



HAL
open science

ENGINEERING AUTONOMOUS AND PROGRAMMABLE BIOSENSORS THROUGH SYNTHETIC BIOLOGY: INTEGRATING MULTIPLEXED BIOMARKER DETECTION AND BIOMOLECULAR SIGNAL PROCESSING INTO NEXT-GENERATION DIAGNOSTICS

Alexis Courbet

► **To cite this version:**

Alexis Courbet. ENGINEERING AUTONOMOUS AND PROGRAMMABLE BIOSENSORS THROUGH SYNTHETIC BIOLOGY: INTEGRATING MULTIPLEXED BIOMARKER DETECTION AND BIOMOLECULAR SIGNAL PROCESSING INTO NEXT-GENERATION DIAGNOSTICS. Bioinformatics [q-bio.QM]. University of Montpellier, 2015. English. NNT: . tel-01615133

HAL Id: tel-01615133

<https://hal.science/tel-01615133>

Submitted on 12 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de
Docteur

Délivré par l'**Université de Montpellier**

Préparée au sein de l'école doctorale CBS2 – Sciences
chimiques et biologiques pour la Santé
Et de l'unité de recherche CNRS FRE3690 Sys2Diag
Complex biological systems modeling and engineering for diagnosis

Spécialité : **Biologie de Synthèse**

Présentée par **Alexis Courbet**

**ENGINEERING AUTONOMOUS AND
PROGRAMMABLE BIOSENSORS
THROUGH SYNTHETIC BIOLOGY:
INTEGRATING MULTIPLEXED
BIOMARKER DETECTION AND
BIOMOLECULAR SIGNAL PROCESSING
INTO NEXT-GENERATION DIAGNOSTICS**

Soutenue le 7 décembre 2015 devant le jury composé de

Pr. Eric Renard, PU-PH, CHRU & CNRS
Dr. Denis Pompon, DR, CNRS
Pr. Joerg Stelling, ETHZ
Pr. Yaakov Benenson, ETHZ
Dr. Franck Molina, DR, CNRS
Dr. François Fages, DR, CNRS

President du Jury
Examineur
Rapporteur
Rapporteur
Directeur de Thèse
Invité

©Copyright 2015 Alexis Courbet
All Rights Reserved

*Pour l'enfant, amoureux de cartes et d'estampes,
L'univers est égal à son vaste appétit.
Ab! que le monde est grand à la clarté des lampes!
Aux yeux du souvenir que le monde est petit!*

*To a child who is fond of maps and engravings
The universe is the size of his immense hunger.
Ab! how vast is the world in the light of a lamp!
In memory's eyes how small the world is!*

Charles Baudelaire, *Le Voyage*

À mes parents

Engineering autonomous and programmable biosensors through synthetic biology: Integrating multiplexed biomarker detection and molecular signal processing into next-generation diagnostics

Abstract

The promise for real precision medicine is contingent on innovative technological solutions to diagnosis. In the post-genomic era, synthetic biology approaches to medicine provide new ways to probe, monitor and interface human pathophysiology. Emerging as a mature field increasingly transitioning to the clinics, synthetic biology can be used to apply engineering principles to design and build biological systems with clinical specifications. A particularly tantalizing application is to develop versatile, programmable and *intelligent* diagnostic devices closely interconnected with therapy. This thesis presents novel engineering concepts and approaches to design synthetic biological devices interfacing human diseases in clinical samples through biomolecular digital signal processing, in light of a need for dramatic improvements in capabilities and robustness. It addresses primarily the engineering of synthetic gene circuits through integrase based digital genetic amplifiers and logic gates, to integrate modular and programmable biosensing of biomarkers and diagnostic decision algorithms into bacteria. It then investigates systematic bottom-up methodologies to program microscale synthetic protocells performing medical biosensing and biocomputing operations. We demonstrate streamlined microfluidic fabrication methods and solutions to implement complex Boolean operation using integrated synthetic biochemical circuits. This contribution also extends to the characterization of protocell design space through novel computer assisted design frameworks, as well as the analysis of mathematical and biological evidence for universal protocellular biocomputing devices. The articulation of biological governing principles and medical implications for the synthetic devices developed in this work was further validated in the clinic, and initiates new models towards next-generation diagnostics. This work envisions that synthetic biology is preparing the future of medicine, supporting and speeding up the development of diagnostics with novel capabilities to bring direct improvement in biotechnologies from the clinical lab to the patient.

Keywords: Synthetic biology, bioengineering, biosensor, biocomputing, gene circuits, biochemical circuits, biomarkers, medical diagnosis, *in vitro* diagnostics, molecular diagnostics, precision medicine, translational medicine

Résumé

Les promesses de la médecine de précision dépendent de nouvelles solutions technologiques pour le diagnostic. Dans l'ère post-génomique, les approches de biologie synthétique pour la médecine apportent de nouvelles façons de sonder, monitorer et interfacer la physiopathologie humaine. Émergeant en tant que champ scientifique mature dont la transition clinique s'accélère, la biologie synthétique peut être utilisée pour appliquer des principes d'ingénierie afin de concevoir et construire des systèmes biologiques comprenant des spécifications cliniques. Une application particulièrement intéressante est de développer des outils diagnostiques polyvalents, programmables et *intelligents* étroitement interconnectés avec la thérapie. Cette thèse présente de nouveaux concepts et approches d'ingénierie pour concevoir des dispositifs biosynthétiques capables d'interfacer les maladies humaines dans des échantillons cliniques en exploitant du traitement de signal au niveau biomoléculaire, à la lumière d'un besoin croissant en termes de capacités et de robustesse. Cette thèse s'intéresse en premier lieu à l'ingénierie de circuits synthétiques de gènes, reposant sur les portes logiques à intégrases, pour intégrer des opérations modulaires et programmables de biodétection de biomarqueur associées à des algorithmes de décisions au sein de population de bactéries. Elle s'intéresse ensuite à des méthodologies systématiques dites *bottom-up*, pour programmer des protocellules synthétiques microscopiques, capables d'exécuter des opérations de biodétection médicale et de biocomputation. Nous décrivons le développement de méthodes simples de fabrications microfluidiques associées à des solutions pour implémenter des opérations booléennes complexes en utilisant des circuits biochimiques synthétiques. Cette contribution s'élargit aussi à la caractérisation de l'espace de conception de protocellules à l'aide d'approches de design assisté par ordinateur, ainsi qu'à l'analyse de preuves mathématiques et biologiques pour l'utilisation de protocellules comme des dispositifs universels de calcul. L'articulation des principes biologiques fondamentaux avec les implications médicales concernant les dispositifs biosynthétiques développés dans ce travail, a été jusqu'à la validation clinique et initie de nouveaux modèles pour le développement de diagnostics de nouvelle génération. Ce travail prévoit que la biologie synthétique est en train de préparer le futur de la médecine, en supportant et accélérant le développement de diagnostics avec de nouvelles capacités, apportant un progrès biotechnologique direct depuis le laboratoire de biologie clinique jusqu'au patient.

Mots-clés: Biologie synthétique, bioingénierie, biosenseur, biocomputation, circuits génétiques, circuits biochimiques, biomarqueurs, diagnostic médical, diagnostic *in vitro*, diagnostic moléculaire, médecine de précision, médecine translationnelle

Acknowledgements

First and foremost I would like to thank my parents, my brother, my grand parents and my family, for supporting me and encouraging me in my decisions and commitments, and bringing me the love and intellectual stimulation that built me.

Secondly, I would particularly like to thank Franck Molina, my PhD supervisor with thousand ideas as precious and incredible as contagious, who offered me the chance to complete my goals, always supported me and my work, and introduced me to an extraordinary scientific world, amazing collaborators and formidable intellectual horizons. Franck was not only a great mentor, but also a talented, energetic and fascinating communicant of his broad passions and overflowing interest he has of the world.

I would like to warmly thank Eric Renard for his help, his engagement and his invaluable support and advice and for serving as a Jury member in my first and this time second thesis. I would like to also thank his team at the Department of Endocrinology of the University Hospital of Montpellier.

I am profoundly thankful to Drew Endy for pulling out his engineer hat, and for his welcoming invitation and precious advice and help. It has been a true pleasure to work in his lab at Stanford, with the great, friendly and intellectually unforgettable company of Brent Townshend, Paul Jaschke, Atri Choksi, and Olivier Borkowski. I hope our paths will cross again.

Although difficult to find sufficient words, I would like to give very special thanks to Jerome Bonnet, whom I am deeply grateful to have met, without whose help much of my work would not have been possible, and for his help, his advice, his sympathy and his patience. Jerome taught me a lot about science, but also gave me an unforgettable glimpse of how to conduct oneself as a good scientist and person.

Many, many thanks to Patrick Amar for his sympathy, his sharp mind, his fresh ideas and broad scientific culture which did not fail to illuminate our passionate discussions. It is always a real joy to share a moment with you.

I feel particularly lucky that my path crossed the one of amazing persons and scientists, as well as amazing mentors which strongly stimulated my interest for knowledge and made my commitment to science possible. Frédéric Jaisser at the Collège de France introduced me to real scientific research and gave me the taste of experimental medicine and biology. Franck Lezoualc'h and Eric Morel at the University of Paris Sud taught me the joy of cell signaling and reprogramming and how to work with animal models. Sophie Brisset and Lucie Tosca gave me the chance to discover the amazing medical science of cytogenetics, which triggered in me deep ethical and biotechnological reasoning. I am also thankful to Florence Doucet-Populaire who gave me the opportunity to deepen my passion for microbiology in a clinical and experimental Lab around formidable people. Last but not least, I am profoundly grateful to Florence Gattaceca, who is a great, talented and engaged mentor, scientist, musician and person, whose friendship mean a lot to me.

I am deeply thankful to Catherine Royer from the CBS, for her precious advice, her support and for her amazing Biophysics classes. I would also like to thank Stéphane Delbec for his great biotechnology classes, and Pascal Nouvel for his remarkable classes on Philosophy of Sciences which brought the joy, freshness and perspective I particularly needed during my PhD.

Pierre Sonigo has been an amazing source of fresh, sharp, precious, and alternative perspective on biology and medicine, and I am deeply thankful for his interest in my work, his kindness and his support and help.

I would like to thank Dinah Weissman, for her precious advice and for sharing her invaluable experience with me, as well as for the interest and support in my work she demonstrated during my PhD.

I am grateful to the entire Sys2Diag laboratory, CNRS and Alcediag staff, researchers and Bio-Rad former associates. I am very grateful and lucky to have been surrounded by skilled, friendly and cheerful colleagues and scientists: first, my dear office mate and brilliant chemist and surfer Christophe Nguyen, the star of the proteome Laurence Molina, and Gudrun Aldrian, Gwendal Lazennec, Francisco Frederico Shneider, Siem Van Der Laan, Nicolas Salvetat, Christopher Cayzac, Yoann Lannay, Sandrine Billouez, Camille Boyer, Thérèse Galindo, Liyan He, Sabine Pérès and David Jean; And former Bio-Rad colleagues: Daniel Laune, Jean Daniel Abraham, Benedicte Jardin, Delphine Merle, Dominique Piquer, Eve Mathieu-Dupas, Florian Salipante, Isabelle Molina, Janette Fareh, Laetitia Rubrecht, Lionel Valera, Pascal Galéa, Sandra Cobo, Sylvie Promé, Thibault Sibourg, Vincent Hamelle, and Véronique Moulin. I am also particularly thankful to Marc Chakiervili for his invaluable help on scripting and coding.

Nadia Vié has given me much of her time and brought me some wonderful help and advice on flow cytometry. I am also very thankful to Julien Cau and all the staff at the RIO imaging platform, that do an amazing job at bringing cutting edge technologies to researchers.

I had the chance to work with François Fages at INRIA, and I am very thankful for our friendly and fruitful collaboration, our very pleasant discussions, and all his precious advice and experience from science and industry he shared with me.

I would like to warmly thank Sarah Guizio for her friendship and wish her the best of luck for the rest of her PhD. I am also very thankful to Guillaume Cambray for his helpful critics, and sharing his precious perspectives on science along with his personal experiences.

I am thankful to Denis Pompon and Michel Desarménien for their willingness to support and take interest in my work during my PhD comitties and later to serve in the defense jury.

Last but not least, I would like to express profound gratitude to Pierre Antoine Bonnet, who has given much time and efforts to accomodate our Residency and make our scientific dreams come true, as well as Sonia Khier and Carine Masquefa for their friendship, cheerfulness, and support.

During my Residency, I was extremely lucky to have the chance to work in the Toxicology Departement at the University hospital of Montpellier, and Olivier Mathieu was a great and kind senior clinician that brought us much joy and highest degree medical expertise during our late night works.

My dear friends and colleagues Julien Bouckenheimer and Guilhem Royer provided me with a great deal of fun, kindness and mental support in the process, and I am deeply thankful to have met them. Tsveta and Konstantin Todorov gave me their unconditional friendship and support, and brought me much joy, shared their amazing wine and rakia with me, and brought me intense fun and philosophical excitement that I will never forget.

Moving towards the end of this list, I would like to especially thank all my unique, amazing and gifted dear old friends outside of science, from Paris and abroad, for filling my life with both joy and meaning.

Many thanks to mice, rats, mammalian cells, bacteria, viruses and patients that gave so much to knowledge and science.

Finally, I would like to thank all the scientists and professors at the University of Paris-Sud and Montpellier, as well as doctors and medical practitioners who during my long studies communicated me their passion for biological and medical sciences. I also want to thank all the public health structures and research administration that support the advancement of knowledge and contribute to human, social and scientific progress: the CNRS, INSERM, the Public Hospital, the University and the structures that support young researchers, which made all my long studies and research possible.

Foreword

La géométrie n'est pas vraie, elle est avantageuse.

Geometry is not true, it is advantageous.

Henri Poincaré, *La science et l'hypothèse*

Biology has operated a natural evolution during the 20th century. Since the foundations of enzymology, through Jacques Monod and the advent of molecular biology and cybernetics, finally enriched by the holistic views of systems biology and quantitative biotechnologies of the 90's, this path finally resolved in the beginning of this century in a modern formulation: synthetic biology. It is constituted as an interdisciplinary approach focusing on the flow of matter, information and energy in biological systems. Successor of molecular biology and genetic engineering, synthetic biology is synonymous to the paradigm shift in life sciences, effectively captured in the expression *understanding by building*. The famous "What I cannot create, I do not understand" by Feynman, or to quote Stéphane Leduc, to analysis, succeeds "synthesis". This later unlikely visionary, proposed in 1912 that "Biology is a science like any other, (...) it must be successively descriptive, analytical and synthetic". Along this idea, the contemporary Jacques Loeb proposed *abiogenesis*, the fabrication of living organisms from matter, as a main objective of biology. The revolution underway was evident with for instance the advent of synthetic chemistry. The interconnection between engineering (building) and science (understanding) is at the origin of predictive models in synthetic biology, enabling to fully exploit the nanoscale at which biological systems operate, fortified by billion years of optimization. Although synthetic biology applies engineering principles to living organisms (standardization, automation, *in silico* design...), the peculiarity of this discipline lies in its substrate, still widely misunderstood and untamed. For this reason it is perhaps one of the most ambitious modern scientific and human adventures, since synthetic biology seeks to understand and design off-balance systems, deconstruct emerging phenomena, read and rewrite the evolutionary history of life and its origins.

Living organisms can be regarded as nanomachines, which are themselves composed of the most effective nanocircuits to manipulate information, matter and energy at the molecular level. With the latter consideration and a biomedical perspective in mind, comes immediately an idea: exploiting living systems to treat. Medical practice has always used biological knowledge to move towards an ever more efficient practice, and as such synthetic biology as a new discipline finds its place: getting the most localized, fast, accurate, and intelligent medical procedure. Specifically, medical diagnosis is an exciting technological field of research that focuses on the most efficient modalities of extraction of physiological information to make it intelligible and meaningful on a clinical plan. In this sense, synthetic biology appears as a wonderful tool to probe patient's biology at the molecular level and interface it with clinical practice. In this work, I thus explored the potential synergy between this new discipline and emerging diagnostic technologies.

Finally, in a global perspective, synthetic biology is a new approach to tackle life sciences. It concentrates tremendous open scientific questions of the 20th century, whose progress does not only provide an increased understanding of nature, but also new technological tools applicable to the living, including Humans and their health. The last decade has thus witnessed the rapid development of synthetic biology to full maturity. Fully grasping the biotechnology shift that is happening is, I believe, of the utmost importance to ensure the best human, knowledge and scientific progress as well as effective and fertile clinical translation.

Contents

Abstract.....	v
Acknowledgments.....	vi
Foreword.....	ix
Table of contents.....	xi
Chapter 1 Introduction	2
1.1. Engineering biology: the promises of Synthetic Biology.....	2
1.2. Synthetic biology as a medical technology	6
1.2.1. Engineering biology brings new insights in human physiology and health.....	6
1.2.2. Synthetic biology: Towards next-generation diagnostics	10
1.2.2.1. Overview and new considerations	10
1.2.2.2. How can synthetic biology be useful to medical diagnosis?	12
1.2.2.3. What has synthetic biology proven so far?	14
1.2.2.4. Concluding remarks.....	14
1.2.3. Fundamental Synthetic Biology approaches to diagnostic biosensor development	16
1.2.3.1. Top-down engineering of biosensors	16
1.2.3.1.1. Microbial systems.....	18
1.2.3.1.2. Eukaryotic systems.....	24
1.2.3.1.3. Viral systems	28
1.2.3.2. Bottom-up engineering of biosensors	31
1.2.3.2.1. Nucleic acid based systems.....	32
1.2.3.2.2. Protein based systems.....	39
1.2.3.2.3. High order assemblies and synthetic cell-like systems	49
1.3. Synthetic biology and biological signal processing for diagnostics: to sense and to compute	54
1.3.1. Synthetic biological circuits operating <i>in vivo</i>	57
1.3.2. Synthetic biological circuits operating <i>ex vivo</i>	60
1.4. Thesis statement: How can synthetic biology serve the engineering of next generation diagnostics?	62
1.5. Approach, summary of contributions and overview of the dissertation	63

Chapter 2 Engineering bacteria as a programmable and autonomous biosensing platform through integrase based synthetic gene circuits.....	68
2.1. Introduction	69
2.1.1. Bacterial biosensing.....	69
2.1.2. Synthetic biology & cell-based Biocomputing: integrating algorithms for decision making	70
2.1.3. Integrase based digital amplifying gene switches and Boolean logic gates for medical biosensing.....	70
2.2. Operational principles, design and architecture of batosensors	74
2.2.1. Behavior and robustness of bacterial chassis in human clinical samples	74
2.2.2. Design consideration of cell-based biosensors & rational improvement in robustness for the clinic.....	76
2.2.3. Multiplexing logic and memory in human clinical samples.....	78
2.2.4. Systematic methodology to achieve tunable biosensing coupled to programmable signal processing.....	80
2.2.5. Thresholding, digitization, and amplification of biologically relevant molecular signals using digital amplifying genetic switches	83
2.3. Developing polymer chemistries to immobilize batosensors within portable formats.....	84
2.4. Analytical evaluation of batosensors for the detection of biomarkers in clinical samples	86
2.4.1. Detection of a metabolized biological signal [glucose] in clinical samples	86
2.4.2. Batosensor mediated detection of pathological glycosuria in diabetic patients samples	89
2.4.3. Developing standards to quantify batosensor robustness in clinical samples	89
2.5. Engineering batosensor quorum sensing for intelligent sepsis diagnosis.....	91
2.5.1. Sepsis: clinical significance and diagnostic challenge	91
2.5.2. Batosensor mediated sepsis detection	92
2.5.3. Preliminary results	93
2.6. Conclusion and discussion	96
2.7. Materials and methods.....	99
2.8. Supplementary Materials	102

Chapter 3 Programming autonomous protocell biosensors via integrated synthetic biochemical circuits	114
3.1. Introduction	115
3.1.1. Protocells and bottom-up engineering of biological systems	116
3.1.2. Implementing biosensing and biocomputing operations using synthetic biochemical circuits.....	118
3.1.3. Computer assisted design and modeling for bottom-up synthetic biology	121
3.1.3.1. HSIM	124
3.1.3.2. BIOCHAM	126
3.1.4. Microfluidics for synthetic biology: methods for protocell fabrication	127
3.1.4.1. Introduction	127
3.1.4.2. Theory: Basic microfluidic concepts	129
3.2. Operational principles, design and architecture of protosensors	133
3.2.1. Architecture and general functioning	133
3.2.2. Programming <i>in vitro</i> algorithms for the differential diagnosis of acute diabetes complication	134
3.3. Molecular programming of protocells: from <i>in silico</i> design to experimental validation of synthetic biochemical circuits.....	135
3.3.1. <i>in silico</i> design, simulation and model checking.....	135
3.3.2. Experimental validation <i>in vitro</i>	139
3.4. Microfluidics approach for protosensors prototyping and fabrication: encapsulating synthetic circuits in protocells	141
3.4.1. Theoretical considerations	141
3.4.2. Experimental set-up and results	145
3.5. Construction and analytical evaluation of medical protosensors	148
3.5.1. <i>In silico</i> optimization of protosensor circuitry.....	148
3.5.2. Digital signal processing and multiplexing logic.....	149
3.5.3. Assaying pathological clinical samples: protosensor mediated diagnosis of Diabetes	153
3.6. Conclusion and discussion	154
3.7. Materials and methods.....	157
3.8. Supplementary materials.....	158

Chapter 4 Engineering universal protocell biocomputers.....	176
4.1. Introduction	177
4.2. Methods	179
4.2.1. Protocells as computation units: definitions	179
4.2.2. Circuit wiring.....	181
4.3. The case study	184
4.3.1. The Boolean satisfiability problem	184
4.3.2. The assembly of the machines.....	186
4.3.3. The computation process.....	189
4.4. Conclusion	190
 Chapter 5 Final Remarks.....	 193
5.1. Summary.....	193
5.2. Discussion and Perspectives.....	194
 Chapter 6 Bibliography	 200
 Chapter 7 Selection of recent advances in synthetic biology of interest to the field of diagnostics	 241
 Chapter 8 Annexes.....	 245

When one has managed to know the physical mechanism of production of an object or phenomena, (...) it is then possible (...) to produce the object of phenomena, science has become synthetic. Biology is a science like any other, (...) it needs to be successively descriptive, analytical and synthetic.

Stéphane Leduc, *La biologie Synthétique*, 1912

Up to now we are working on the descriptive phase of molecular biology. (...) But the real challenge will start when we enter the synthetic biology phase of research in our field. We will then devise new control elements and add these new modules to the existing genomes or build up wholly new genomes. This would be a field with the unlimited expansion potential and hardly any limitations to building “new better control circuits” and (...) finally other “synthetic” organisms (...).

Waclaw Szybalski, *Control of gene expression*, 1974

The work on restriction nucleases not only permits us easily to construct recombinant DNA molecules and to analyze individual genes but also has led us into the new era of ‘synthetic biology’ where not only existing genes are described and analyzed but also new gene arrangements can be constructed and evaluated.

Waclaw Szybalski & Anna Skalka, *Nobel prizes and restriction enzymes*, 1978

Chapter 1

Introduction

The present chapter is intended to give an overview of medical synthetic biology, and more particularly focusing on emerging approaches for the engineering of innovative diagnostic devices. The aim of this Chapter is to allow the reader to familiarize with the major concepts and advances in the field that will help to put into perspective the work discussed and analyzed in the following chapters.

