour of the PEP utilizer and more generally PykA changes in response to metabolism.

Overall, we conclude that PykA is a new type of regulator of DNA replication that helps the temporal positioning of replication in the cell cycle through processes that varies with the richness of the carbon sources provided in the environment. This work, combined with previous work in this field suggest that the temporal control of replication in a large range of metabolic growth conditions is achieved at least in part via a network of moonlighting CCM enzymes with important implications for our basic understanding of cell biology and human health.