1.1. Engineering biology: the promises of Synthetic Biology

The last decades can be regarded as the descriptive phase of molecular biology and functional genomic research, which in the last two decades permitted the advent of synthetic biology^{1 2 3 4}. Although still widely debated, and while official and non-official definitions have been given, the hallmark of synthetic biology is probably the common interest to build biological entities that do not exist in nature yet. According to the High-level Expert Group from the European Commission, the following definition best applies:

Synthetic biology is the engineering of biology: the synthesis of complex, biologically based (or inspired) systems which display functions that do not exist in nature. This engineering perspective may be applied at all levels of the hierarchy of biological structures – from individual molecules to whole cells, tissues and organisms. In essence, synthetic biology will enable the design of ‘biological systems’ in a rational and systematic way.⁵

Said differently, synthetic biology applies the classical engineering strategies to the biological substrate: standardization (compatibility and exchangeability of components), decoupling (dissecting a complex problem or systems into multiple subtasks), and abstraction (streamlining

to address relevant facets of a problem). Another key concept involves modularity, which refers to the capacity of systems components to be separated and recombined, or the capacity to construct large systems by combining basic orthogonal elements. This approach is extensively employed to rapidly engineer scalable and cost-effective devices and structures.

Therefore, synthetic biology has become a science of designing biological components, devices, systems and organisms in a systematic and rational manner to create predictable, useful and novel biological functions. Systematically defining, cataloguing, engineering and standardizing large sets of modular biomolecular components based on always increasing amount of data, in easy accessible databases provide well-characterized standard biological parts enabling hierarchical abstraction of biological functions^{4 6 7 8}. These abstract components, or biological *parts*, can then be assembled at the systems level to construct new biological systems with user-defined functionalities (**Figure 1**), capitalizing on accessible combinatorial diversity. In synergy with major technological improvements, synthetic devices and systems can be easily designed and simulated *in silico*, before synthesis, transfer, and assembly in complex systems, according to an iterative explore-standardize-build process. The most notable achievement in this perspective is probably in the management of the registry of standard composable genetic parts, or BioBricks⁹ which by classifying and sharing promoters, coding sequences, ribosome binding sites or other parts, has for more than a decade supported the exponential growth in capabilities of synthetic biology.

As a mature engineering discipline, the true limits lie in the *read* and *write* capacities, or in other words in the available technology to manipulate biological structures. Interestingly, synthetic biology first crystallized around key enabling technologies, namely DNA synthesis and sequencing, boosted by computer modeling. The recent advances in synthetic genomics (i.e. or the capacity to now write and read megabase-scale information stored in nucleic acid polymers) have permitted to gain control on living organisms with unprecedented precision. Amongst a large example of impressive achievements, it is notable that this approach recently enabled the construction of a whole synthetic yeast chromosome¹⁰, or the first bacteria species living with a synthetic genome engineered through chemical synthesis of a computer assisted reduction of a natural *Mycoplasma* genome¹¹. The high throughput and high precision edition and programming capabilities on biology is also particularly relevant in technologies such as multiplex-automated genome engineering (MAGE)¹² for programming cells via accelerated evolution, the recent Cas9-CRISPR technology for universal genome editing¹³, or the success of optogenetics¹⁴ for remote *in vivo* light control of phenotypes and genotypes.

In the emergence process, systems biology's comprehensive perspective on biology played a key role¹⁵. It led to view natural processes as systems of interacting components, where synthetic biology could play the complementary approach: building and testing the systems. This view led biological processes to be increasingly understood through the prism of information processing. In other words, synthetic biology sees the engineering of life at the interplay of biology and information technology. This has prompted the use of genetic modules and synthetic gene circuits and DNA as information storage, illustrated in the early success using transcriptional circuits in the model chassis bacterium *Escherichia coli* around the toggle switch¹⁶, oscillators¹⁷, counters¹⁸, cell-cell communication¹⁹, or basic Boolean logic gates²⁰. These advances have lately

been considerably augmented in terms of scale, with for instance large information storage systems²¹, layered biomolecular logic gate circuits²², DNA biocomputing that can mimick brain-like behavior²³, and now proves of concept of analog and digital domain signal processing across all kingdoms of life²⁴. Available substrates to support rationally programmed circuitry is now evolving towards non-transcriptional RNA based and post-translational elements, to support novel modalities of biological control. Additionally, while evolution was traditionally considered a hurdle to biotechnologies, recent approaches propose to exploit this intrinsic property of living organisms²⁵. Interestingly, these approaches bring formidable insights and understanding of fundamental natural design principles from the building of artificial ones²⁶.

While the top-down engineering of life can now almost be seen as straightforward and reliable, there still remains a long way to go before mastering bottom-up construction of synthetic biosystems. In other words, we are still much better at repurposing naturally occurring living systems than rationally building *de novo*. However, both approaches contemplate the same ultimate goal which is to realize fully orthogonal biology. As it happens, the first steps towards this goal, implying unnatural nucleotides and amino-acids and their polymers (i.e. Xeno Nucleic Acids and synthetic enzymes)^{27 28 29 30}, are already underway. Likewise, the first minimal synthetic replicative systems have been described^{31 32 33}.

Synthetic biology thus provides a method for systematic and rational assembly of synthetic parts into on-purpose systems, and as such can be defined as the science of structuring biological matter to achieve control on biological energy and information processing. It has already proven of outstanding capabilities towards applications, for instance in bioremediation³⁴, or foods³⁵, but mostly industrial microbiological bioproduction³⁶ through metabolic pathways engineering for the production of biofuels and high value biomaterials and biomolecules³⁷, and of course, biomedicine^{1 38}.

Although most synthetic biology labs do not focus their efforts on biomedical applications, fundamental advances in the design of new medically oriented molecular devices are symptomatic of the growth of biomedical synthetic biology. Traditional biotechnologies attempting to bridge the gap between research and patient clinical care have too often been burdened with issues of reproducibility and standardization. We suspect that these typical hurdles can be addressed by synthetic biology methodologies to allow safe, robust and reliable clinical applications. Maturation of the field and technological development indeed enhances our ability to study and control biosynthetic systems to be used for health applications^{39 40 41}. It is now slowly transitioning into the clinics and has already yielded successful applications, for example vaccine development and production⁴², unprecedented synthesis strategies for high value drugs^{43 44 45 46 47} such as artemisinin⁴⁸, synthetic opioids⁴⁹, or novel biotherapeutics^{50 51}, high value synthetic medical biomaterials⁵², gene delivery⁵³, control of parasite vectors⁵⁴, or a vast range of proof of concept *smart* cells for therapeutic purposes^{55 56 57 58 59 60}.

Nevertheless, the progress seen during the last decade to augment the synthetic biology *toolbox*, with for instance the modular resource of the Biobrick repository, has mostly been insufficient to provide actual real world tools via a promised *plug-and-play* strategy. Tedious trial and error steps, fine-tuning and extensive human supervision are often required to engineer working biosynthetic systems. Consequently, applications in the medical field remain limited, synthetic

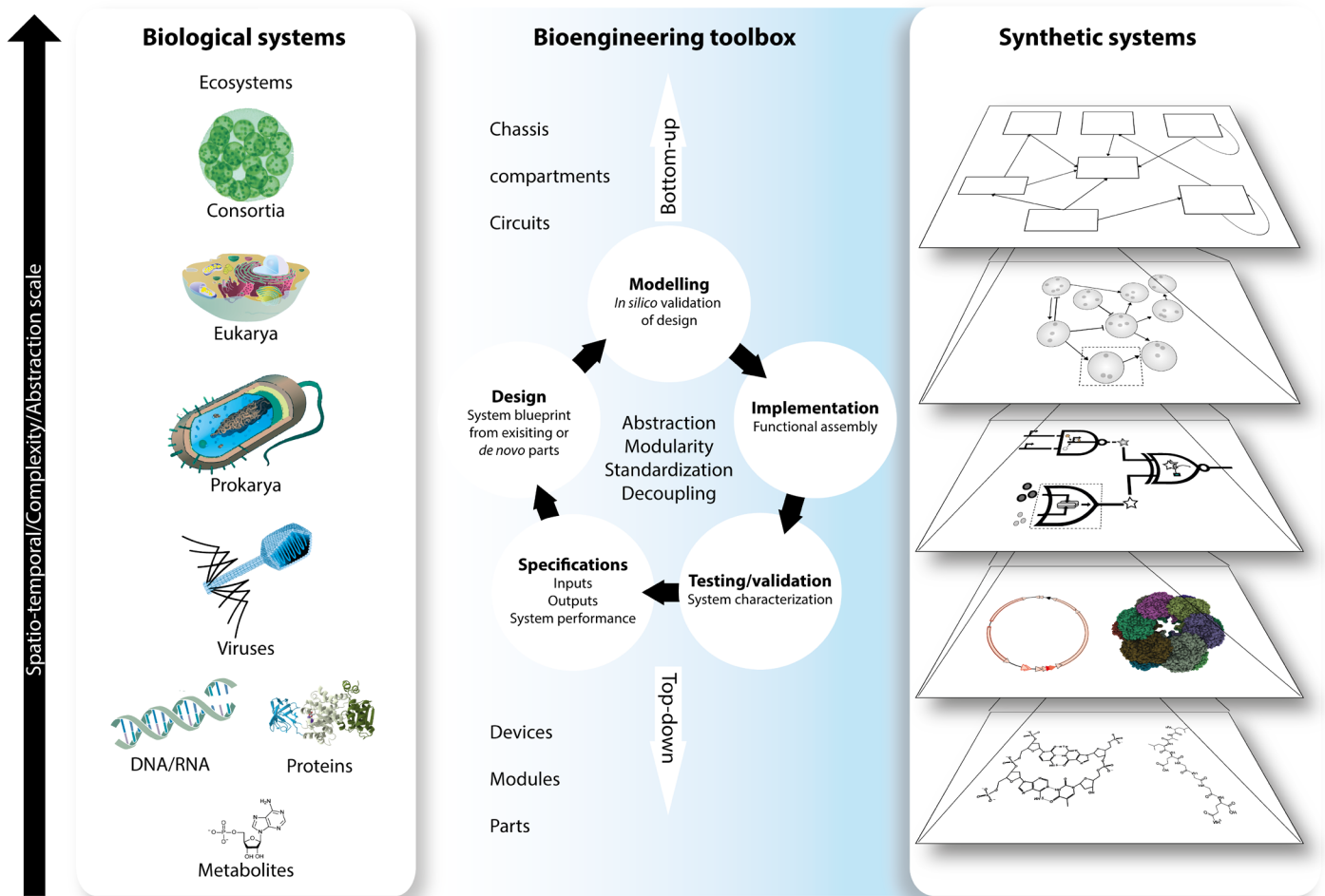


Figure 1: General framework that incorporates top-down and bottom-up perspectives in the synthetic biology design process. The increasing knowledge of biological systems, their deconstruction, and the design of synthetic systems across different levels of complexity, is an iterative process that incorporates both top-down and bottom-up design considerations. The study of biological systems enables the accumulation of increasing amount of data that feed a *bioengineering toolbox* with standardized elements such as chassis, compartments, circuits, devices, modules and parts. Systematic engineering methods and mathematical tools constitute the conceptual framework by which synthetic biology operates. First, a design objective is formulated, taking into account functional constraints and specifications in terms of systems performance. Then, a synthetic biological system is designed by composing with well-characterized components with known properties, either *ab initio* (bottom-up) and/or combined with a larger biological context (top-down) and modeled *in silico* to identify potential modes of failures. The synthetic system is then constructed experimentally (implemented) and performance of the system is assessed. If the system fails to meet performance requirements, this new information can be used to refine the design and iterate the design process. This process constantly improves understanding of biology and reduces the number of iterations necessary to achieve a specific design objective. Knowledge based design infuses each levels of the abstract hierarchical scale of synthetic systems, which are in that sense parallel to “natural” biological systems.

biology faces challenges toward human clinical applications^{56 61} and most research tools have yet to reach clinical trial. Identifying and developing universal approaches, biological devices and methodologies that would ease the way towards the clinic appear necessary. However, the research landscape is moving, and specific applications are becoming realities. In particular, and as we will discuss in detail in the following sections, synthetic biology could offer a drastically new approach towards applications in medical diagnosis (**Figure 2**).

Last but not least, synthetic biology, as a novel perspective on life sciences, also encompasses important aspect of human sciences, arts and politics, and raises new attractive economic opportunities. This aspect is of importance if one is to envisage industrial and health application in order to anticipate misusages and address growing ethical and biosafety concerns.

1.2. Synthetic biology as a medical technology

A physician's subject of study is necessarily the patient, and his first field for observation is the hospital. But if clinical observation teaches him to know the form and course of diseases, it cannot suffice to make him understand their nature; to this end he must penetrate into the body to find which of the internal parts are injured in their functions. That is why dissection of cadavers and microscopic study of diseases were soon added to clinical observation. But to-day these various methods no longer suffice; we must push investigation further and, in analyzing the elementary phenomena of organic bodies, must compare normal with abnormal states. We showed elsewhere how incapable is anatomy alone to take account of vital phenomena, and we saw that we must add study of all physico-chemical conditions which contribute necessary elements to normal or pathological manifestations of life. This simple suggestion already makes us feel that the laboratory of a physiologist-physician must be the most complicated of all laboratories, because he has to experiment with phenomena of life which are the most complex of all natural phenomena.

Claude Bernard, *An Introduction to the Study of Experimental Medicine* (1865)

1.2.1 Engineering biology brings new insights in human physiology and health

A considerable need exists for improving understanding of diseases, along with the discovery of biomarkers for differential diagnosis, prognosis of diseases and monitoring of therapeutic interventions. Different strategies have thus been pursued to get insights on molecular pathophysiology, to unveil mechanisms and potential therapeutic targets, but also to discover predictive biomarkers of pathology development. Feynman's famous quote, *What I cannot create, I do not understand*, in other words analysis-by-synthesis⁶², take in this context all their meaning. From a medical science perspective, the more we tinker with biology, the more we gain understanding of the complex behavior and organization of biological systems, the more likely we will be to engineer operational medical devices. Hence, synthetic biology represents a powerful approach towards new models and tools to explore and probe pathophenotypes. The rational and systematic *reverse* engineering of biosynthetic pathways, biological parts, synthetic genes and networks constitute valuable resources for the multi-level screening of disease mechanisms. It allows the iterative design and *in vivo* implementation of quantitative and dynamic models to test molecular hypotheses, and to perturb and probe biological networks topologies^{63 64}.

For instance, Yagi *et al.* recently shed new insights on breast cancer pathogenesis and approaches to diagnosis using a synthetic biology strategy to reconstitute G protein-regulated networks in breast cancer cells. They stably expressed an engineered G α_i -coupled GPCR, which had gained the ability to respond to a synthetic agonist, enabling them to probe the signaling pathways downstream of specific G proteins⁶⁶.

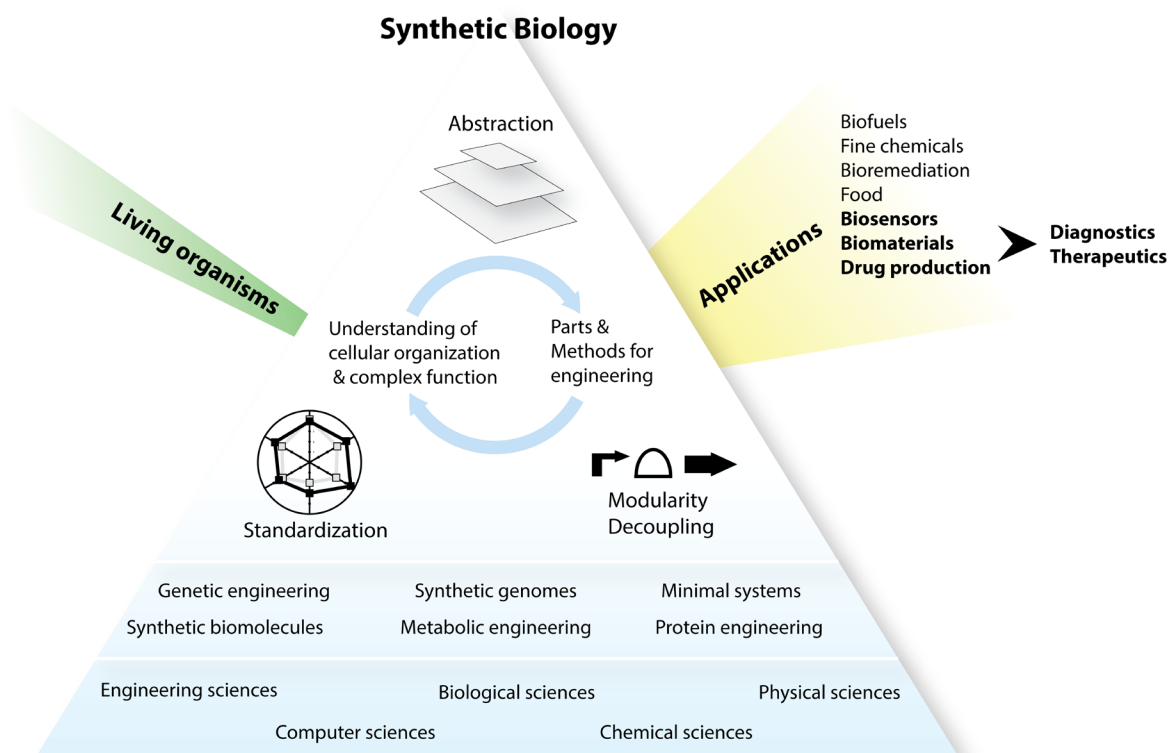


Figure 2: Synthetic biology generates knowledge and tools for technology and medicine. The vast array of interdisciplinary methods and substrates that can be manipulated via synthetic biology enables the engineering of biological systems to gain new insights on governing principles of biological structures, as well as to develop various applications with increased design space, which can be targeted towards human health and diagnostics.

Synthetic biology recently also enabled the systematic synthesis of whole pathogens such as SARS or Influenza viruses^{67 68} or their components through complex DNA-gene synthesis and whole genome assembly techniques^{11 69}. This methodology offered fast access with low efforts to address pathogenicity mechanisms and provided new diagnostic targets. Novel immunoassays, as well as DNA arrays were developed for known or potential pathogens and newly described infectious agents^{70 71 72}. For example, gene synthesis has recently been translated to clinical diagnosis with the discovery of Merkel Cell Polyomavirus and its association with Merkel cell carcinoma, a rare human skin cancer⁷³. Systematic gene synthesis also enables synthetic codon sequence optimization of genes, and enhanced the production of multi-epitope and chimeric antigens. Synthetic biology enables simplified screening and improved diagnostic performance via standardized and robust antigens, thus reducing assay variability and achieving high levels of sensitivity and specificity in serologic immunoassays of infectious agents⁷⁴ or autoimmune diseases⁷⁵. These strategies have been used to mimic specific epitopes from pathogens in many diagnostic systems. For example, a synthetic protein combining four different immunodominant epitopes from *Borrelia burgdorferi* generated an improved serological test for the diagnosis and monitoring of Lyme disease⁷⁶. In the same way they provide more sensitive methods for

detecting patient antibodies in diagnostic immunoassays, peptide synthesis through multi-epitope and chimeric genes can be valuable for the direct identification of new autoantigens⁷⁷. A method relying on synthetic representation of the proteome using phage display combined with high-throughput sequencing permitted to identify novel autoantigens in neurological syndromes⁷⁸. These synthetic approaches have also recently yielded comprehensive insights in human viral immunology. Xu *et al.* recently developed a high-throughput method to exhaustively explore the human virome relying on massively parallel DNA sequencing of a bacteriophage library displaying proteome-wide peptides from all human viruses⁷⁹ (**Figure 3A**). These advancements now allow for the high-throughput identification and quantification of unique antibodies biomarker of patient's serological repertoire, but also bring unprecedented progress towards unraveling mechanisms and dynamics of adaptive humoral immunity, serological memory and response for emerging infections and vaccine development, as well as therapeutic antibody discovery⁸⁰. Taking these approaches further, the capability to rationally engineer the immune response, has recently been proposed to treat immune disorders in humans⁸¹. This approach coined synthetic immunology proposes to repurpose immune cell effector functions through the use of synthetic proteins or engineered immune cells, and is already transitioning to preclinical and clinical trials.

While clinical management of complex diseases is increasingly relying on biomarkers, our ability to discover relevant ones remains limited by our dependence on endogenous molecules. The lack of specific, predictive or robust biomarkers still limits the diagnosis of many pathologies. Thus, recent attention has been given to the engineering of disease specific synthetic biomarkers. These exogenous agents are administered in the circulatory system where they record molecular events associated with pathological states. As such, they enable the non-invasive monitoring of non-classical parameters by producing new molecular signatures that can then be retrieved in clinical samples such as blood or urine. Several teams recently developed protease-sensitive biomarkers that respond to pathological enzymatic activities at diseases sites, and release reporters in circulation that are then concentrated in hosts' urine to be measured. The potential for early disease stage detection and monitoring compared to classical blood biomarkers has been reported with murine models of liver fibrosis, cancer and solid tumors, or cardiovascular diseases^{82 83 84 85} (**Figure 3B**). These preliminary studies are important steps toward use of injectable synthetic biomarkers in the clinic, and could be generalized in a multiplex diagnostic platform and tailored for the diagnosis of various diseases.

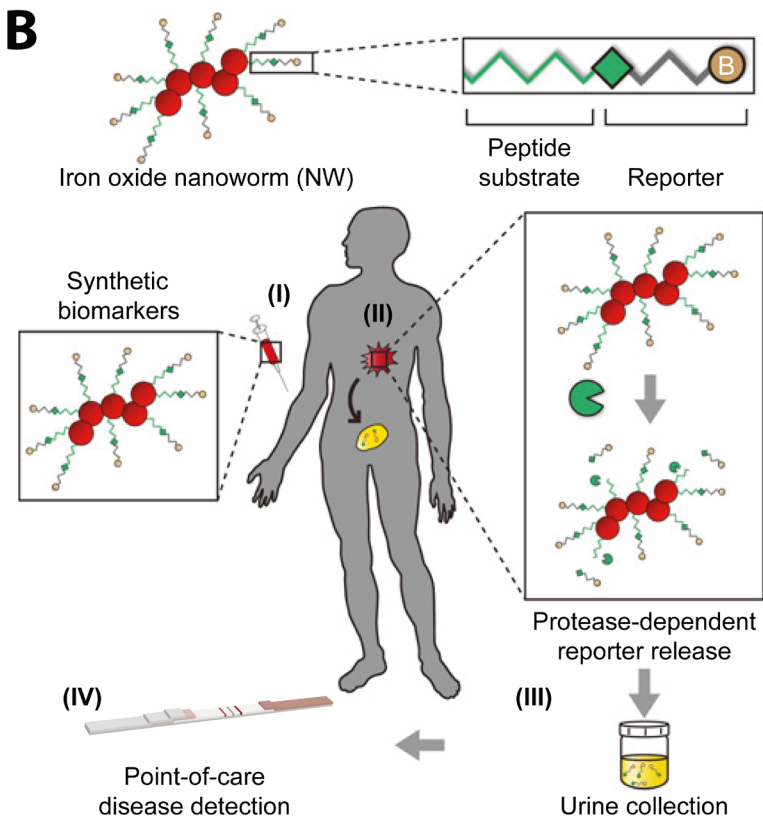
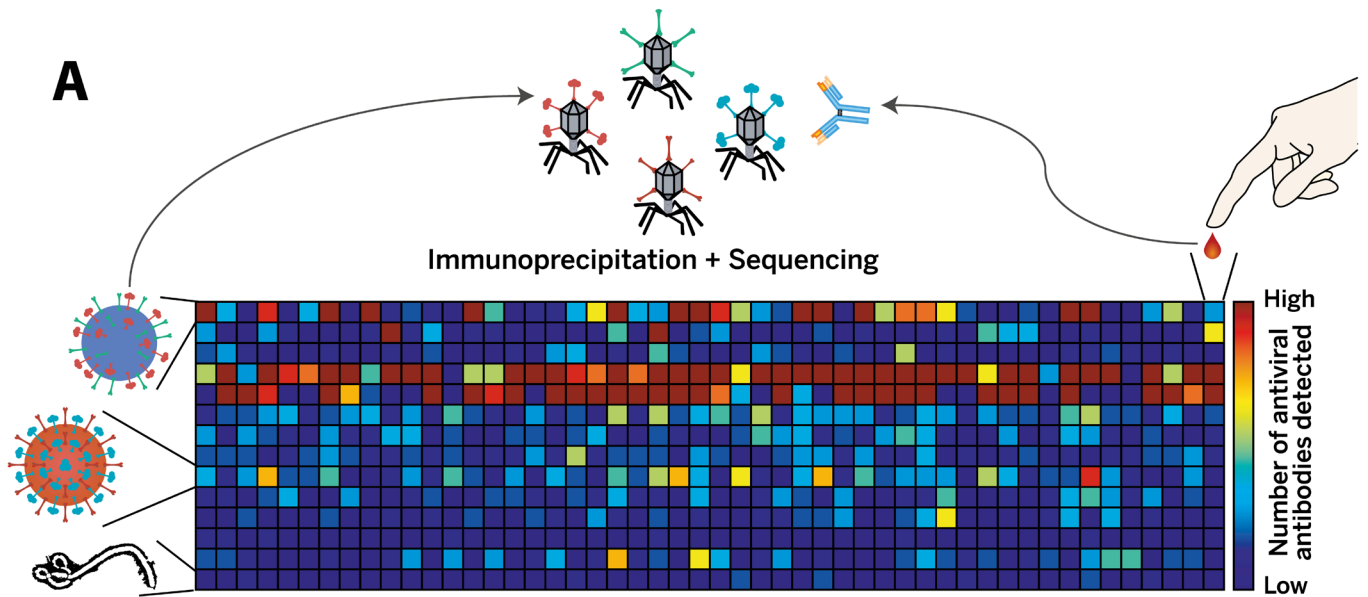


Figure 3: Synthetic biology to explore human pathophysiology. (A) Systematic viral epitope scanning. This method allows near exhaustive analysis of antiviral antibodies in human sera from 1 μ l of blood, through the coupling of DNA microarray synthesis and bacteriophage display to create a uniform synthetic representation of peptide epitopes of the human virome. The hitmaps depicts the relative number of antigenic epitopes for a virus (Adapted from Xu *et al.*⁷⁹) (B) Synthetic biomarkers for non-classical monitoring of diseases. Protease sensitive synthetic bionanoparticles are synthesized by conjugating substrate-reporter tandem peptides to a carrier nanoworm particule. Proteolytic cleavage of the linking peptide substrate at the disease site liberates ligand-encoded reporters that filter into urine. (Adapted from Warren *et al.*⁸².)

1.2.2 Synthetic biology: Towards next-generation diagnostics

The section below was adapted from *Bringing next-generation diagnostics to the clinic: fulfilling the promise of synthetic biology* by Courbet *et al.*, to be published in *EMBO Molecular Medicine* (under review)

Diagnostic applications have recently attracted great interest from synthetic biologists. In order to achieve highest medical impact, we believe it is now time to communicate to a general audience of scientists and medical practitioners about the diagnostic capabilities offered by synthetic biology and trigger interest, commitment, and translational dialogue. This section is therefore intended to be of broad biomedical interest, from fundamental researchers and clinical biologists to medical specialists. Raising awareness, increasing understanding and knowledge of emerging diagnostic technologies for this particular audience is critical, and we therefore believe it is important to highlight these new insights.

Here, we envision that synthetic biology's most imminent medical impact is in the revolution of diagnostics, and its relation to personalized and precision medicine through point-of-care and companion diagnostics. The aim of this section is (i) to demonstrate the importance of present and future synthetic biology approaches to medical diagnosis (ii) to map the landscape of novel biodiagnostic strategies and technologies emerging from synthetic biology (iii) to propose future orientations that could accommodate medical, socio-economical, industrial and legal requirements.

1.2.2.1 Overview and new considerations

Developing high clinical value diagnostics remains a major technological problem of medical sciences. Financial imperatives, health facility resources, as well as medical information management, geographic and economic misdistribution further exacerbate a context of increasing life expectancy and risk factors⁸⁶. Medical solutions offering non-invasive and systematic screening of populations at risk in resource-limited settings, diagnosis at the patient bedside and close monitoring are thus of first importance⁸⁷.

The ever-increasing understanding of biological systems, as well as medical care evolution towards personalized solutions also place evolving imperatives on medical bioanalytical technologies⁸⁸. At the same time, convergence of precision medicine, diagnosis and therapy led to the development of personalized medicine, point-of-care, companion diagnostics, and theranostics. Realizing these novel approaches would benefit the individual as well as society, improving therapeutic outcome and reducing healthcare costs, while also benefiting regions with poor infrastructure.

So far, centralization of conventional *in vitro* diagnostics in clinical laboratories has been required in order to match modern standards, achieve specific, sensitive, multiplexed and high-throughput measurements, or generate results with high robustness and reliability. However, these diagnostic tools are either non-portable, high maintenance and costly, or are restricted to the detection of single biomarkers with mostly low sensitivity and specificity. These molecular signatures can be

of various biochemical natures ranging from genetic and epigenetic markers to changes in complex evolution of proteome, genome or metabolome. Since individual biomarkers are limited in providing optimal diagnostic sensitivity and specificity, they cannot accurately account for complex molecular pathophenotypes and testing for multiple biomarkers at once can thus save time and resources while improving diagnostic accuracy⁸⁹. Furthermore, the current organization of medical diagnosis remains incapable of accommodating emerging clinical need for real precision diagnostics, promising finest monitoring and intervention on pathophysiology.

Consequently, the last decade witnessed important efforts to identify predictive biomarkers of diseases, as well as an interest for innovative diagnostic technologies providing with rapidity, versatility, robustness, easiness-to-use, portability, and last but not least, cost-effectiveness⁹⁰. Yet, novel diagnostics capable of decentralizing the biochemistry lab to the patient without scarifying medical service are still to emerge.

To achieve highest value, novel diagnostic devices would perform autonomous biodetection of pathological biomarkers with high sensitivity, specificity, robustness, rapidity, and possibility of direct analysis in complex matrices without sample pretreatment. Moreover, future diagnostics could be implantable and passive, wirelessly connected to the clinician, while providing with near real time or continuous measurements. Diagnostic information about new types of biological parameters could be processed at the micro/nanoscale regarding clinically relevant intercellular, cellular, and subcellular events. Precise detection of biomolecules in close relevance to human pathophysiology could be achieved via the scoring of transient and weak signals, as well as integrated signal processing to enable complex measurements. Biosensor approaches integrated with information technologies and biological/electronic interfaces are likely to support novel solutions⁹¹.

Yet, the engineering of such integrated, stand-alone *expert* biosensing device as medical decision support remains a challenge. Innovative and robust methods to engineer biosensing systems are thus of tremendous importance.

We propose that this could be supported by the emerging biotechnological field of synthetic biology. Benefiting from a constantly increasing capability to systematically inform and interface biology, we envision that synthetic biology's key capabilities will serve and accelerate the engineering of a generation of diagnostics with novel capabilities (**Figure 4**).

Although most synthetic biologists do not focus their efforts on biomedical applications, fundamental advances, maturation of the field and technological development recently enhanced our ability to control biosynthetic systems to be used for human health^{92,93}. It has already yielded successful medical applications, for example for the production of complex drugs^{43,44}, high value synthetic medical biomaterials⁵², or proof of concept *smart* cell therapeutics⁹⁴. Although applications in the medical field remain in their infancy, the field has been under extensive investigation and key achievements have announced translation of synthetic biology into clinical biosensing.

1.2.2.2 How can synthetic biology be useful to medical diagnosis?

Firstly, the field has demonstrated extensive value to explore human pathophysiology and served for the discovery of new predictive biomarkers. Indeed, as we have previously discussed, a considerable need exists for improving understanding of diseases. In that perspective, synthetic biology represents a powerful approach towards new tools to explore and probe pathophenotypes. For instance, synthetic biology enabled simplified screening and improved assays diagnostic performance via standardized and robust antigens⁷⁵, synthetic representation of the proteome and human virome^{78,79}, synthetic immunology⁸¹, or the use of synthetic *intelligent* biomarkers⁸³.

Secondly, since biological systems naturally behave as biosensors integrating various kinds of biochemical signals, they can be reprogrammed to support diagnostic operations. They can thus assess molecular pathophysiology by biorecognition of biomarker patterns and respond conditionally with phenotypic readouts. In addition, biological systems have interesting characteristics for diagnostics, such as the ability to perform ultrasensitive and specific response to stimuli. They are also autonomous, self-powered, miniaturizable, auto-organized, amenable for high throughput and can function in complex biological contexts at different scales. Moreover, synthetic biology provides tools and methods to mine and efficiently re-engineer the vast repertoire generated by evolution, in order to retrieve useful biological functions for clinical biosensing.

Last but not least, biological systems are efficient problem-solving systems that rely on biological signal processing modules. Synthetic biosensor systems can thus be designed to exploit a modular architecture with high composability, in which 3 modules are exchangeable: sensor, processor, and reporter (**Figure 4**). The rationale behind biological integrated signal processing for diagnosis is to achieve re-programmability, multiplexing of pathological signals, to provide with quantitative, semi-quantitative or qualitative measurements, as well as amplification, thresholding, noise filtering, or logic operations. While trading a simple for a more complex design would be counterproductive, modular and standardized interface between sensing and reporter components can speed up the design, and increase versatility and capabilities. Moreover, the signal sensing event of biosensors can be associated to a computation process that integrates compiled medical knowledge in the form of a decision algorithm. Biological circuitry can thus be easily reprogrammed to integrate varying clinical constraints, different medical agendas and a vast range of pathologies. Synthetic biology thus supports the integration of tailored signal detection, processing and reporting by means of modular construction. Biological systems can thus be systematically repurposed into autonomous devices that assess diagnostic rules *in situ*.

In fact, the diagnostic process attempts to classify patient conditions into distinct clinical categories that support medical decisions regarding treatment and prognosis⁹⁵. Medical diagnosis can thus be regarded as a logical problem, or an elementary computation process leading to medical decision making. In other words, the pathophysiological state of a patient is a function of molecular patterns of disease associated biomarkers with can be computed. Since medical

diagnosis identifies with the process of making decisions about the state of human physiology, and biological systems can implement the logical operations of medical diagnosis, it is possible to exploit biological systems for diagnostic applications. In this perspective, synthetic biology enables the full integration into operational diagnostic devices (**Figure 5**).

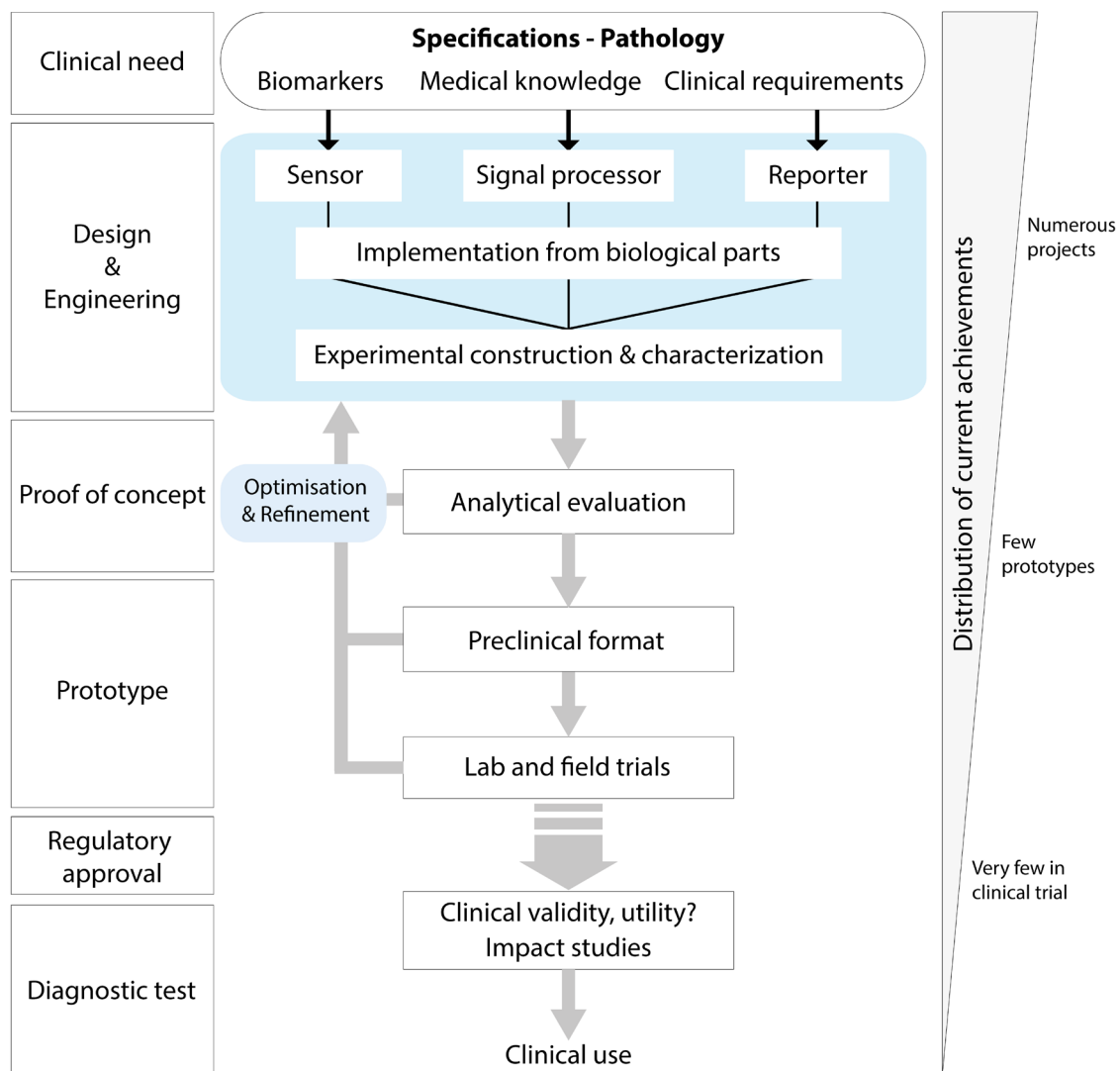


Figure 4: Systematic workflow for the development of diagnostic devices using synthetic biology, from bioengineering considerations to clinical use. A bioengineering solution can be formulated corresponding to certain clinical specifications as a technological support to disease diagnosis. Such specifications can be systematically implemented within biological substrate using standardized biological parts. An iterative process between analytical properties of engineered diagnostic systems, design and construction optimizes the process to eventually lead to effective approval and clinical use.

1.2.2.3 What has synthetic biology proven so far?

This framework has been recently successfully applied to engineer diagnostics such as cell-based biosensors screening patient's urine or blood for metabolic disease⁹⁶, performing precise profiling of allergies in human whole-blood samples⁹⁷, or serving as sentinel probiotics *in vivo* to detect the onset of liver cancerogenesis⁹⁸. Other biological systems such as bacteriophages have been engineered into accurate microbial diagnostics providing with near-real-time detection in clinical samples and antibiotic susceptibility testing⁹⁹. Alternatively, proof of concept devices made of networked synthetic nucleic acids or proteins have been recently developed, such as autonomous DNA machines for the *in vivo* multiplexed analysis of cancer cell-surface markers^{100,101}, logic circuits for identification of specific cancer cells¹⁰², programmable paper based synthetic gene networks for strain specific Ebola diagnostics¹⁰³, or protein sensors for inexpensive point-of-care companion diagnostics for various narrow therapeutic range drugs¹⁰⁴.

By nature biologically compatible and interfaceable, synthetic biological systems offer the possibility to develop implantable devices sensing pathological stimuli *in situ* that immediately offer a therapeutic response (i.e. theranostics). This approach proves extremely valuable in many clinical situations where therapeutic benefit depends on the delay in analytical methods, clinical information management and interpretations, and effective patient care. For instance, engineered T cells have already been tested in clinical trials for cancer theranostics¹⁰⁵ and could be further improved using synthetic biology strategies in terms of safety and capabilities⁵⁵, and proof of concept theranostic synthetic circuits in mammalian cells were engineered to perform urate homeostasis control, or to monitor and manage diabetes and metabolic disorders¹⁰⁶.

1.2.2.4 Concluding remarks

Synthetic biology has grown and advanced enormously in the past few years, providing with robust methods allowing the assembly of engineered molecular and cellular devices with biosensing and information-processing capabilities. For these reasons, diagnostic applications have recently attracted great interest from synthetic biologists. Researchers and clinicians could now begin to translate the bioengineering framework into the medical field, to ultimately realize *intelligent*, autonomous, and programmable diagnostics. Synthetic biology could serve as a methodology and technological support to help interface medical biology, through more precise, personalized, and sophisticated diagnostic tools increasing versatility and portability and simplifying medical decision rules. It provides with an opportunity to decrease the scale of diagnostic devices, getting closer to the patient and giving access to real-time, personalized and physiologically meaningful diagnostic measurements. We envision that in the future synthetic biomolecular devices will make precise decisions leading to enhanced medical care and therapy (**Table 1**). The prospect for synthetic devices to act as self-contained diagnostics is almost established, and could evolve toward multipurpose implantable system for *in vivo* theranostics.

In this perspective, we envision that synthetic biology’s most imminent medical impact is in diagnostics. We suspect that these advances are likely to announce a profound technological evolution similar to next-generation sequencing technologies or antibodies in the development of immunosensors. The relation to personalized medicine and therapy through point-of-care, companion diagnostics and theranostics will most likely bring evolution to the patient as well as into the clinical laboratory. In addition, while the field of mobile health and point-of-care is rapidly growing and likely to become widespread reality through the use for example of connected devices¹⁰⁷ (or even maybe more provocative *mind-controlled* designer synthetic devices¹⁰⁸), new supports may be required to achieve full potential, and synthetic biological systems stand as promising alternatives.

	Conventional diagnostics	Synthetic biology enabled diagnostics
Diagnostic procedure	Centralized clinical biochemistry lab, high resource requirements	Ambulatory, close to patient, potentially <i>in vivo</i> , low resource
Sample management	Pre-treatment, large volumes	Raw, small volumes
Nature of biomarkers	Parallelized, static, disconnected from patient pathophysiology	Multiplexed, dynamic, <i>in situ</i> , close to patient pathophysiology
Data transmission	Delayed, complex interfaces	Real-time, integrated signal processing, local/remote readout
Link to therapy	Delayed, through physician evaluation	Direct, <i>in situ</i> , through programmable decision algorithms: remote supervision
Data management	Files, registries	Embedded memory
Medical benefit	Robustness, gold standard	Patient comfort/care, personalized solutions, patient commitment

Table 1: Conceptual differences in medical procedures between conventional versus synthetic biology enabled diagnostics.

1.2.3 Fundamental synthetic biology approaches to diagnostic biosensor development

The framework considering biological entities as systems of interacting components capable of input detection, information processing, executing logical operations, and producing an output⁶, has led to the engineering of *intelligent* systems for biodetection purposes and as such be used for diagnostics applications. Such systems can be developed from the top-down perspective using modular biological parts assembled *in vivo* to generate useful synthetic phenotypes, or be assembled *ex-vivo* from a bottom-up perspectives. In this section, we propose to review the advances in synthetic biology that proved of outstanding relevance to the field of diagnostics.

1.2.3.1 Top-down engineering of biosensors

The engineering of cell-based biosensing system has arisen as a major focus in the field of synthetic biology¹⁰⁹, and proved to be useful as a versatile and widely applicable method for detection and characterization of a wide range of analytes in biomedical analysis^{1 110 111 112 113}. These systems are capable of producing dose-dependent detectable signals in response to the presence of specific analytes in a given sample. However, the first generation of cell-based biosensors mostly relied on native cellular sensor modules without extra signal processing abilities, and thus can only detect isolated signals with low signal to noise ratio and poor robustness when used in complex matrices¹¹⁴.

Consequently, synthetic biology efforts have focused on streamlining the construction of robust cell-based biosensors for biomedical applications. A wide range of modules have emerged through genetic engineering, and enhanced these systems in terms of modulation of sensitivity, specificity and dynamic range, near-real-time signal processing, multi-input (multiplexing) and logic operations, or toward the integration of orthogonal biological and electronic components^{115 116 117 118 119}. Cell based biosensors capable of multiplexing detection enable to classify complex conditions specified by combination of several signals. Many proofs of concept have highlighted the great advantage of *in vivo* integration of algorithm using biological logic circuits, in order to customize cell sensing and signaling into decision making systems, to be used for various clinical applications. In this way, sensor/reporter modules can be interfaced with fine signal processing such as digital logic and memory (see Section 1.4) carried out *in vivo* by synthetic gene networks. This strategy can be used to enhance sensing specificity and accuracy of the output response. In addition, engineering frameworks have been developed for the optimization of cell-based biosensors, such as directed evolution through MAGE or phage assisted continuous evolution. Even though synthetic gene circuits have been used for a decade to construct cells that respond to biological signals in a programmable fashion^{115 120 121}, current commercially available or proof of concept cell biosensors have so far been mostly used in contexts irrelevant of medical applications¹²². In this perspective, I suspect advances in synthetic biology will enable a new era of robust, stand-alone and integrated smart biosensing devices for medical diagnosis.

These biosensing devices are supported by a chassis, or host cell, which supplies necessary resources for functionality. The engineering of cell-based biosensor devices has been conducted in different cellular chassis, either plant¹²³, algae¹²⁴, mammal¹²⁵⁻¹²⁶, yeast¹²⁷, and a wide spectra of bacteria species¹²⁸. They have been widely investigated for environmental and medical diagnosis because they enable cheap and simple large-scale field screening and measurements. However, they have other properties that make them interesting as diagnostic devices. They are relatively easy and inexpensive to prepare and store through cell culture, require low-cost reagents, and have evolved increased stability compared to biochemical probes (e.g. DNA, proteins...) when exposed to perturbation (temperature, pH, ionic strength...). Moreover, cell-based assays are non-destructive, and provide more comprehensive and complex functional and physiological information than classical analytical methods, such as bioavailability¹¹⁰⁻¹²¹. They can provide insights into the pathogenic mechanisms, potentially giving estimation of clinical risks associated with specific molecular events¹²⁹⁻¹³⁰. Because of the auto-replication of biological systems and self-powering a cell-based diagnostics system could be portable, and have reduced production costs compatible with systematic and widespread deployment. Last but not least, cell-based sensor systems could be implanted directly *in vivo*, thus permitting noninvasive detection of conditions in live organisms over time, which could prove particularly powerful for diagnostic applications.

Additionally, cell-based biosensors can be further integrated into high density devices to perform high through-put analyses and are amenable for miniaturization and incorporation into portable, μ TAS devices¹³¹⁻¹³⁵. In fact, micro-engineering, bioelectronics and microfluidic strategies enable the use of population of engineered cells, where a *cell-based chip* provides solid and fluidic support for long term maintenance and reagent/sample manipulation, acting simultaneously as a sensor, a processor analyzing complex data, and an output device that translates the detection of diseases into information intelligible to humans. For example, cell-based biosensors have been integrated *on-chip* with microelectrode arrays, photodiodes, field effect transistors, impedance or potentiometric sensors¹³⁶.

Commonly used reporter modules rely on colorimetric, fluorescent, or luminescent readouts, but can also be further interfaced with electronic transducers such as acoustic detection, surface plasmon resonance, and electrochemical methods. Their choice mostly depends on assay specifications, in terms of sensitivities or technical resources. Importantly, colorimetric outputs are human readable, a property of interest for integration into low-cost, easy-to-use point of care devices, while luminescent signals offer ultrahigh sensitivities and wide dynamic range of detections. However, instead of measuring traditional end point signals, other biosensing frameworks exist, and can be achieved thanks to properties inherent to biological systems, where information processing capabilities of genetic networks *in vivo* can be exploited (see section 1.4). It is thus possible to define different modes of readout, such as linear, frequency, or threshold, or multivalued modes of detection. For example, a riboregulated transcriptional cascade counter that uses multiple regulatory layers, enables a cell-based biosensor to give an output that is the function of the number of successive time delayed input signal events. These *counting* systems could offer new modalities of biosensing where the output is the exact sum of signal triggers in time and not concentrations¹⁸⁻²¹. Other authors have developed frequency-modulated cell-based

biosensors, and suggested that oscillatory sensors could confer a number of advantages over traditional ones. Cell based biosensors relying on optical reporter can for example be improved by frequency measurement, which is less sensitive to environmental factors compared to bulk intensity measurements that require normalization and calibration^{137 138} (**Figure 6**).

Synthetic biology is thus advancing the design of genetically programmed cell-based biosensors by increasing the diversity of readout modes that can be implemented, the nature and complexity of molecular biomarker patterns that can be detected and processed.

1.2.3.1.1 Microbial systems

The first microbial systems designed for the detection of various molecular cues such as organic chemicals, heavy metals, drugs, or toxics were developed early and proved useful in many applications, such as the MicroTox (Modern Water) and BioTox (Aboatox) assays. In some cases they could operate in complex matrixes such as human serum^{139 140} and urine¹⁴¹, where they were used to measure for instance biomarkers of toxic exposure, or *in vivo* where it was shown that exposure to antibiotics could be measured *in situ* in the rat gut, as well as other mammalian body fluids and tissues^{142 143 144} and on the field to assay complex foods¹⁴⁵ or soil samples¹⁴⁶.

The microbial sensor module determines selectivity and sensitivity of detection of pathological signals, and is traditionally derived from bacterial sensory systems such as transcriptional regulators-inducible promoters from stress responses or degradation pathways. For this reason, natural systems used in first generation of biosensors often lacked suitable selectivity/sensitivity required for biomedical applications, which motivated the increasing development of orthogonal sensing parts and devices through synthetic biology¹²⁰. The engineering of orthogonal sensing modules allowed more flexibility for tailoring detection specificities, sensitivity, and transfer functions. For example, the rational engineering of RNA riboswitches^{147 148}, or periplasmic binding protein^{149 150 151} enabled detection of various small molecules ligands such as the drug theophylline¹⁵², metal ions, nucleic acids, and proteins^{153 154} or extracellular biomarkers such as glucose, trinitrotoluene, L-lactate respectively. A growing repertoire of orthogonal synthetic parts dedicated to the engineering of biosensing systems is constantly emerging, such as ncRNAs^{155 156}, two-components systems¹⁵⁷, and intracellular protein transcriptional regulator-promoter pairs^{158 159}.

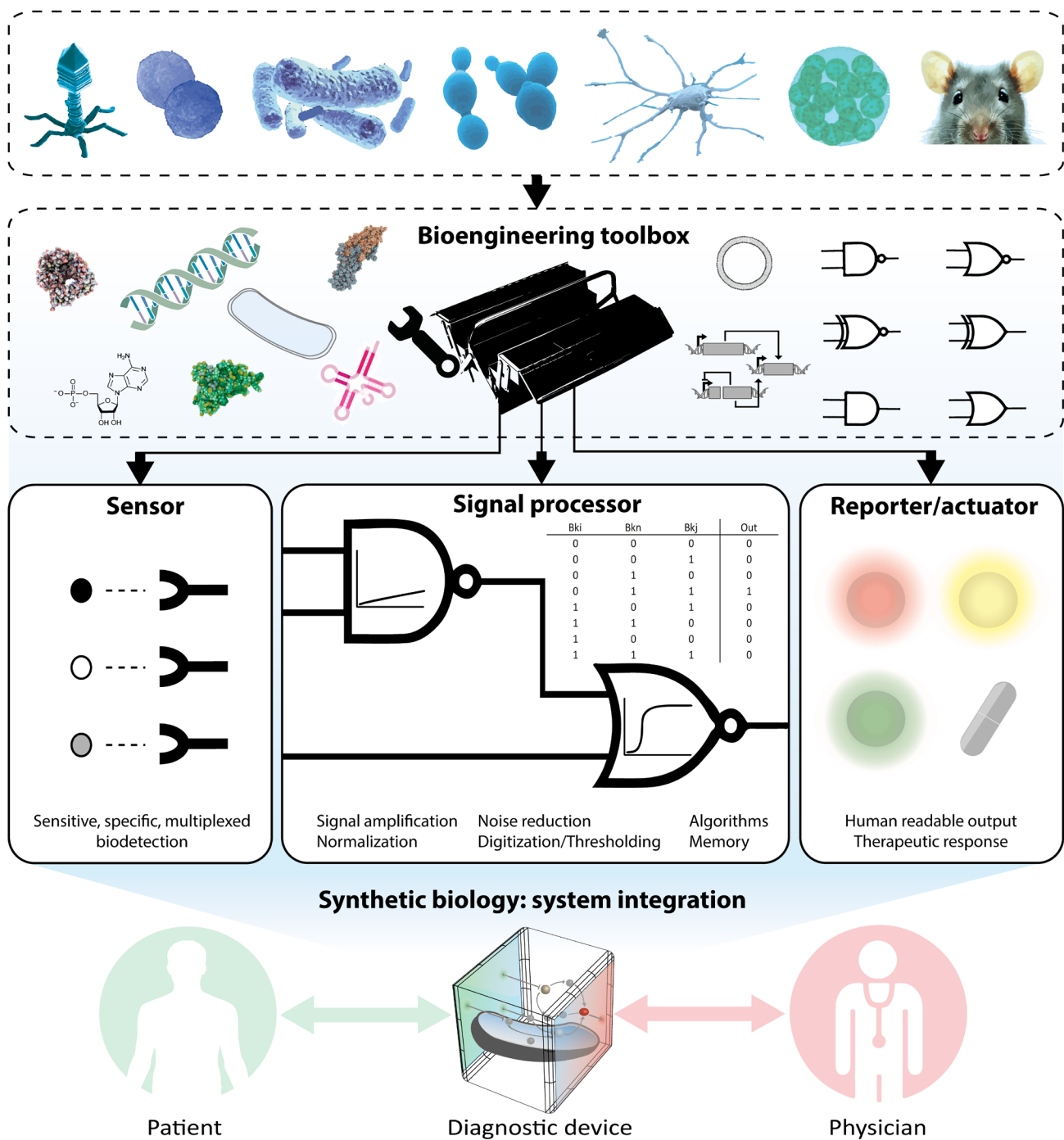


Figure 5: Synthetic biology enables the engineering of next-generation diagnostics through the system integration of sophisticated biological capabilities. Biological systems have evolved powerful molecular modules to sense and process biological signals and inform their phenotypes accordingly. Synthetic biology enables the systematic re-engineering, standardization and cataloguing of useful biological parts. It supports hierarchical abstraction from biological complexity for efficient assembly of parts into useful, composable and programmable modules. Medically relevant modules for sensing biomarkers, achieving signal processing and reporting can then be further integrated into biosensing systems to develop novel diagnostic devices meeting clinical specifications to aid medical decision making.

Synthetic biology provided methods for the assembly of complex genetic circuits to achieve reduction of expression noise and improve the signal-to-noise ratio, through the fine tuning of promoter strength^{160 161} to the integration of multiplexing circuits in single or multiple cell consortia¹¹⁹. Gene circuits can be integrated in microbial cell to develop biological filters and amplifiers to enhance biosensing selectivity and sensitivity and to develop logic gated bacterial sensors. For example, in this thesis I engineered a bacterial biosensor system we called Bactosensor, as an aid to diagnosis associated medical decision (Courbet *et al.*⁹⁶) (**Figure 8: case 1, See Chapter 2**). This approach offers interesting advantages that we believe could have consequences in medical practice. Bactosensors could provide simple use, cost-effectiveness, high sensibility and specificity, multiplexing and built in memory capacity, as well as embedding medical algorithms, while needing no clinical sample preparation¹¹⁷. Additionally, we proposed that encapsulation of bacteria in stable hydrogels could provide a disposable and portable format. We showed that bactosensors could operate in urine and serum, and demonstrated that their use could be of interest in the non-invasive screening for glycosuria and diabetes in urine samples. Although the use of bacterial biosensors in clinical samples had already been described^{140 142 162 139 141 144 145}, the robustness and reliability of living biosensors toward effective use in the clinic had not been addressed. Assaying complex *real world* samples is challenging because of the matrix effects of chemical mixtures on biosensor's behavior. In our study, we thus proposed a systematic method to evaluate the operational robustness of bacterial biosensors for the clinics and optimization of biosensing, signal processing and readout synthetic modules.

However, classical approaches do not enable cell-based devices to sense all species of clinical relevance, such as protein biomarkers (albuminuria, antibodies, antigens...) which do not naturally enter bacterial cells. In order to interface robustly with host physiology, cell based devices necessitate the engineering of cell-surface sensor modules. Interestingly, bacteria are able to sense and respond to extracellular analytes via two-component systems, which constitute precious elements to implement new biosensing frameworks in prokaryotes. These receptors are intrinsically modular, and have already been successfully re-engineered for different biodetection purposes¹⁶³. Moreover, programmable bacterial cells with alternative sensory modules such as mechanical, electrical and chemical systems to detect external stimuli via ion channel, or magnetosome for example, could be in the future exploited for a variety of diagnostic applications^{164 165}.

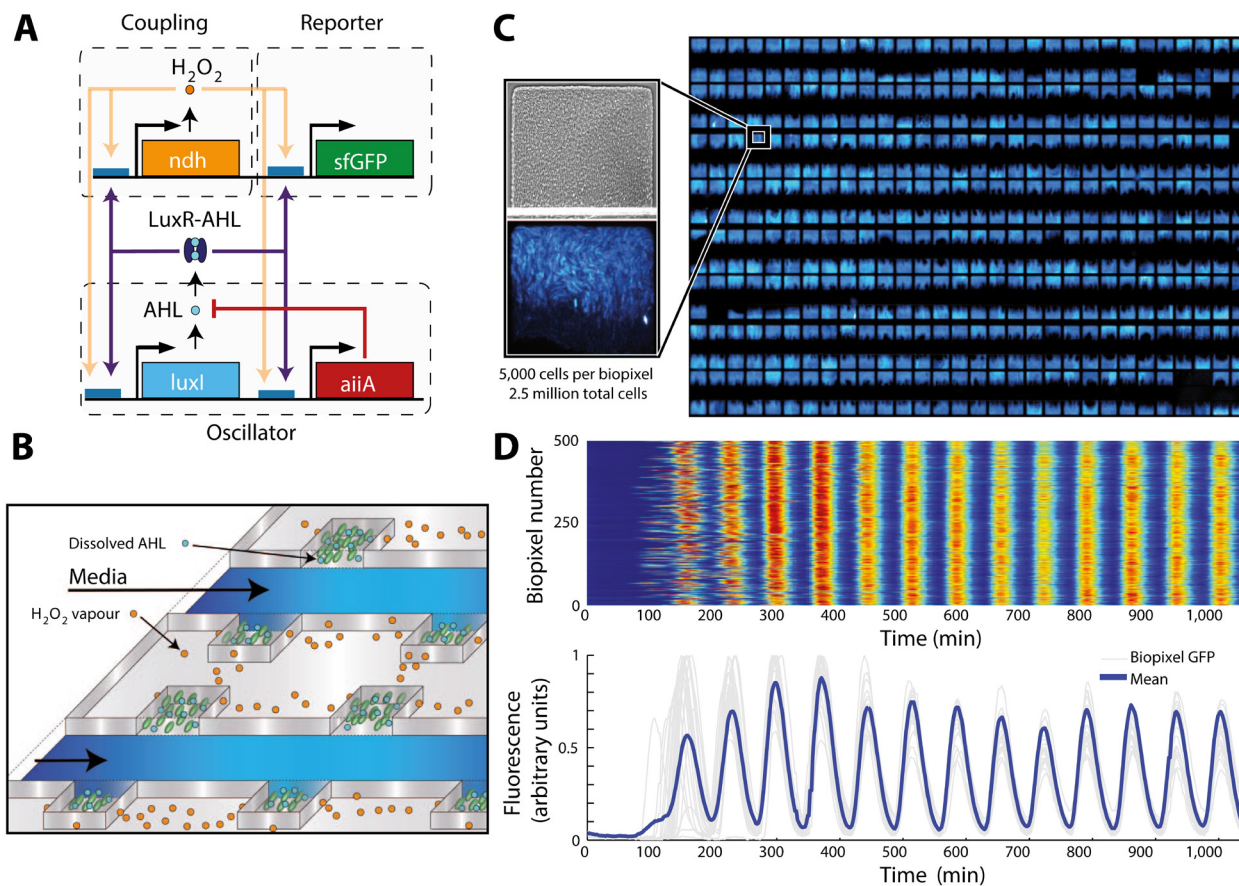


Figure 6: A sensing array of radically coupled biopixels. (A) Proposed synthetic biosensing circuit diagram: The luxI promoter drives expression of luxI, aiiA, ndh and sfGFP. The ndh gene codes for H₂O₂ producing enzyme, which is an additional activator of luxI promoter, which with AHL sensing via the quorum sensing genes luxI and aiiA triggers synchronized oscillations at the colony level. (B) Conceptual design of the sensing array: AHL diffuses within colonies while H₂O₂ diffuses between adjacent colonies through the PDMS. (C) Fluorescent image of an array of 500 *E. coli* biopixels (~2.5 million cells). (D) Heat map of the trajectories depicting time-lapse output of 500 biopixels undergoing rapid synchronization. (Adapted from Prindle *et al.*¹³⁶)

Synthetic biology efforts also permitted to further advance the engineering of new microbial biosensor systems through chassis optimization¹⁶⁶. Chassis can be engineered to behave appropriately in the desired environment. For instance, a microorganism biosensor could be designed to operate and withstand a particular physicochemical stress in harsh environments such as human serum. Moreover, a particular task or biosynthetic device may operate differently across chassis, and most laboratory strains of microorganism would not fit requirements for clinical applications. A promising solution is to develop synthetic streamlined chassis¹¹ with minimal functions required for their operation in media. Most approaches made use of *Escherichia coli*, which still remains the model platform of choice for synthetic biology for its ease of use, vast biological description, specific synthetic parts optimized, and engineering experience. However, one drawback of using *E. coli* as a chassis is the limited repertoire of relevant biosensing modules (e.g. promoters, transcriptional or posttranslational systems) to sense biomarkers. *Bacillus subtilis* is a promising and adaptable alternative chassis for synthetic biology^{116 167} and could be of great interest for biosensing purposes as the number of parts and devices available increases, and considering it offers interesting characteristics like genome minimalisation, assembly of genome-

scale heterologous DNA fragments, a wide range two-component and quorum-sensing systems, and the ability to sporulate after what it can be simply harvested and dried for long term storage and distribution. *B. subtilis* is a promising chassis to develop bactorsensors for its ability to engineer synthetic membrane receptors connected to orthogonal signal pathways to drive signal processing of pathological signals. In addition, biotechnological domestication of new chassis through synthetic biology, for example *Pseudomonas*¹⁶⁸, is likely to promote the emergence of new, robust microbial platforms with interesting physiological and stress-endurance characteristics for biosensing in clinical conditions.

Freeze drying of bacterial cells has also been proposed as a convenient way for the long term storage and distribution of most bacterial species for biosensor assays. However, it may add some complexity to the manufacturing process. Contrarily, the properties of spores make their use interesting for the development of cell-based diagnostics. Sporulation enables stable storage format, handling and shipment of biosensors with extended shelf-life¹⁶⁹. Spores can be integrated in miniaturized portable devices where spore germination, incubation with clinical samples, and signal detection are all integrated. For example, Date *et al.* have developed a μ TAS device for the detection of arsenite and zinc using engineered *B.subtilis* spores. Germination of spores and quantitative response to the analyte could be obtained at room temperature in 2.5–3 h with detection limits of 1×10^{-7} M for arsenite and 1×10^{-6} M for zinc in serum samples¹⁷⁰. In another study, properties of spores themselves have been used to develop a real time biosensor, or label-free exponential signal-amplification system¹⁷¹. The authors showed that this technique could be used to detect bacterial contamination in platelet concentrates with kinetics of the order of minutes. Like formation of spores, immobilization of cells has received much attention. In this thesis, I propose that this strategy could improve the analytical performance, handling, storage and preservation of microbial biosensors without the need of continuous cultivation, and to make them suitable for integration into deployable and *ready to use* devices¹⁷². Different strategies have been proposed as a way to obtain stable microbial biosensors encapsulation, covalent binding, adsorption, and cross-linking on various substrates. Although interesting formats have been proposed like paper strips¹⁷³ we suspect that the encapsulation of bactorsensors in hydrogel beads increase robustness and preserve viability and response characteristics of sensing cells under the harsh environmental conditions they are exposed to by protecting them, prevent their spread, and enable multiplexed biosensing as well as the combination of algorithmic operations in different population of beads at the same time.

Microbial cells thus offer a rich playground to engineer novel diagnostic tools, and we believe new biomedical technologies allowing new usages are likely to emerge. For example, as natural commensal microbiome flora, prokaryotes could be used in the form of diagnostics probiotics to monitor for example gut pathologies *in situ*. A recent study showed that bacteria could be engineered to detect and record biological signals inside the mammalian gut in a programmable way⁴¹. Bacteria could be also engineered as theranostic microrobots targeting solid tumors *in vivo*¹⁷⁴. More recently, Danino *et al.* engineered a probiotic *E. coli* strain as an orally administered diagnostics to noninvasively monitor liver cancerogenesis⁹⁸ (**Figure 7**). Their microbial diagnostic platform was capable of recording signals arising from metastasis *in vivo* and generated an output signal measurable in the urine, for extended periods of time without deleterious health effects in mice.

Alternatively, other approaches to diagnostics development using engineered microbial cells are emerging. It has been recently demonstrated that microbial cells could be engineered to generate synthetic tunable multiscale nanomaterials (such as gold-particle patterning to create nanowires and nanorods) that can be conjugated with target ligands and drug molecules for diagnostic applications¹⁷⁵.

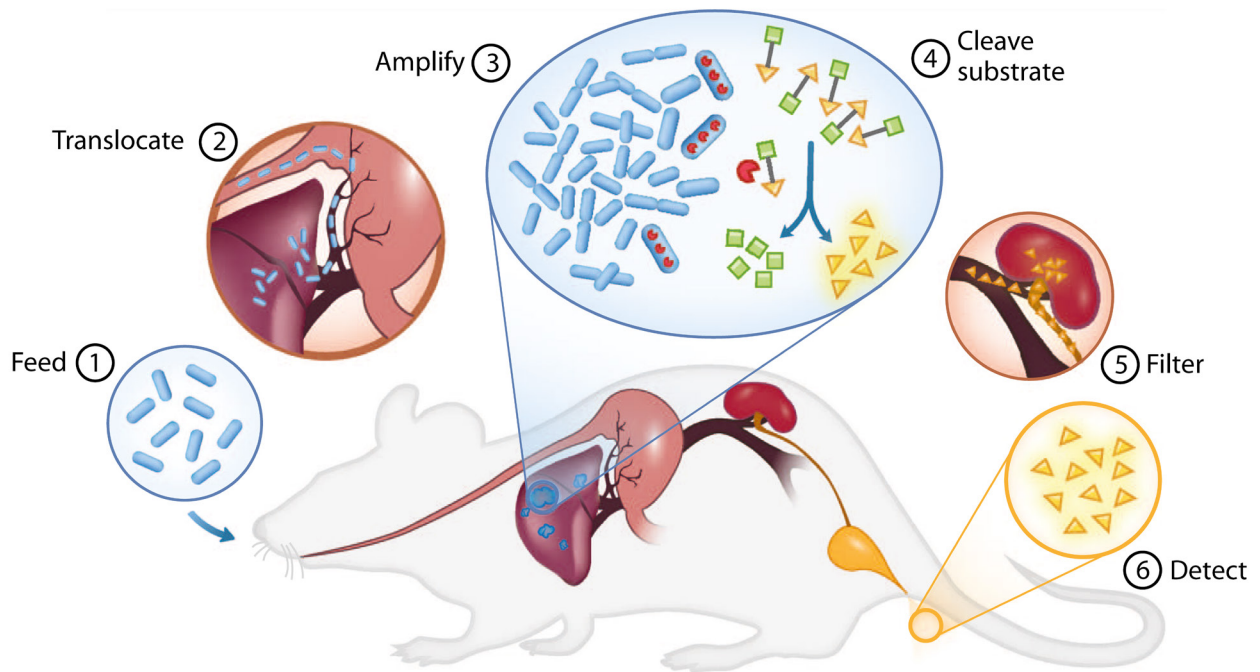


Figure 7: Programmable probiotics for non invasive detection of cancer. The PROP-Z diagnostic platform is made up of probiotic *E. coli* N bacteria transformed with a dual-stabilized, high-expression lacZ vector as well as a genomically integrated luxCDABE cassette that allows for luminescent visualization without providing exogenous luciferin (blue). (1) PROP-Z is delivered orally. (2) Probiotics rapidly (within 24 hours) translocate across the gastrointestinal tract and (3) specifically amplify within metastatic tumors present in the liver. (4) PROP-Z expresses high levels of the enzyme lacZ (red), which can cleave systemically injected, cleavable substrates (green and yellow). Cleavage products of the substrates (yellow) filter through the renal system (5) into the urine for detection (6). (Adapted from Danino *et al.*⁹⁸)

1.2.3.1.2 Eukaryotic systems

Eukaryotic systems are physiologically closer to humans with a similar metabolism, and compared to prokaryotes benefit from a more sophisticated genome, proteome and cellular organization. While it increases the available bioengineering space, eukaryotes are also more complex to engineer. Interestingly, this implies that the extrapolation of biosensing measurements can be more informative and of greater relevance for certain detection agendas. Although more recent, the toolbox of biological parts and devices that operate in eukaryotic and more specifically in mammalian cells is rapidly expanding^{176 177}.

Similarly to microbes, natural eukaryotic or systematically prokaryote derived¹⁷⁸ nucleic acid and protein-based sensor modules have been developed to detect diverse ligands such as small molecules: vitamins and metabolites^{179 180 181 182}, gases like acetaldehydes and nitric oxide^{183 184}, pH¹⁸⁵ or hypoxia¹⁸⁶ or combinations of such molecular signals. Some recent mammalian biosensing systems made use of RNA-based sensors to detect clinically relevant biomarkers¹⁸⁷. RNA aptamers are interesting as sensing modules because they can be easily engineered *de novo* to target either small molecules, proteins, or other RNAs inside live cells, through various selection strategies¹⁸⁸. For example, RNA aptamers could detect increased levels of intracellular protein inputs in the NF- κ B- and Wnt-signaling pathways¹⁸⁹ by linking detection events to alternative splicing of an output gene. Emerging frameworks involve synthetic protein modules exploiting programmability of zinc finger binding, which can for instance detect specific DNA sequences and trigger an intracellular response via coupling to a synthetic gene circuit¹⁹⁰.

Auslander *et al.* recently reported a mammalian cell-based biosensor capable of precise profiling of allergies in human whole-blood samples⁹⁷ (**Figure 8: Case 2**). This device, a histamine biosensor, consisted of a synthetic histamine-responsive signaling cascade in which the G protein coupled receptor HRH2 senses extracellular histamine levels and then triggers Gs-protein-mediated intracellular signaling and activation of a reporter gene. By exposing patient's whole-blood samples to an array of allergens, basophil cells undergo an allergen-specific release of histamine which replicates the specific allergic reaction in the body. The serum was analyzed by the designer cells that could precisely score the allergen-triggered release of histamine, thereby integrating histamine levels with interesting sensitivity and dynamic range of response. This strategy proved very interesting, when current *in vivo* and *in vitro* diagnostic methods to determine the molecular aetiology of allergic syndromes suffer from lack of reproducibility, patient discomfort, bulky experimentation, low dynamic range and poor correlation with clinical symptoms. This study emphasizes the interest of such devices as it provides non-invasive, personalized allergy profiles, and pioneers the use of engineered cell-based biosensors for novel diagnostic methodologies.

Engineering of eukaryotic systems also enables the *intracellular* diagnostics approach, which involves genetically encoded non-invasive detection of combinations of small molecules, nucleic acids, and proteins in live cells over time. This strategy could allow the measurement of intracellular molecular and genomic markers, while taking into consideration the cellular context. For example, instead of probing the chemical nature of a cell's genome, this new approaches can

account for cell and genome and epigenome topology and regulatory organization *in situ*, which is known to be of functional physiological and pathological relevance^{191 192}. For instance, instead of measuring averaged signals of a cell population in the steady state, intracellular synthetic gene circuits can give access to time and space resolution, while enabling the monitoring of the cell's gene-phenotype relationship¹⁹³, which is a fundamental challenge in human health.

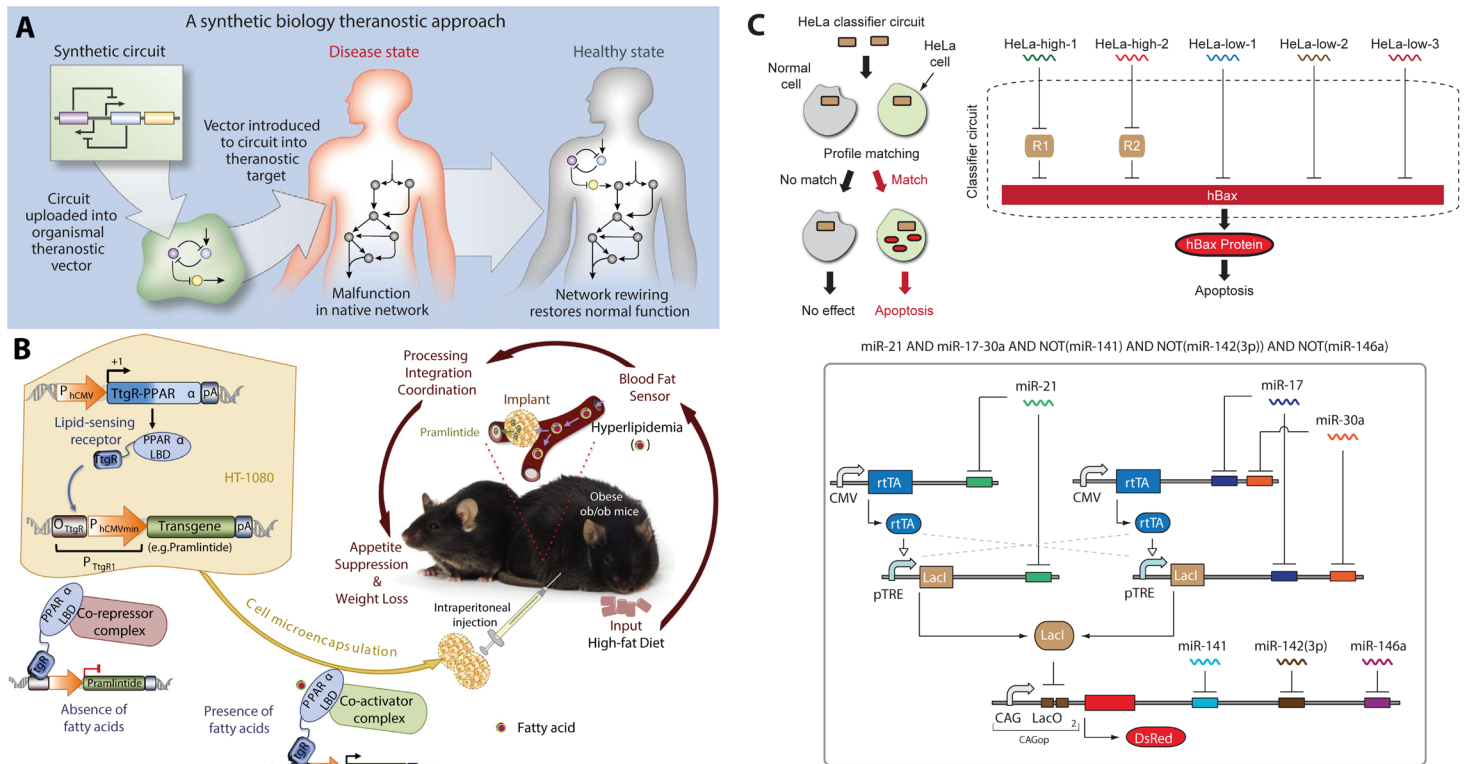


Figure 9: Synthetic biology approaches to theranostic and prosthetics. (A) Conceptual approach to synthetic circuit development for theranostics and prosthetics. Synthetic gene/biochemical circuits are uploaded into cells to therapeutically target the body's endogenous networks, causing a transition from disease to healthy state. (Adapted from Ruder *et al.*⁵⁶) (B) Theranostic engineered cells for obesity. A fusion protein combining the phloretin responsive repressor (TtgR) and the human peroxisome proliferator-activated receptor (PPAR α) which bound a TtgR-specific operator (OTtgR) linked to a minimal promoter (PhCMVmin) (PTtgR1 as a whole) to control transgene expression, was expressed in mammalian cells. In the absence of fatty acids, the lipid sensing sensor (LSR) associated with the co-repressor complex to repress transgene expression, but switched to full induction in the presence of fatty acids when LSR associated with the co-activator complex. Mice with diet-induced obesity were implanted with microencapsulated cells engineered for LSR-driven expression of pramlintide. These mice showed a significant reduction in food consumption, blood lipid levels and body weight when maintained on a high-fat diet. (Adapted from Kojima *et al.*¹⁰⁶) (C) Synthetic gene circuit for cancer cell type classification. Top: Schematic representation of a HeLa specific classifier circuit operation, and detail of the synthetic circuit taking cancer specific miRNA biomarkers as inputs to trigger Bax dependent apoptosis of cancer cells. The entire circuit implements a multi-input AND gate. Bottom: Detailed genetic implementation of the classifier circuit and logic formula taking into account 5 pathological miRNAs. (Adapted from Xie *et al.*¹⁰²).

Taking these considerations further, clinical synthetic biology has long been interested in the promise that engineering of mammalian cell-based biosensing devices could enable diagnosis of pathological states conjugated to therapeutic modulation of human physiology. Synthetic biologists are thus trying to develop and integrate synthetic gene circuits *in vivo*, which would directly link the detection of molecular disease signals to targeted therapeutic activities. This strategy involves synthetic gene circuits also known as *prosthetic networks*. Prosthetic networks act as intracellular molecular prosthesis that sense, monitor and score disease-associated biomarkers and coordinate an adjusted diagnosis, and timely preventive or therapeutic responses for increased efficacy and safety. In the form of implantable devices, they could act as self-powered, autonomous circuits that trigger pharmacological systemic or *in situ* responses to restore deficient phenotypes. This recent strategy that has yet to prove applicable in the clinics, could offer potential applications in the long-term surveillance and intervention of cancerogenesis, but also infectious or chronic diseases, such as gout and diabetes (**Figure 9A, 9B**).

The potential of medical prosthetic networks was demonstrated in a pioneering example reported by Kemmer et al.¹⁹⁴ Gout is associated with non-regulated, pathological levels of uric acid. The authors showed they could engineer a synthetic mammalian genetic circuit to sense, and maintain uric acid homeostasis in the bloodstream of mice. In their design, a modified *Deinococcus radiodurans*-derived protein that senses uric acid levels triggers a dose-dependent de-repression of a secretion-engineered *aspergillus flavus* urate oxidase that eliminates uric acid. The authors also showed they could insulate the circuit in transgenic cells by immunoprotective microencapsulation. Implantation of these designer cells could treat animals by reducing the levels of uric acid to subpathological levels. Similar proves of concept have been demonstrated for metabolic diseases such as Diabetes¹⁹⁵, or diet induced obesity¹⁹⁶.

However, precise discrimination between clinical states is essential for such autonomous decision-making devices. Again, combination of multiple context-specific promoters has proved more efficient than single input approaches that suffer from linear responses and limited control of specificity and efficacy. In oncology, more and more routine diagnoses are based on molecular signatures rather than symptomatology. Nissim *et al.* engineered the mammalian two-hybrid system to act as an autonomous logical AND gate that integrates as inputs signals arising from cancer-related promoters and expresses a killer (or reporter) gene specifically in cancer cell lines. This approach provided increased response tunability and revealed a digital-like response of input amplification following a sharp activation threshold, providing robustness, minimizing input noise and false-positive identification of cell states¹⁹⁷. In another key study, using gene expression levels that are commonly used to diagnose prostate and small-cell lung cancer, Benenson *et al.* designed a DNA circuit that computed five yes/no molecular sensing events *in vivo* in order to detect biomarkers of prostate cancer. This biomolecular computer was designed to be conditionally responsive to the presence of five positive biomarkers to generate a therapeutic output¹⁹⁸. In another example of intracellular prosthetic diagnostics, a platform that integrates logic and sensing could detect pathogenic patterns of miRNAs *in vivo*¹⁰² (**Figure 9C**). The authors generated a classifier system through straightforward engineering of nucleic acid hybridization reactions, which could assess whether the transient expression profile of six endogenous miRNAs matched a specific profile characteristic of cervical cancer. This genetic logic circuit could identify cancerous cells and triggers apoptosis in response. This approach could be in

principle extended to the detection of complex molecular pathophenotype and connected to *in vivo* therapeutic actions. Synthetic gene network built using CRISPR-Cas9 technology in mammalian cells, also showed capable of integrating cellular pathophysiological information from two cancer specific promoters. Using these cancerous triggers as inputs, the system could then activate an output gene following a AND Boolean operation. When using a luciferase output, the authors could detect bladder cancer cells or induce cell death using functional apoptotic genes as outputs¹⁹⁹. Last, Wroblewska et al. recently showed that RNA binding protein circuits could be easily delivered in cells to exert anticancer theranostic regulation *in cellulo*²⁰⁰. These studies brought promising proves of concept toward the clinical use of personalized designer theranostic cells, which could be further engineered to produce different responses, such as the *in situ* production of imaging agents to aid the diagnostic of tumors and metastases, associated with an anticancer action.

However, physiologically relevant cues are often extracellular, and require tools to sense a vast range of ligands in complex environments: cytokines, hormones, various proteins, pathogens, hypoxia, inflammation, or pH. At the same time, keeping high orthogonality in sensing components is required to avoid modes of failure and interface robustly with the patient host. The engineering of such novel mammalian sensor systems can be achieved through different strategies: redirecting the output of natural receptors, or engineer existing transmembrane sensor proteins to recognize small molecule inputs or user specified antigens (reviewed in²⁰¹). While the first approach in this direction showed successful demonstration to detect endogenous molecules via the rewiring of Notch, GPCRs or RTK signaling to elicit novel responses, diagnostic applications may require receptors that detect biomarkers for which there are no endogenous receptors²⁰¹. To address this need, some authors recently developed a technology they termed Modular Extracellular Sensor Architecture (MESA). It consists in a fully orthogonal architecture where independent, tunable protein receptor modules undergo ligand binding-induced dimerization, which further results in proteolytic trans-cleavage of the intracellular part, releasing a transcription factor previously sequestered at the plasma membrane. They developed a systematic platform for conditional transmembrane ligand detection that produces outputs in the form of either transcriptional regulation or reconstitution of enzymatic activity, and enable straightforward engineering for the detection of user defined ligands¹⁷⁹. Another interesting extracellular receptor that has received attention as a recognition element are G-protein coupled receptors (GPCRs). GPCRs represent the largest family of membrane receptors, are highly modular and their customization could benefit from a large range of natural binding repertoire ranging from small molecules to peptides and glycoproteins biomarkers. Moreover GPCR in cell-based biosensing can be connected to various cellular processes to be used as the sensor readout. For example, directed evolution of GPCRs permitted to obtain receptors with specificities for small molecules²⁰². This strategy has been employed in mice with good success and could be interesting for novel diagnostic or analytical purposes²⁰³. In addition, the engineering of novel immune receptors, with the same modularity, diversity and selectivity as antibodies thus capable of sensing a wide range of disease-associated antigens, such as protein biomarkers of cancer, infections or cardiovascular risk, was achieved with chimeric antigen receptors (CARs). CARs are designed with single-chain antibodies (scFvs) that are fused to cytoplasmic regions of intracellular signaling elements (the CD3 zeta chain), which linkage leads to a novel modular input/output sensor that activates upon binding to the target. These synthetic receptors also enabled the

tailored re-programming of T cells to respond to defined ligands, and proved clinically extremely promising for cancer immunotherapy^{204 205 206}. Therefore, these synthetic receptors could open the way for novel cell-based biosensors for diagnostic applications.

While various eukaryotic chassis have been engineered into cell-based biosensors, mammalian cells have dominated synthetic biology medical proves of concepts. However, the yeast *Saccharomyces cerevisiae* also constitutes a potentially interesting chassis for biosensor development¹²⁷, and can be stored and distributed in a dry active state. As a model organism, many genetic engineering tools are available. Extracellular yeast mating peptide sensing systems are G-protein coupled receptors, which initiate an intracellular signaling pathway, and could be good target for the engineering of new biosensing devices²⁰⁷. To date, *S. cerevisiae* remains an underexploited yet promising platform for biosensor development.

1.2.3.1.3 Viral systems

The ability of viruses and more specifically bacteriophages to specifically infect and lyse their bacterial host has been exploited for many decades to reveal and identify bacterial species. Phage-based diagnostics have been recently further investigated as an emergent technology for the clinical diagnosis of infectious bacterial diseases, and synthetic biology approaches have already played a major role in the engineering of phage based technologies for the detection of human pathogens^{208 209}.

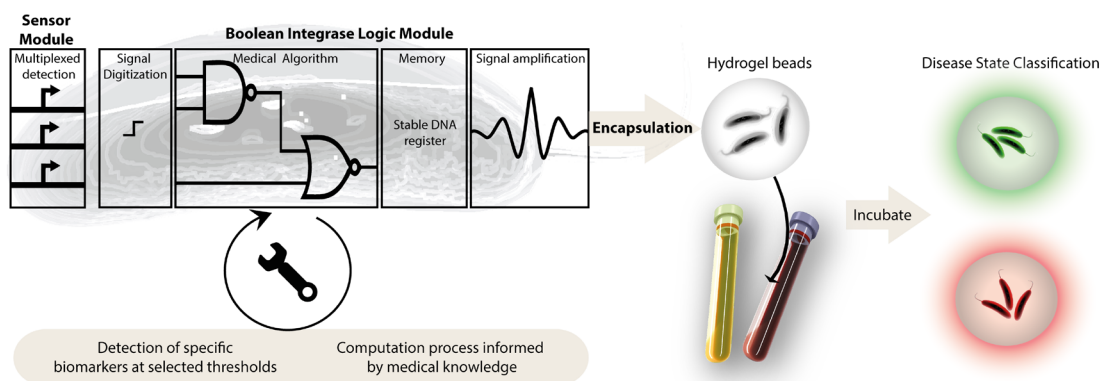
Near-real-time microbial diagnostics remain of critical interest in the clinics, where timely detection of pathogens and delivery of species specific evidence based therapy is a life-threatening issue²¹⁰. Microbial diagnostics currently suffer from well-recognized limitations, since they requires an enrichment step during which pathogens are amplified over incubation times ranging from 10 to 48 h, or even more than 10 days for certain pathogens such as *Mycobacterium tuberculosis*. Moreover, standard techniques such as microscopy lack sensitivity, nucleic-acid amplification tests such as PCR offer molecular specificity but have complex sample preparation and poor reliability (inhibition, false positives...), and immunoassays although highly sensitive, are labor intensive and challenging to implement multiplexed detection. To date, bacterial culture isolation remains the standard for species identification and confirmation. Consequently, there is a greater emphasis on the direct detection of pathogens from clinical specimens, without the need for tedious and slow isolation of pure bacterial cultures. Phage-based diagnostics can be regarded as a versatile, widely applicable and valuable solution to timely microbial diagnostics, and synthetic biology has already shown its potential to dramatically improve this technology (**Figure 8: Case 3**).

Natural phages can be engineered to deliver genetic information into specific bacterial species, thus exploiting their metabolism for the production of readable molecular signals (fluorescent, or luminescent proteins.)²⁰⁸. Synthetic viruses can be rationally designed and engineered into modular viral scaffolds^{211 212} or modified via directed evolution²¹³ and chemical and genetic modification can be used to generate numerous functionalities²¹⁴ and cell target specificity²¹⁵.

Different phage-based assays formats and detection methods have been investigated: phage amplification with bacterial lysis^{216 217}, phage/DNA amplification followed by quantitative PCR to identify phage DNA amplification^{218 219}, dot blot assay²²⁰, phage-integrated colorimetric, fluorescent, and bioluminescent reporter genes^{221 222 223 224 225 226 227 228}, phage/protein amplification detection with phage-specific antibodies²²⁹. More recent developments include quantum dot reporting, electrochemical and optoelectronic methods (for an extensive review see²³⁰), or innovative biophysics methods²³¹. Diagnostic sensitivities as low as 10 cells/mL with a response time of 1 hour in a clinical sample matrix have been described, and a number of proof of concept and commercial products showed a very good response time and sensitivity in medical context²³⁰. The utilization of cocktails of phages or the assembly of phage-derived recognition proteins has been proposed to specifically detect desired bacterial spectra. The advantage compared to other detection method like hybridization based assays, is that it doesn't require an enrichment step and sample pretreatment to achieve maximum specificity and sensitivity, and provides discrimination between living cells and dead cells. In addition, the wide bacterial selectivity range, host specificity, ease of use, straightforward production and extremely low reagent cost, seem to make phages ideal candidates to exploit as bacterial detectors in a variety of culture, food, water, clinical and environmental matrices²³². Phage diagnostics can also give information about the genetic nature of the host, and thus can be used for antibiotic susceptibility testing²³³. For example, identification of *M. tuberculosis* by culture on solid or liquid media takes more than 10 days, requires specialized and costly equipment, and technical expertise and show poor sensitivity for identification. Mycobacteriophage amplification technology or reporter mycobacteriophage technology allows *M. tuberculosis* detection in less than 48 hours, along with providing antibiotic susceptibility testing²³⁴. As another example, blood culture tests such as KeyPath™, allows for simultaneous identification of *S. aureus* and differentiation between MSSA and MRSA (Methicillin sensitive and resistant *S. aureus*, respectively)²²⁹. Phage-based platforms are also currently clinically used for the detection of *Yersinia pestis*, *Bacillus anthracis*²⁰⁸.

However, few prototypes have been fully translated from laboratory to the clinics and have been successfully commercialized. Key bioengineering advances provided by synthetic biology are required for full maturation of this technological field to achieve enrichment free, sensitive, specific, straightforward phage based diagnostic tests. High throughput and genetic engineering tools, libraries of robust and reliable devices and parts such as reporter genes, sensitive sensors and synthetic gene circuits may enable the engineering of the huge natural phage *repertoire* chassis (over 10³²) at a much more higher pace than achieved so far^{99 235}.

Case 1: Bacterial whole-cell biosensor



Case 2: mammalian whole-cell allergy profiler

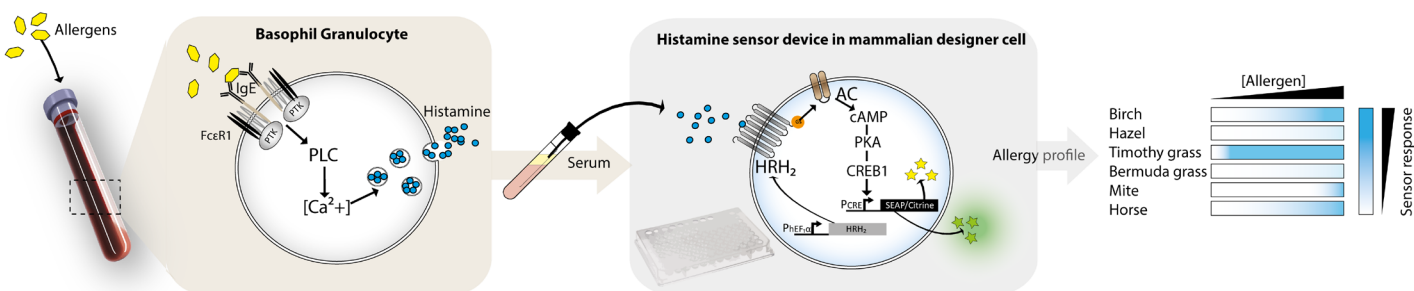


Figure 8: Case studies: recent synthetic biology research strategies to provide novel diagnostic tools. Case 1: next-generation bacterial biosensors for medical diagnosis detecting biomarkers in human clinical samples with a robust, programmable, and reliable behavior for clinical use (adapted from Courbet et al.⁹⁶). **Case 2:** mammalian cell-based biosensors that score the allergen-triggered release of histamine from whole-blood-derived human basophils. A synthetic signaling cascade engineered within the allergy profiler rewires histamine input to the production of reporter protein, thereby integrating histamine levels in whole-blood samples (Adapted from Ausländer et al.⁹⁷).

Synthetic virus based devices have also been shown to be useful for the rapid typing and monitoring of specific eukaryotic cell phenotypes. Until now, they have been extensively used for therapeutic purposes and virus-mediated delivery of effector genes and payloads²³⁶. Similarly, prosthetic decision making circuits embedding diagnostic algorithms could be delivered via viral vectors *in vivo* into mammalian cells, injecting molecular computers probing the internal state of a cell. As previously discussed, such payloads supported by synthetic gene circuits can then sense, score, monitor and store disease-relevant molecular information. For example it could contain cancer specific promoters²³⁷⁻²³⁸, and an actuating device transmitting the cell's pathological state to human readable information.

These design principles have also extensively been investigated for *in vivo* imaging diagnostic strategies²³⁹⁻²⁴⁰. They were applied for example for different imaging modalities: insertion of key genes in melanogenesis in a vaccinia virus vector allowed improved MRI and optoacoustic imaging, in a tumor specific manner²⁴¹. Another method for non-invasive optical imaging of

tumors *in vivo* was successfully developed and uses engineered viruses that carry genes and probes to allow deep tissue molecular imaging²⁴² or further encodes enzymes (β -gal and glucuronidase) that can be monitored in the serum of tumor-bearing mice²⁴³ as well as in the blood of humans with cancer²⁴².

Another interesting field of virus engineering research enabled by synthetic biology is the engineering of synthetic viral nanoparticles and their genome-free analogues: virus-like particles. A broad range of genetic and chemical engineering methods have been developed to exploit virus nanoparticles as biomedical imaging diagnostics reagents, and the inclusion of peptide ligands on the particle surface permitted the improvement of current *in vitro* diagnostic assays based on the conventional enzyme linked immunosorbent assays²¹⁴. In such assays, the viral nanoparticle helps guiding the antibodies to achieve maximum capture of the biomarkers. In addition, high densities of antibodies on the surfaces of the nanoparticles lead to greater binding of biomarkers, which enhances detection sensitivities. For example, it was recently showed that by combining viral nanoparticles, which are engineered to have dual affinity for troponin antibodies and nickel, they could detect troponin levels in human serum samples that are seven orders of magnitude lower than those detectable using conventional enzyme linked immunosorbent assays, exhibiting properties that could prove valuable in the early detection of the protein marker troponin in patients with a higher risk of acute myocardial infarction^{244 245}. Other viral nanoparticles could perform similar highly sensitive diagnostic assays and could be implemented for a variety of biomarkers.

While phage based diagnostics technologies are maturing and transitioning toward real world diagnostics, it is very likely that further engineering eukaryotic viruses will lead synthetic biologists to major medical developments toward the clinic²¹⁴. Viral nanotechnologies for diagnostic have now come of age and we believe that it will not be long before novel assays reach a prominent role.

1.2.3.2 Bottom-up engineering of biosensors

Following the advances in the construction of programmable biosensing circuits in living organisms, synthetic circuits assembled *ex vivo* in minimal systems from the bottom-up constitute viable approaches for designing, understanding, and exploiting dynamic biochemical circuitry for biodetection and thus diagnostic purposes. Cell-based biosensing systems often rely on intracellular passive diffusion of analytes, or kinetics of transcriptional and translational processes that result in slow sensor responses. In addition, non-orthogonal gene circuits constitute a load in engineered cells that can interact with chassis components and result in unpredictable and noisy response profiles. On the contrary, bottom-up synthetic systems that rely on nucleic acid, protein, or metabolites can reach temporal dynamics in the order of seconds or minutes. Released from unwieldy complexity, context dependencies, and unpredictability that burden the use of living systems, *ex vivo* systems allow researchers to directly access and manipulate modular biomolecular parts with unprecedented control and design space (**Figure 5**). Advances in such bio-inspired functional systems²⁴⁶ include diverse capabilities including: biosensing,

biocomputing, memory, signal amplification, and various biological functionalities. Progress in this field demonstrated that cell-free synthetic biology is a promising field for the fundamental understanding of native biological systems but most interestingly for the engineering novel biotechnological tools for the clinics^{247 248}.

1.2.3.2.1 Nucleic acid based systems

Nucleic acids are versatile molecules capable of information processing and storage. They are governed by simple, predictable and programmable rules driven by Watson-Crick base-pairing interactions and strand displacement that enable their straightforward nanoscale synthesis and engineering with important design space²⁴⁹. The past decades witnessed the development of complex *in vitro* nucleic acid circuits and devices highlighting the potential of using nucleobases and their polymers as building blocks to generate useful architectures^{250 251 252}. Nucleic acid based *in vitro* systems have made numerous contributions to biodiagnostic as well as biotechnology research, with the best example probably being the development of polymerase chain reaction. As signal detection, amplification and transduction depend on the programmability of Watson-Crick base pairing, nucleic acid circuits can be tuned and adapted to various applications compared to other biomolecular signal amplification reactions. Moreover, novel methods to select and amplify sequence-specific nucleic acids with specific recognition sites (aptamers) for low-molecular-weight analytes, macromolecules or whole cells and development of catalytic nucleic acids (DNAzymes or ribozymes) are promising and likely to provide new analytical tools²⁵³. Meanwhile, the field of DNA biocomputing and molecular programming has taken an increasing importance for analytical applications²⁵⁴. The modularity of nucleic acids, as well as their capacity to directly interact with a wide range of analytes, especially other nucleic acid biomarkers, enables the implementation of decision making circuits that are programmable functions between selected inputs and outputs, which are of outstanding relevance to diagnostic applications²⁵². A variety of sensing systems relying on nucleic acid devices have been developed during the past decade, with particular interest for riboswitches, aptamers, and catalytic nucleic acids (DNAzymes and ribozymes) coupled to more complex nucleic acid reaction networks.

Aptamers are synthetic single-stranded nucleic acids that selectively bind to a broad range of specific targets ranging from proteins to peptides, amino acids, drugs, metal ions, and even whole cells, and benefit from systematic and robust methods for their obtention through a combinatorial directed evolution method called SELEX²⁵⁵. They have demonstrated great promise in diagnostic biosensor development during the last decade, since they possess unique characteristics compared to antibodies or other biomimetic receptors, comparable or even better affinity, easy and cost-effective synthesis with high reproducibility and purity, simple and straightforward *de novo* design, engineering and chemical modification^{256 257}. Aptamers are thus powerful alternatives to antibodies or other biomimetic receptors for the development of diagnostics²⁵⁸. They proved valuable as diagnostic tools in several diagnostic applications and assay formats such as biomarker detection from cancer clinical testing to detection of infectious microorganisms and viruses (reviewed in²⁵⁹). For example, the possibility of using aptamers as an

alternative molecular recognition element in ELISA has received great interest, which gave rise to an ELISA-derived assay called enzyme-linked apta-sorbent assay (ELASA)²⁶⁰. Taking the versatility of aptamers further, recent studies proposed to develop intelligent aptasensors that embed boolean logic. For instance, Zhou *et al.* engineered biocomputing systems with aptamer-based biochemical sensing controlling a self-powered biofuel cell that process molecular information. This proof-of-concept could detect patterns of thrombin and lysozyme inputs and generate an electrochemical output following a NAND truth table²⁶¹. Even though these logic biosensors were shown to operate effectively in complex physiological sample, they still require significant engineering efforts prior to a potential practical application. Moreover, while most diagnostics are still under the supremacy of immunoassays, further studies are needed to evaluate clinical robustness of aptasensors in clinical sample matrices and to provide new sensing formats (reviewed in²⁶²).

Similarly, the discovery of natural riboswitches has inspired application to ligand detection, exploiting the ability of RNA to recognize molecular targets and harnessing the ligand-dependent structural rearrangement of RNA to generate a measurable signal^{263 264 265}. Riboswitches are RNA aptameric elements in RNA devices that control gene expression, refolding, or allosteric ribozyme activities *in vivo* in response to a broad range of specific ligands²⁶⁶. Riboswitches are integrated into RNA, and are mostly constituted of an internal ribosome entry site accessible for the ribosome only in the presence of a specific ligand, while it is inhibited in its absence. Because synthetic riboswitches make it possible to regulate any gene or RNA enzyme with an arbitrary molecule, they function as biosensors, in which the output is easily detectable protein expression or enzymatic activity that reflects the concentration of the corresponding ligand²⁶⁷. Rational design strategies for constructing novel riboswitches that work in cell-free translation systems have been described, and their systematic engineering for different biosensing targets, such as FMN, tetracycline and sulforhodamine B have been demonstrated²⁶⁸. In another approach, Olea *et al.* described a general analytical method for the detection of target ligands based on self-replicating aptazymes. These *autocatalytic aptazymes* are constituted of an aptamer domain linked to the catalytic domain of a self-replicating RNA enzyme²⁶⁹. Ligand-dependent self-replication of RNA proceeds in a self-sustained manner, undergoing isothermal and protein free exponential amplification. The rate of exponential amplification is a function of the concentration of the ligand, thus enabling quantitative ligand detection.

Catalytic nucleic acids, or DNAzymes, that can also be employed diagnostic reagents, and were extensively used as amplifying labels for optical and electrochemical sensing platforms. A vast repertoire of synthetic catalytic nucleic acids have been recently engineered, such as metal-ion-dependent DNAzymes, aptamer inducible DNAzymes and cofactor-dependent DNAzymes that catalyzes cleavage or ligation of oligonucleotides or mimic native enzymatic functionalities. Furthermore, DNAzyme have been employed to trigger catalytic cascades and thus used for amplified autonomous sensing of miRNAs and mRNAs within living mammalian cells and DNA logic gate cascades and computing circuits²⁷⁰. For instance, a method supporting nanomolar detection of histidine was reported, using a L-histidine-dependent RNA-cleavage DNAzyme²⁷¹. Likewise, a HRP-mimicking DNAzyme cascade was engineered for the amplified aptamer mediated detection of PDGF²⁷². DNAzymes also provided a colorimetric method to detect telomerase activity as a cancer specific cellular biomarker^{273 274}.

Other strategies recently developed, rely on the binding of single stranded DNA signals to a partially double-stranded complex by a single-stranded domain called a *toehold*, and then release the originally bound strand after branch migration has occurred. In this way, an output signal can be activated upon the arrival of an input signal, and the reaction rate can be controlled by the length and nature of the toehold. This concept permitted the development of many DNA strand displacement circuit strategies, resulting in a wide range of applications for *in vitro* diagnostics²⁷⁵²⁷⁶. For example, Chen *et al.* have recently developed a toehold exchange mechanism working with double-stranded nucleic acids, which they show can be used as a novel programmable diagnostic device to detect single nucleotide polymorphism. They demonstrate that conditionally fluorescent DNA probes are capable of detecting variations of a single base in a target dsDNA, reliably over a wide range of conditions²⁷⁷. They then successfully apply this principle to diagnose individual point mutations in Rifampicine bacterial antibiotic resistance genes in *E. coli*. This technology could also prove interesting to screen extended genetic regions and multiplex SNP detection.

Moreover, toehold mediated strand displacement mechanism permitted to develop novel enzyme free nucleic acid amplification circuits for different diagnostic detection strategies, such as entropy-driven catalysis (EDC) circuits, seesaw gates, catalytic hairpin assembly (CHA) reactions and hybridization chain reactions (HCR)²⁷⁸. In such circuits, single-stranded nucleic acid inputs produce refolding of kinetically trapped substrates via toeholds and strand exchange reactions, thus enabling conditional molecular interactions. Outputs of EDC, CHA and HCR are constituted of independent ssDNA, multiple duplexes and concatemers of increasing length respectively, and they can thus easily be coupled to different analytical modalities, with signal transduction characteristics that are suitable for diagnostics especially when the concentrations of input molecules are low. They include transduction to fluorescent, luminescent, electrochemical, enzymatic activity via DNAzymes, and colorimetric signals²⁷⁹. Researchers have already been able to use *in vitro* DNA circuits to amplify signals and detect RNA, proteins and small molecule analytes using different reporting methods combined in a plug-and-play way²⁸⁰. This methods have provided new paradigms for the design of enzyme-free biosensors for point-of-care diagnostics²⁷⁹. CHA and HCR have been developed and adapted into novel diagnostic tools, where they showed improvement in sequence-specific detection of amplicons generated by enzymatic amplification²⁸¹.

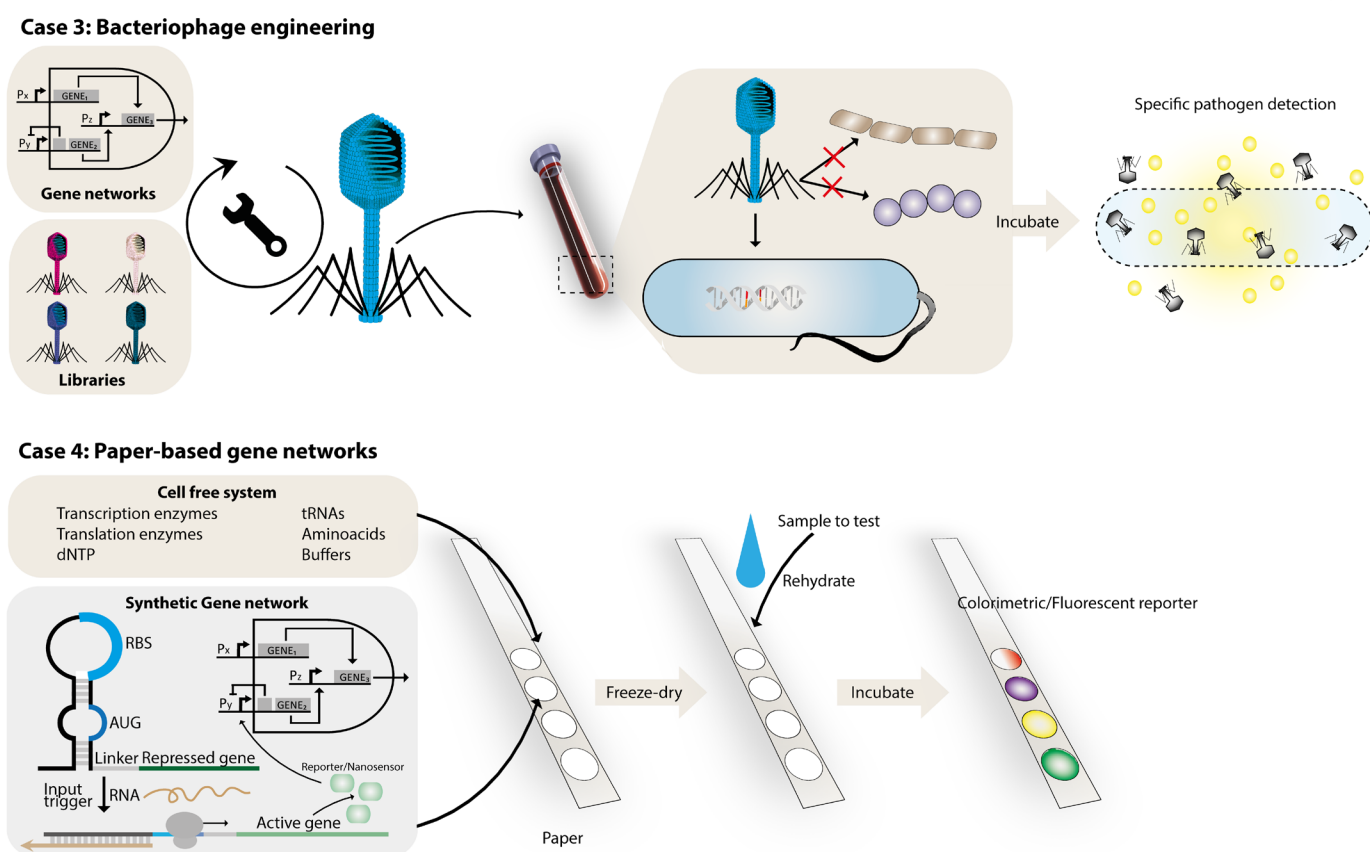


Figure 8: Case studies: recent synthetic biology research strategies to provide novel diagnostic tools. Case 3: Engineering bacteriophages as near-real-time microbial diagnostics by using them to transform target specific viable bacteria into factories for detectable molecules (adapted from Lu *et al.*⁹⁹). **Case 4:** Toehold RNA switches biosensors, *in vitro* paper-based platform that provides an alternate, versatile venue for synthetic biologists to operate and a medium for the safe deployment of engineered gene circuits beyond the lab. Commercially available cell-free systems are freeze dried onto paper, enabling the inexpensive, sterile, and abiotic distribution of synthetic-biology-based technologies for the clinic (adapted from Pardee *et al.*¹⁰³).

For example, CHA demonstrated improved signal-to-background ratio, while providing several hundred-fold amplification within a few hours detecting less than 10 copies/ μl of a target sequence. Compared to conventional enzyme based amplification reactions, CHA provided high sequence specificity and false-positive signals arising from non-specific binding to templates was greatly suppressed. In another example, a non-nucleic acid small molecule analyte, lead, could be detected with sensitivities approaching ~ 10 pM, which was 4 orders of magnitude better than the previously reported biosensors without amplification. In another study, CHA amplification reduced by more than 2 orders of magnitude the detection limit for thrombin aptamers to 20 pM, a sensitivity comparable to conventional ELISA. These nucleic acid circuits showed also capable of improving conventional immuno-assays methods. Immuno-HCR strategies notably increased the sensitivity of carcinoembryonic antigen detection cytokines and chemokines, as well as performing multiplex analysis²⁷⁸. HCR reactions can also be used for a detection of protein biomarkers²⁸², as well as an imaging tool, and proved extremely useful to enhance signals from *in situ* hybridization and for imaging mRNA expression *in vivo*^{283 284}. These methods demonstrated

high sensitivity and specificity but also great versatility and could be readily programmed and adapted to different applications. In addition, some nonenzymatic or enzymatic cascades could stand as potential alternatives for polymerase chain reaction in terms of sensitivity. However, the timeframe in which these amplifier circuits generated an output is situated between 2 and 50 hours²⁷⁸, a delay that could still prevent usage in specific diagnostic set-ups. All these strategies can be coupled to develop complex biosensing modalities. For example, extensive efforts were directed to apply the enzymatic and nonenzymatic nucleic acid cascades for amplified sensing and logic gated detection of nucleic acids and aptamer substrate complexes. Analytical advantages of cascaded amplification and sensing include: isothermal conditions, no requirement in terms of special instrumentation, generation of human readable colorimetric signals, and increased versatility. They could thus be amenable for or point-of-care diagnosis or extended diagnostic modalities.

Integrating medical algorithms into DNA circuits for disease diagnosis has been performed to tackle different pathologies such as infectious diseases, cancer, or metabolic disorders. Clinically relevant biomarkers can be detected as inputs to nucleic acid circuits via riboswitches or aptamers that translate the recognition to DNA/RNA conformational change, which triggers a computation process following a diagnostic algorithm. Nucleic acid circuits originated from efforts to develop nucleic acid computation, and besides signal amplification they have other properties that prove useful in diagnostic assays. Nucleic acid circuits are particularly capable of implementing decision making algorithms by including logic gates, thresholding and bandpass elements, and as such be useful for background suppression and noise reduction, to provide novel diagnostic devices. For example, autonomous molecular computers have been engineered to distinguish pathological states, by integrating the detection of disease biomarkers such as mRNAs, miRNAs, proteins, and small molecules into a programmable detection algorithm²⁸⁵. In addition, the advantage of nucleic acid circuits is that they can be scaled up and extended to encompass basically any diagnostic agendas²⁷⁵, as highly complex sensing and computing circuit can be needed to assess complex pathophenotypes and achieve quantitative discrimination between healthy and disease states with high resolution. Such autonomous complex circuits with the capability to recognize patterns of molecular events, make decisions and respond to the environment have already been successfully developed, for example by mimicking neural network computation with considerable power²³.

Cell types, both healthy and diseased, can be classified by inventories of their cell-surface markers using aptamers and nucleic acid circuits²⁸⁶. In a recent approach, You *et al.* developed DNA nanorobots for programmable analysis of multiple surface markers to enable the phenotype profiling on whole cells. They engineered a device combining structure-switching DNA aptamers with toehold-mediated strand displacement reactions to perform autonomous Boolean logic-based analysis of multiple cancer cell-surface markers with production of a diagnostic signal, associated with a targeted therapeutic effect¹⁰¹. In a similar approach, Rudchenko *et al.* engineered a molecular automata capable of scanning lymphocyte surfaces using a combination of antibodies and DNA circuits to assess the presence or absence of cell surface markers on living human cells²⁸⁷.

Nucleic acid diagnostic devices have proved capable of operating in solution but also on solid surfaces such as paper^{288 289}. The use of transcriptionally generated RNA circuits along with post translational components as transducers might further simplify the production of nucleic acid circuits for point-of-care applications: instead of producing, purifying and storing multiple kinetically trapped nucleic acid substrates, double-stranded transcription templates could be used to generate these circuits *in situ*. For example, Pardee *et al.* recently developed toehold RNA switches integrated on paper-based biosensors that provide an alternative and versatile platform for synthetic biologists (**Figure 8: Case 4**). This format enables the safe deployment of synthetic gene circuits beyond the laboratory. In this approach, they proposed that commercially available cell-free systems freeze-dried on paper could enable the inexpensive, sterile, and abiotic distribution of synthetic biology DNA-based biosensing technologies for the clinic. They demonstrated this technology with the detection of small-molecule and nucleic acids, prototyping of complex gene circuits for programmable *in vitro* diagnostics, including glucose and strain-specific Ebola virus biosensors¹⁰³.

Synthetic nucleic acids can be also used as probes in higher order structure constituted of amplifying probes. For example, branched DNA assays, in which alkaline phosphatase labeled nucleotides bind branched DNA structures (bDNA) generating a chemiluminescent signal, have shown to increase the specificity of conventional assays, such as the VERSANT assay (Siemens healthcare, USA). The more accurate, automated, highly sensitive and broad dynamic range of bDNA assays, have proved them useful for the diagnosis, prognosis, monitoring of viral load, and effect of HIV, HCV and HBV antiviral therapy, when variability associated with the PCR assay made it less useful for monitoring patients on antiretroviral therapy^{290 291}.

Similarly, other architectures using synthetic nucleic acids probes have been described, such as DNA hydrogels biosensors: ssDNA sensing devices made of hybrid DNA-hydrogel respond to stimuli by altering shape and swelling properties after toehold-mediated DNA displacement reaction. This strategy has been implemented for the detection of various chemicals or proteins^{292 293}. Algorithmic control on assembly and operation of DNA nanostructures and machineries²⁹⁴, have also yielded synthetic molecular machinery from DNA, or nanomachines that can be activated by interactions with specific molecular signals or by changes in their environment²⁹⁵. For example DNA origamis were proposed to be assembled into logic-controlled nanomachines capable of autonomous *in situ* diagnosis and therapy delivery^{100 296}, or stand-alone biocomputers capable of *in vitro* diagnosis²⁹⁷. In the first example, switchable DNA nanocapsule closed by DNA strands hybridized to aptamer sequences could open upon recognition of certain cell surface proteins (**Figure 10**). More recently, following an *ex vivo* prototyping phase, this approach was successfully transitioned DNA origami robots operating in living cockroaches and is now being evaluated for patient use in a clinical trial²⁹⁸.

Furthermore, orthogonal nucleic acid chemistries have also been proposed as new tools for diagnostics development. Novel synthetic nucleobases and their genetic polymers, known as XNA (xenonucleic acids) increase the chemical and structural diversity of nucleic acids, and open up the way for increased affinity and stability against enzymatic cleavage, expanded functionality such as enzymatic activity, and improved synthesis and selection procedures^{299 27}. For instance, selection experiments against two human target proteins, VEGF and IFN- γ yielded XNA

aptamers that bind with affinities that are >100-fold improved over those of aptamers containing only natural bases³⁰⁰. Other authors developed nanomolar to subnanomolar affinities to clinically relevant protein targets including PDGF and pro-inflammatory cytokine IL-6³⁰¹, or small molecules such as camptothecin³⁰². Recent studies also demonstrated the advantage of using XNAs detection probes in biological fluids, particularly because they permit to achieve significant improvement in stability by providing resistance to nucleases. For example, expanded nucleic acids aptamers showed promising properties as probes for *in vivo* tumor imaging. These authors developed a novel locked nucleic acid (LNA)/DNA chimeric aptamer probe that showed a great improvement in performance and serum stability compared to conventional aptamers³⁰³.

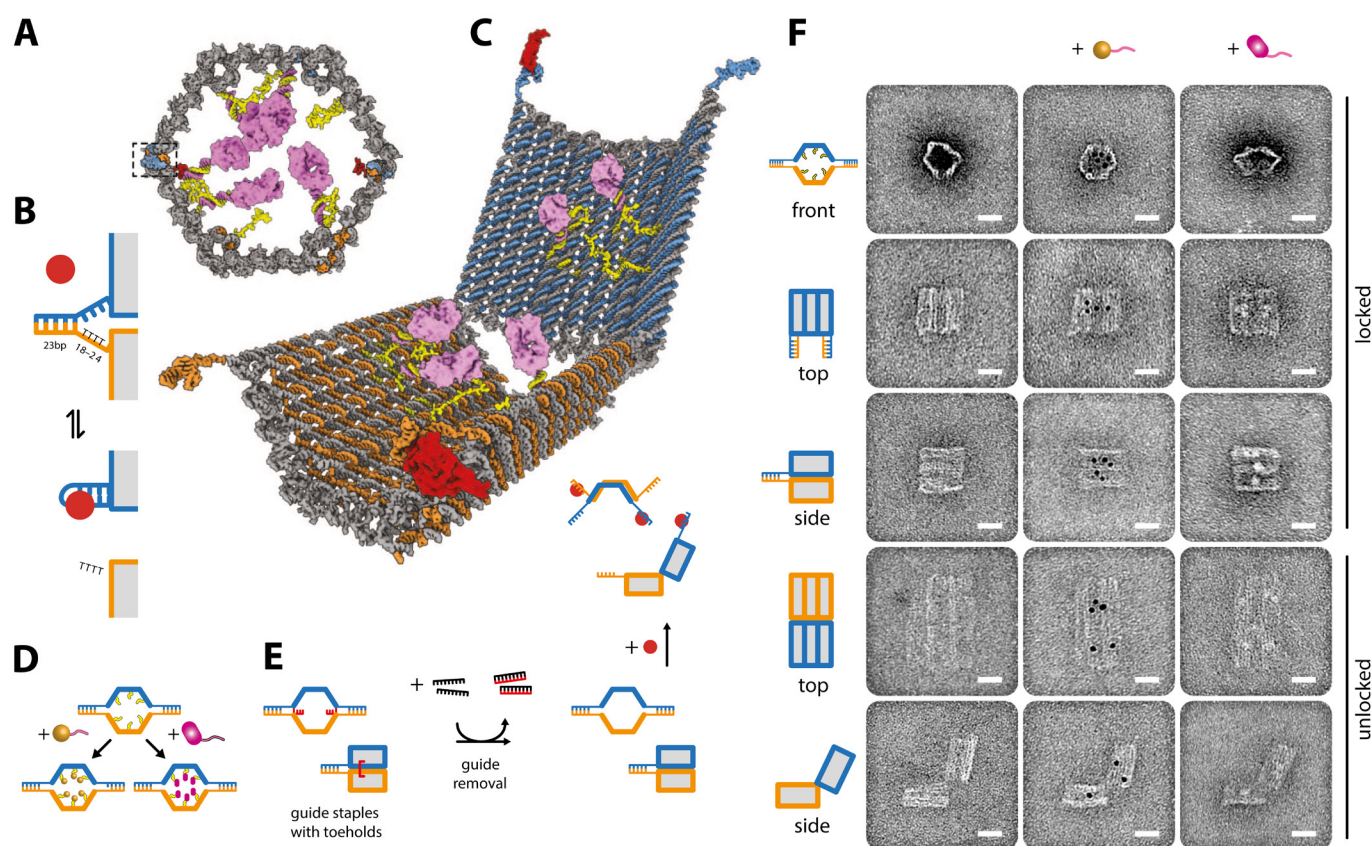


Figure 10: Autonomous, self-assembling, aptamer logic gated DNA Nanorobot for targeted transport of molecular payload. (A) Schematic front orthographic view of closed nanorobot loaded with a protein payload. Two DNA-aptamer locks fasten the front of the device on the left (boxed) and right. (B) Aptamer lock mechanism, consisting of a DNA aptamer (blue) and a partially complementary strand (orange). The lock can be stabilized in a dissociated state by its antigen key (red). Unless otherwise noted, the lock duplex length is 24 bp, with an 18- to 24-base thymine spacer in the nonaptamer strand. (C) Perspective view of nanorobot opened by protein displacement of aptamer locks. The two domains (blue and orange) are constrained in the rear by scaffold hinges. (D) Payloads such as gold nanoparticles (gold) and antibody Fab' fragments (magenta) can be loaded inside the nanorobot. (E) Front and side views show guide staples (red) bearing 8-base toeholds aid assembly of nanorobot to 97.5% yield in closed state as assessed by manual counting. After folding, guide staples are removed by addition of fully complementary oligos (black). Nanorobots can be subsequently activated by interaction with antigen keys (red). (F) TEM images of robots in closed and open conformations. Left column, unloaded; center column, robots loaded with 5-nm gold nanoparticles; right column, robots loaded with Fab' fragments. Scale bars, 20 nm. (Adapted from Douglas *et al.*¹⁰⁰)

These strategies showed that chemically expanded genetic alphabets can yield aptamers with greatly augmented affinities and stability, suggesting the potential of synthetic XNAs as a powerful tool for creating novel, highly functional nucleic acids. Orthogonal nuclease-resistant version of nucleic acids amplification reactions systems and probes, for example based on L-RNA molecules, were also described to have gained increased robustness²⁶⁹. It constitutes an alternative approach that has been applied for example to the autocatalytic aptazymes to construct enzyme entirely from non-natural L-ribonucleotides²⁶⁹. The mirror-image enzyme behaves identically as the D-RNA, but has gained complete resistance to ribonucleases.

Future advances in synthetic biology methodologies for the synthesis, characterization and evolution of synthetically augmented genetic polymers should help resolve numerous arising clinical questions, as well as providing fully programmable substrates for biosensing and molecular computing. XNAs technology is also likely to provide a growing bioengineering toolbox of biochemical encoding and manipulation of biological information, while also enabling to fully exploit their expanded range of physicochemical properties, orthogonality, and biostability. Additionally, *In vivo* circuits operation could further benefit from the use of orthogonal nucleic acid chemistries or even expanded nucleic acid alphabets

1.2.3.2.2 Protein based systems

Proteins are versatile and modular tools that operate naturally as near real time effectors, and have been widely used in many biomedical applications. At the molecular level, many biological response functions are allosterically regulated protein functions that couple an input to an output. Compared to nucleic acids that have limited functional diversity, and gene circuits that are intrinsically slower, the kinetic properties as well as the possibility to implement almost all biological functions: sensing, catalysis, signal processing, memory, among others, define polypeptides as powerful substrate for synthetic biology³⁰⁴. Post-translational tools defined as amino acids and their polymers offer a vast engineering playground for synthetic biologists^{305 246}. Thus, protein based biosensors provide attractive tools for the real time monitoring and control of molecular events in complex biological environments. However, their rational and systematic bottom-up engineering is often more delicate and error-prone than with nucleic acids. Although protein based strategies remain hindered by the difficulty to tailor signal transducers and receptors that can be readily compiled into defined diagnostic circuits, a true engineering approach for the design of protein sensors and circuit devices with standard functional and structural protein modules that sense, process, and amplify specific molecular signal of clinical interest, is emerging³⁰⁶.

Protein–ligand interactions are part of almost every biological process and are of tremendous importance in diagnostics. However, current protein based sensors are still largely based on single probes often isolated from naturally occurring proteins. Many synthetic biology approaches have thus tried to manipulate protein interfaces to enhance diagnostics performances and have enabled the development of new probes with improved capabilities in regard to straightforward

integration in on-purpose formats, coupling of effector functions, robustness in biological samples, and specificity and sensitivity, among others.

Antibodies have been the long lasting paradigm of binding proteins with desired specificities and high affinities, but they have intrinsic limitations related to their molecular properties: large, bivalent, multidomain protein, dependence on disulphide bonds and complex glycosylation pattern, poor heat stability, and are difficult and expensive to manufacture. In recent years, engineered versions of antibodies and even orthogonal binding schemes have entered successfully translated towards clinical application. In addition, new synthetic approaches for further improvements are likely to accelerate translation of novel protein probes and sensors. For example, the ability to conditionally direct antibodies could prove extremely useful. In a recent study, Gunnoo *et al.* showed that they could engineer antibodies displaying gated binding through site-specific, chemical phosphorylation of a recognition domain³⁰⁷. This gated binding could perform Boolean logic operations, such as induction in an enzyme-AND-antigen conditional manner (**Figure 11F**). In this case the simultaneous expression of a cell surface antigen and secreted enzyme were used to conditionally generate binding function. This strategy permits to generate antibodies active only in the presence of specific biomarker inputs of different nature to enhance diagnostic precision.

Immunodetection can also be engineered to integrate environment cues, or also provide straightforward manipulation of sensor binding characteristics by the user. For example, pH gated antibodies have been recently developed by Strauch *et al.*³⁰⁸. They described a strategy to design pH-dependent protein interfaces and showed that they could design a protein that binds antibodies in a pH dependent way. This could prove extremely interesting for antibody affinity purification and certain diagnostic formats. This approach demonstrated how protein engineering could increase versatility and efficiency off conventional diagnostic reagents. Alternatively, manipulation of synthetic antibody genes could allow for the creation of new immunoglobulin devices for novel detection frameworks, such as multi-specific antibodies, that are already moving towards diagnostic applications³⁰⁹.

Directed evolution of proteins as enabled by synthetic biology, is a powerful and versatile bioengineering solution for selecting proteins with desired functionalities³¹⁰. Site-directed mutagenesis creates libraries of rationally designed protein variants that can be screened, to allow quick understanding of protein structure and its effects on function while looking for enhanced forms, all in one experiment. It has been extensively used, either alone or in combination with other methodologies such as computational design, to generate useful probes and diagnostic reagents³¹¹. For instance, a recent study presented a method they called antibody diagnostics via evolution of peptides to evolve diagnostically efficient peptides for *de novo* discovery and detection of serum antibody biomarkers without knowledge of disease pathophysiology. As pathological antibodies repertoire are known to change in diverse diseases, this methods has proven useful to create diagnostics for early disease detection, stratification, and therapeutic monitoring, and enabled effective identification of a critical environmental agent involved in celiac disease³¹².

Meanwhile, a new generation of sensor proteins has been described, derived from small and robust non-immunoglobulin scaffolds that can be engineered with defined binding functions

using the methods of combinatorial protein design, and assembled with modular composability (**Figure 11**). As shape complementarity is an important part of molecular recognition, the capacity to precisely tune the shape of a binding scaffold to match a target of interest enables the generation of high-affinity protein based diagnostics³¹³. Many protein scaffolds have been proposed and consolidated as smaller sets capable of multiple targeting and operation in different settings as diagnostic reagents, such as engineered affibodies, adnectins, anticalins, or DARPins⁸¹³¹⁴ (**Figure 11C, 11D**). They combine the binding properties of antibodies with improved properties such as small size, high stability, absence of cysteins, high yield bacterial expression and the possibility of building higher order and multispecific constructs.

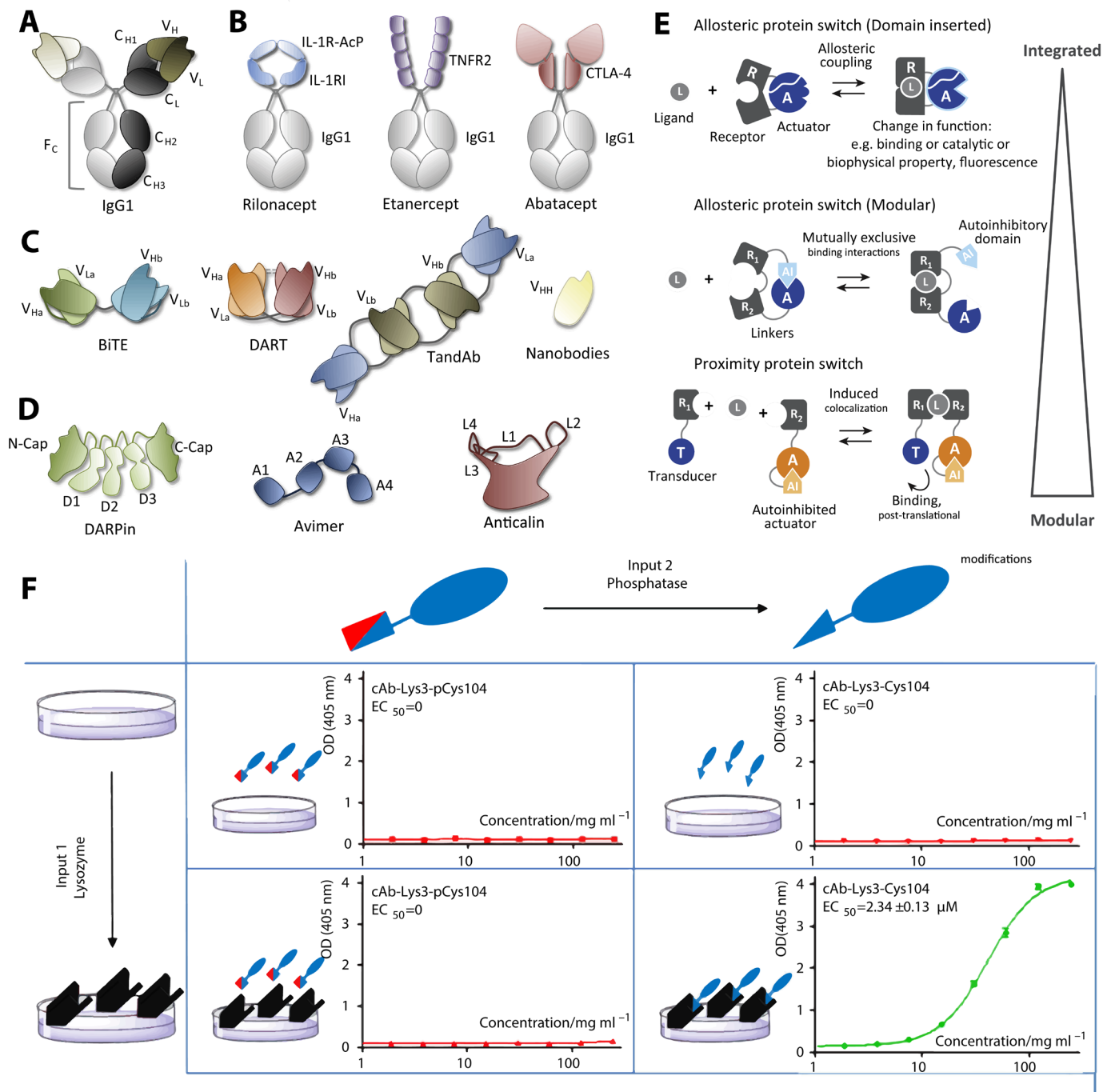


Figure 11: Example of protein based synthetic biology strategies to diagnostics. (A, B, C, D) Synthetic proteins related to immune binding functions. (A) Structure of a bispecific IgG1. IgG1 comprises four polypeptide chains that are linked by disulfide bridges. The variable domains (V_H and V_L) bind antigens, and the constant domains (C_L, C_{H1}, C_{H2}, and C_{H3}) form the rigid backbone. (B) Synthetic proteins are chimeric molecules linking IgG1 to target-binding proteins. Riloncept, etanercept, and abatacept comprise the Fc region of IgG1 coupled to the interleukin-1 (IL-1) receptor type 1 (IL-1R1) and the IL-1R accessory protein (AcP), tumor necrosis factor receptor 2 (TNFR2) or cytotoxic T lymphocyte antigen- 4 (CTLA-4), respectively. (C) Antibody derivatives: Bispecific T cell engager (BiTE), dual affinity retargeting technology (DART) antibody, tetravalent tandem antibody (TandAb), and nanobody (D) Antibody mimetics are small scaffold proteins with high affinities. Designed ankyrin repeat proteins (DARPs) consist of 3–5 fully synthetic ankyrin repeats (D1 to D3/D4/D5) flanked by N- and C-terminal Cap domains. Avimers comprise multiple A domains (A1–A4), each with low affinity for a given target, resulting in high-affinity proteins. Anticalins are based on the rigid structure of lipocalins, with four accessible loops, the sequence of which can be randomized. (Adapted from Geering *et al.*⁸¹) (E) Principle designs of protein-based switches increasing in modularity. (Adapted for Stein *et al.*³¹⁹) (F) A synthetic gated two-input activation of an AND antibody. Input 1 is presentation of the corresponding selective antigen (black shape, here lysozyme) that engages as a ligand for the Complementarity Determining Region (CDR) in its unblocked state. Input 2 is the presence of the enzyme that unblocks that CDR by removing a blocking group (red shape, here removed by phosphatase). Enzyme-linked immunosorbent assay data for binding shows functional ‘output’ only upon the presence of both inputs (green). All other input states fail to generate activity (red). (Adapted from Gunnoo *et al.*²⁹⁹).

Also described as interesting post-translational strategies for controlling the flow of information in biochemical reaction networks, synthetic protein scaffolds are particularly attractive because of the modular nature of the design, and permit spatial organization of enzymes, and have thus been employed to create orthogonal interaction domains for assembly of synthetic metabolons. They have been shown to improve biochemical reactions in multi-enzyme complexes through substrate channeling³¹⁵ and programmable fine-tuning of enzymatic reaction and yields³¹⁶.

Instead of relying on natural antibody production and associated tedious methods, manipulation of biomolecular recognition between ligands and proteins can also be performed *in silico*. Computational design of proteins has successfully been extended to new folds, new catalysts³¹⁷³¹⁸, on existing scaffolds³¹⁹, and even non-natural reactions³²⁰ with defined specificities and affinities³²¹³²². Computational design of proteins enables the systematic engineering of binding sites, protein structure and function³²³. A decade ago, Looger *et al.* presented the first structure-based computational method to redesign protein ligand-binding specificities. Multiple soluble proteins receptors binding a number of small-molecule ligands with high selectivity and affinity, such as trinitrotoluene, L-lactate, serotonin, and the nerve agent pinacolylmethylphosphonic acid have been reportedly built in the periplasmic binding protein protein¹⁵¹. These *de novo* engineered receptors can then be used as biosensors for their new ligands although the systematicity and reliability of the method has been questioned³²⁴. More recently, Tinberg *et al.* demonstrated an approach for designing *de novo* proteins that bind small molecules and use it to create specific binders for digoxigenin³²⁵. The method relies on the design of highly energetically favorable, defined interactions with the ligand in customizable protein scaffolds. Binding was further mapped using library selections and deep sequencing, and enabled to optimize affinity to a picomolar level, comparable to conventional antibodies. Moreover, the selectivity for digoxigenin over the related steroids digitoxigenin, progesterone and b-oestradiol, could be rationally programmed by manipulation of rational design of hydrogen-bonding interactions. The authors also found that these synthetic sensors had increased stability for extended periods at ambient temperatures, and could be expressed at high levels in bacteria, properties that provide a more robust and cost-effective alternative compared to antibodies. Thus, these computational methods should enable the development of a new generation of biosensors and diagnostics for the detection of small molecule compounds. The computational design of protein-protein interaction, although suffering from shortcomings in current approaches, is now transitioning to reality, and recent successes show we could soon be capable of modulating, reengineering and designing on demand protein-protein interaction networks³²⁶.

Protein switches are used in natural biological signal transduction systems, and enable cells to sense, integrate and respond to a variety of molecular signals. Consequently, the re-engineering of tailored protein switches could enable real time, *in situ* detection of clinically relevant inputs. Recent progress in constructing protein-based switches is likely to define a new generation of molecular diagnostics. For instance, the engineering of ligand binding protein sensor switches has led to many interesting devices (**Figure 11E**). Protein switches and sensors can be built from simple, modular components, yet display highly complex signal-processing behavior³²⁷. Enzymes are of particular interest, as they can implement detection, signal processing and amplification and are amenable to modular engineering. Engineering of synthetic allosteric control in proteins, orthogonal protein building blocks, control of switchable protein-protein interactions or

designing switchable enzyme are thus major fields of investigation³²⁸. In cells, kinases and phosphatases are inactive by default and get switched by specific signal to be processed. Modular autoinhibition is a natural occurring form of enzymatic regulation in which autoinhibitory domains conformationally inhibit the activity of another domain within the same molecule. Covalent modifications such as phosphorylation are then capable of relieving inhibition and confer a switch like behavior to enzymatic activity. For example, Dueber *et al.* in pioneering work^{329 330 331}, explored how modular domains can be assembled to build switches with nonlinear input/output function. They integrated the autoinhibitory interaction module of the yeast kinase N-WASP with several domain-peptide interactions from unrelated signaling proteins: Src homology 3 (SH3) and PDZ peptide-ligand interactions. These authors managed to fuse constitutively interacting domain-peptide pairs to generate a N-WASP protein responsive to peptide ligands, where different combinations of input modules could produce logic gated behaviors (AND, OR) and ultrasensitive, near-digital switching dynamics with signal amplification. The same domain fusion strategy was later also successfully applied to re-engineer guanine nucleotide exchange factors³³².

Modular protein switches can also be engineered with orthogonal regulation processes. The synthetic coupling of overlapping protein domains, or domain fusion, so that small ligand, peptide or protein binding partners can then regulate allosteric activity of a enzymatic switch, have generated useful devices. For instance, ligand-sensing domains have been fused with dihydrofolate reductase, β -lactamase^{333 334 333} and Src, p38, and focal adhesion kinase^{335 336} generating estrogen analogs, maltose or rapamycin inducible versions of these proteins respectively. Sallee *et al.* developed a method to systematically construct two-domain fusion proteins using naturally occurring sequence overlaps between interacting domains, which displayed mutually exclusive binding properties to ligands³³⁷. Although still suffering from lack of standardized protocols, issues with folding unpredictability and dynamics and relying on empirical optimization³³³, the coupling with screening strategies enable to fully exploit this approach, and in the future new tools could enable the straightforward engineering of such sensor systems.

Mutually exclusive binding interactions have also been used to develop protein sensors where ligand interacting fluorescent or bioluminescent modules modulate the efficiency of resonance energy transfer^{338 104}. Recently, an interesting and innovative approach was described by Griss *et al.*, in which semisynthetic bioluminescent protein sensors with a new mechanism could be used for inexpensive point-of-care biosensors for companion diagnostics¹⁰⁴ (**Figure 8: Case 5**). This technology also known as LUCIDs (Luciferase Based Indicators of Drugs) permitted precise quantification of specific drugs in patients serum by spotting drops of clinical sample on a paper format and recording the signal using a basic digital camera. LUCIDs have a modular design and consist of 3 basic blocks: a protein-based receptor, a luciferase and a synthetic part containing a fluorophore and a specific ligand. Upon ligand binding to the receptor module, the fluorophore is maintained in close contact with the luciferase permitting efficient bioluminescent resonance energy transfer. A competing specific analyte can displace the binding and hence extinguish BRET efficiency. By measuring the ratio of light emitted from the luciferase and the synthetic fluorophore, one can quantify the concentration of the target analyte, in such a way that it does not depend on sensor concentration and signal intensity. These modular devices were

integrated on paper format to generate portable devices, and engineered for the detection of a wide range of drugs: Methotrexate, Tacrolimus, Sirolimus, Cyclosporin, Topiramate, and Digoxin. They proved efficient and accurate with human samples, and promising for the development of new generations of portable companion diagnostic assays.

Case 5: Paper-based biochemical networks : luciferase-based indicators for companion diagnostics

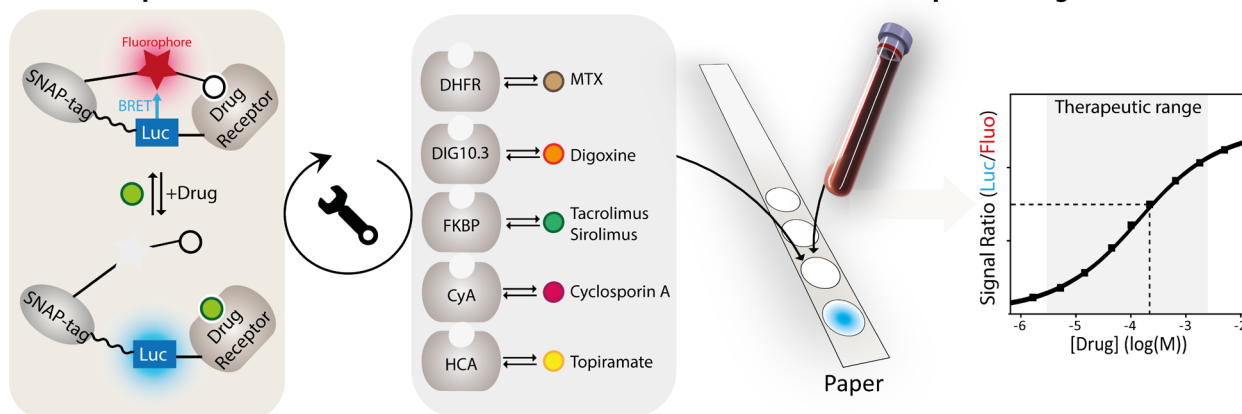


Figure 8: Case studies: recent synthetic biology research strategies to provide novel diagnostic tools. Case 5: Semisynthetic bioluminescent protein sensor approach proposed as an entirely new mechanism for inexpensive point-of-care biosensors. This enables the quantification of specific drugs in patients samples by spotting minimal volumes on paper and recording the signal using a simple camera (adapted from Griss *et al.*¹⁰⁴).

Similar sensors systems were developed that relied on complementation of luciferase fragments or on domain insertion within the luciferase structure permitting the monitoring of molecular physiology within living cells³³⁹. In another strategy, a β -lactamase fused to its inhibitor protein, and connected via a linker to a ligand receptor module, permitted to detect specific molecular cues via measurement of enzymatic activity³⁴⁰⁻³⁴¹. Additionally, Stein *et al.* recently reported a strategy for the construction of modular protein biosensors based on synthetic autoinhibited proteases whose activity can be modulated by specific proteolysis, ligand binding, or protein-protein interactions. They demonstrated that such protease-based ligand receptors and signal transducers could be assembled into different types of integrated signal sensing and amplification circuits. They relied on structure-guided design and directed protein evolution to create signal transducers and also demonstrated the modular design of an allosterically regulated protease receptor following recombination with an affinity clamp peptide receptor. They engineered high functional plasticity in protein switches, not previously observed in naturally occurring receptor systems.

De novo design of synthetic protein networks can also mimic some of the basic logic functions of the more complex in biological networks, and integrate biosensing and signal processing capabilities³⁴² (See **Chapter 3**). Enzymes can also enable the construction of biochemical circuits where they are used to implement a *metabolic logic*, in which the inputs and the outputs are enzyme substrates and products³⁴³⁻³⁴⁵. Such biomolecular logic systems for bioanalytical purposes can be designed to operate in a digital way, and process multiple biochemical information at once in cascades of biochemical reactions, to generate a final output in the form of a yes/no response, thus leading to high-fidelity decision making compared with traditional sensing devices operating

in parallel. Biochemical reaction circuits can thus be seen as the most direct and kinetically favorable way of coupling of the signal sensing with biochemical reporters. In such systems, biomarkers are biochemical entities that can interact and be processed by the enzyme network to generate a final colorimetric, fluorescent, luminescent or electrochemical output. The timely detection of complex patterns of multiple biomarkers with such biochemical systems could positively impact diagnosis and treatment of diseases³⁴⁶. This approach is fundamentally new regarding the sensor design and operation and careful attention to the biocomputing substrates and interface with other systems and electronic transducers have been explored. Enzyme-based reaction networks have further been interfaced with signal-responsive materials and electrodes and immobilization schemes have been reported for that purpose^{347 348 349 350 351}. A few examples of biochemical reaction networks of coupled enzymes implementing Boolean logic functions have been described as proof of concept to provide medical diagnostic solutions³⁵². For example, biochemical reaction networks could detect complex patterns of pathophysiological biomarkers from liver, brain, hemorrhagic shock, oxidative stress, or abdominal trauma injury^{353 354 355 356 357}, or release a drug upon sensing and integrating pathological stimuli in a complex molecular algorithm^{358 359 360} (**Figure 12**). Moreover, in order to increase confidence level of such biosensors³⁶¹, the scaling up and concatenation of enzymatic boolean logic gates (e.g., AND, OR, XOR, NAND, NOR, etc.) in networks, information storage, or threshold filters have been implemented^{362 363}. Although such strategies for the construction of tailored reaction networks still lack general robustness due to the small repertoire of enzyme and orthogonal functionality as well as the complexity and lack of knowledge on enzyme dynamics, extensive theoretical analysis has suggested ways of coping with noise and uncertainty in biochemical reaction networks^{364 365}, and computational tools for automated design, analysis and model checking are increasingly efficient and promising^{366 367 368 369}. Furthermore, coupling protein and nucleic acid-based devices can be achieved, and could generate useful devices in biological circuit engineering for diagnostic applications.

Similar to nucleic acids, the genetic code expansion for synthesis of proteins containing non-canonical amino acids is a rapidly growing field in synthetic biology^{370 371}. Synthetic amino acids could enhance stability, activity³⁷², and provide extended functionalities and overall operability of protein based diagnostic reagents. Already around 100 distinct non-canonical amino acids using orthogonal translation systems have been established, and enabled straightforward *in vivo* or *in vitro* production with synthetic post-translational modifications. This high control from synthetic genes to orthogonal post-translational machineries enables the fine design of novel protein probes with user defined properties. For instance, photocaged phospho-aminoacids have provided access to time-resolved *in vivo* measurements³⁷³, and new possibilities in site-specific fluorescent labeling provided enhanced new protein probes. In another example, Wang *et al.* described a method relying on combination of unnatural amino acid mutagenesis and selective chemical modification that offered the possibility of integrating multiple designer fluorescent labels on polypeptides. This study described the first modular method to introduce multiple probes into proteins at any genetically controlled pair of sites in proteins at physiological temperature, pressure and pH³⁷⁴. This preliminary work suggests that further expansion and applications are possible.

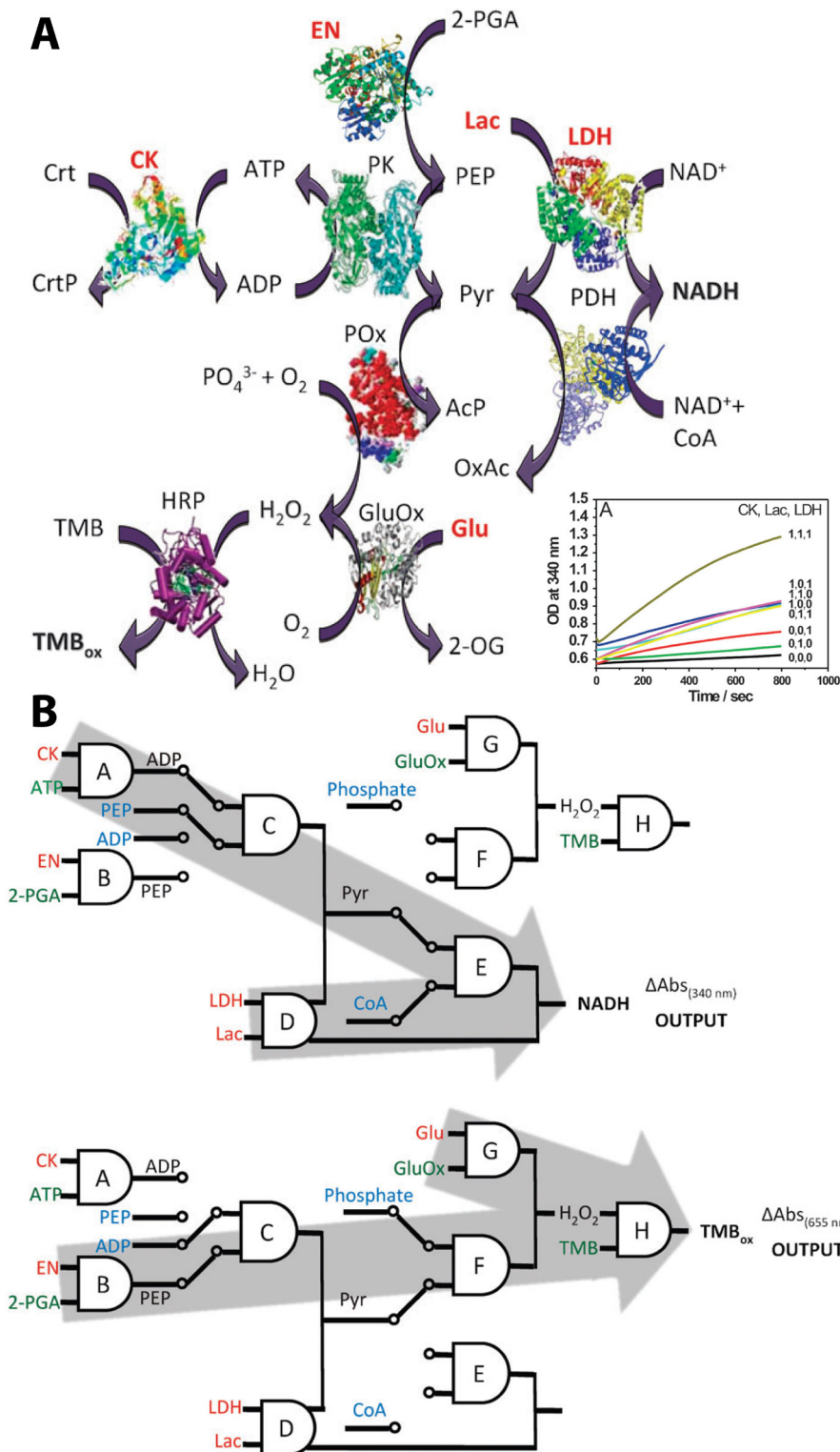


Figure 12: Synthetic biochemical reaction circuits for diagnostics.

(A) Multi-enzyme biocatalytic circuit for the analysis of soft tissues injury (STI) and traumatic brain injury (TBI). Biomarker-inputs for STI (CK, Lac, LDH) and for TBI (EN and Glu) are labeled red. Output signals for STI and for TBI are NADH and TMB_{ox}, respectively. Other products of the biocatalytic cascade are the following: acetyl phosphate (AcP), oxaloacetate (OxAc), 2-oxoglutarate (2-OG), creatine-phosphate (CrtP). Bottom right: Optical detection of the output signal (NADH) generated by the logic system operating for the STI analysis obtained upon different combinations of the injury biomarker input signals (CK, Lac, LDH).

(B) Equivalent logic schemes for the concatenated logic gates analyzing STI and TBI corresponding to the biocatalytic cascade shown above. The system switched to the STI analysis mode. Bottom: The system switched to the TBI analysis mode. Biomarker inputs are red labeled, auxiliary inputs are green labeled, and switching inputs are blue labeled. The switches-regulated pathways for the STI and TBI operational modes are indicated by arrows. (Adapted from Halamek *et al.*³⁴⁶)

The increasing ability to rationally control synthetic genes and the sequence-structure relationships enables to use proteins as potential nanomaterials with a variety of sizes and shapes and functionalities. Protein devices and assemblies can now be engineered into highly homogeneous and precisely patterned nanostructures³⁷⁵, and offer advantages over traditional nanomaterials such as carbon nanotubes, silicon or metallic particles for their low cost and straightforward production, increased biocompatibility, functionalization and interfaceability³⁷⁶. The design of protein self-assembling nanostructures and protein nanomaterial has thus emerged as promising reagents with applications in medical diagnosis. Examples involve protein nanowires, nanotubes, nanocapsules, nanopores, or hydrogels, that could show promising capabilities in biosensor design^{377 378 379 380}.

Although highly amenable for incorporation into integrated devices, protein based biosensors could have potential disadvantages regarding the storage capabilities, transport and shelf life. Translating these approaches towards eukaryotes and prokaryotes may open new avenues in protein-based biosensing and biocomputing. Moreover, direct coupling of biosensing and therapeutic activity in engineered proteins is paving the way for extremely interesting clinical applications, such as the recent synthetic glucose-responsive insulin³⁸¹. Extensive research efforts have so far yielded useful protein based sensor systems, but systematic methods for the engineering of novel devices will require further advances. Moreover, before the promises of synthetic biology approaches can be fully realized, the total connection between amino acid sequence and protein structure and function still remains to be elucidated.

1.2.3.2.3 High order assemblies and synthetic cell-like systems

Living, emergent systems rely on regulatory processes as a central feature of their biological instructions. As we have seen, numerous strategies exploited a variety of their sensing mechanisms involving biochemical pathways, nucleic acids or proteins for the design of biomolecular logic gates *in vitro* or in living cells that can be further organized in biocomputing systems to develop intelligent diagnostics.

In vitro reaction networks can thus be designed for the sensing, processing and reporting of biomarkers, by exploiting biological species and their molecular functions. However, it is also possible to exploit the more complex architecture of living systems, which can be reassembled via bottom-up design in nonliving, on purpose systems^{382 383 384}. Although most biosensing systems previously reviewed here rely on simple architecture of few components, higher functional assembly of synthetic building blocks are possible, mimicking the natural architecture of living cells and giving access to complex features of living organisms.

A key feature of biological systems is the compartmentalization of information. Complex systems have evolved ways to cope with complexity of higher order architectures through the use of compartments (See **Chapter 3**). This strategy allows parallel chemical reactions and higher-level functions to be performed efficiently and simultaneously without loss of information content. New kinds of biotechnological supports arising from advances of synthetic biology and nanoscience give the opportunity to approach, interface, engineer, and assemble components and systems at the small working scale of biology, leading to the emergence of new strategies to diagnostics. The collusion of synthetic biology and nanomaterials will be key to realizing full potential³⁸⁵. Attempting to assemble synthetic parts in compartments approaching biological-scale functional density, such systems could prove capable of assuming near-cell like behavior³⁸⁶, efficient transduction of information and energy that permit complex molecular detection, signal processing, and biochemical actuation, while being autonomous and self-powered.

From the bioengineering perspective, this strategy has been extensively used in natural cells, where the host provides the compartment, building blocks and infrastructure to allow for the execution of instructions supported by the synthetic systems, but also mostly production, expression, maintenance and amplification. In the bottom-up design approach, however, compartmentalization only supports the user defined function without further energetic, metabolic, evolutionary, and regulatory cost, hence increasing the design space. The construction of fully multipurpose, conditional biosensing devices from biological components requires dealing with natural complexity emerging from biological systems. Tackling such challenges would thus require considering the design and engineering of organized, encapsulated systems from rationally assembled components³⁸⁷. These concepts have stressed the need for compartmentalization in bottom-up synthetic biology. Encapsulating complexity is an interesting framework for the conception of integrated systems with the ability to sense and transduce signals from their clinical environment and the ability to generate new biosensing devices with control on programmability. These would be multicomponent, compartmentalized, non-replicating systems. This approach will necessarily require full expertise in design, engineering,

and characterization of membrane systems and the modeling of complex systems. These approaches have been often captured under the concept of synthetic minimal cells, which potential for biosensing and biocomputing has been widely emphasized^{388 389 390 391}.

At the moment, synthetic vesicle-based systems of submicrometer scale, operating as high density intelligent biochemical sensor/effector systems have been proposed to perform diagnostic processes in physiological environments. Combining sensing and effector functionality at the nanoscale, they generate a conditional response that depends on environmental factors such as biomarker concentrations, pH or temperature at the target site³⁹². They are basically composed of a carrier platform and a payload embedding circuitry for sensing, processing signal and reporting. Such stimuli-responsive hybrid nanostructured particles in a range of sizes from nanometers to a few micrometers include liposomes, polymerosomes, core-shell structures, nanogels, and more complex architectures. The controlled assembly of synthetic polymer structures in vesicles is now possible with an unprecedented precision and modularity³⁹³.

Synthetic vesicles have been extensively used for therapeutic strategies as drug nanocarriers, and proved efficient and successful in the treatment of diverse pathologies. Alternatively, they have also progressed toward analytical application as biosensors for bioanalysis for their ability to carry complex diagnostic reagents and electrochemical, fluorescent or chemiluminescent probes. Synthetic vesicles can also integrate synthetic biological parts such as engineered transmembrane and pore proteins, enzymes, nucleic acids or metabolites to integrate stimuli responsive behaviors³⁹⁴. Encapsulation of hydrophilic compounds in their aqueous cavities and the insertion of fragile hydrophobic compounds in membranes offer protection and stabilization from harsh physiological conditions and allowed to act *in situ*³⁹⁵. Synthetic vesicles are known to enhance biochemical reactions, as thermodynamics of synthetic reactions are known to be favored by compartmentalization in picoscale volumes³⁹⁶, stabilize enzymatic processes, and provide signal amplification. The ability to functionalize vesicle surface to perform recognition functions, and targeting, selective transport and sensing is another important aspect of their use in bioanalysis³⁹⁷. Moreover, their small scale provides the opportunity to take advantage of patterns or multimodal molecular factors of the microenvironment *in situ*. Moreover, compartmentalized processes in different segregated spatial localizations can then be put under interactions with one another and create more complex biochemical networks³⁹⁸ (See **Chapter 3 & 4**).

As the first described synthetic compartment, liposomes have been used for a wide spectrum of sensing modalities with a wide range of analytes. Many liposome-based assays have been reported such as liposome immunoassay (LIA), liposome immunolysis assay (LILA), liposome immunosorbent assay (LISA), flow-injection liposome immunoanalysis (FILIA), and cytolysin-mediated liposome immunoassay (CyMLIA), as well as chromatic polydiacetylene liposome based assays^{399 400}, providing low detection limits for analytes including hormones, viruses, bacteria, DNA/RNA segments, pesticides, tumor markers, proteins, antibodies and some drugs^{401 402} (reviewed in⁴⁰³). Liposomes with engineered biological pores have also been extensively used for nanopore-based biosensing applications. Rational modifications by directed evolution or biochemistry have been carried out to reengineer mutant channels for desired biodetection purposes. For example, α -hemolysin, MspA or FhuA, and more recently phi29 derived synthetic nanopores have been engineered for sensing a wide range of analytes, from metal ions to organic

molecules to DNA, RNA and peptides⁴⁰⁴. Further efforts have been conducted to associate these architectures into point-of-care formats.

Important efforts have been conducted to engineer devices such as orthogonal polymeric vesicles with other membrane properties for diagnostic, to protect reagents but allow them to interact *in situ*⁴⁰⁵. Polymeric vesicles structures similar to lipid vesicles can be engineered using synthetic block copolymers and stand as interesting candidates to develop orthogonal nanosystems for medical applications^{406 407}. They are more stable, more versatile, and less immunogenic than liposomes. Control over block copolymer chemistry enables tunable design of polymersome material properties. Optimization efforts allow scientists to design smart compartments encapsulating sensing and biocomputing biochemical networks made of nucleic acids, enzymes, and metabolites, and control on size, encapsulation of species, membrane properties and permeability to enhance sensing sensitivity and specificity, and allow insertion of membrane proteins⁴⁰⁸. Recent advances are shifting these active nanosystems towards smart-complex synthetic parts and polymer assemblies, like multi-compartment cascade reaction⁴⁰⁹.

Crucial to innovation in diagnostics is the development of new platforms that combine multifunctional compounds with stable, safe and implantable devices for close to patient strategies. As discussed before, theranostic strategies could decrease health burden of many pathologies by enabling the simultaneous detection and treatment of pathological events through interactions manipulated at the molecular level, by that mean achieving less side effects and timely delivering of therapy. Along with *in vitro* assays, synthetic vesicles based systems have been proposed to work as intelligent nanocarriers for theranostic. While surface functionalization enables selective targeting, theranostic nanocarriers could improve disease diagnostic and treatment because of their ability to execute conditional biological functions at targeted diseased sites⁴¹⁰. Additionally, targeted nanodelivery systems would greatly beneficiate *in situ* imaging diagnosis⁴¹¹. Such injectable systems can process pathological signals and release *in situ* specific signals and/or drugs based on analysis of multiple signals. Several types of injectable diagnostics based on vesicle systems have been proposed, such as liposomes and synthetic polymeric systems. For instance, polymersomes have proven as excellent non-invasive intelligent fluorescent probes carrier for diagnostic imaging⁴¹². Another recent study obtained success in developing a platform based on polymeric artificial organelles to target specific cells for subcellular delivery of drugs, enzymes, nucleotides, and diagnostic agents⁴¹³.

Synthetic nanobiological assemblies have been exploited to construct new diagnostic assays with increased specificity and sensitivity. Assays relying on conventional assemblies can display important sensitivities for single molecular targets, whereas the engineering of multifunctional nanoplatfoms for sensing, imaging of biomarkers can prove capable of multiplexing input detection for a more efficient discrimination between complex disease phenotypes⁴¹⁴. Self-assembled nucleic acids nanostructures can provide templates for the spatial ordered patterning of enzymes to develop high sensing efficiency and sensitivity of biocatalytic cascades for nanoscale devices. For instance, spherical nucleic acids have during the last decade constituted a major technological advance in the field⁴¹⁵. Such approaches have been used to develop for example glucose, ethanol or cocaine biosensing devices^{416 417 418 419 420}. Synthetic bionanoparticles can also perform Boolean logic operations using two proteolytic inputs associated with unique aspects of tumorigenesis⁴²¹. Konry *et al.* also reported the integration of microarray sensor

technology with algorithmic capability for the gated screening of proteins and DNA markers in a biological sample. The system they developed performed simple Boolean logic operations by coupling multiple molecular recognition inputs like IL-8 and specific genes to a fluorescence signal output⁴²². Similarly, Janssen *et al.* recently developed synthetic antibodies for molecular diagnostics that are peptide DNA conjugates, enabling the control of antibody activity in a DNA based logic gated behavior⁴²³. In another study, hybrid biochemical reaction networks exploiting enzymes and oligonucleotides with a computing functionality were applied to the identification of bacteria exhibiting multi-drug resistance. This approach enabled the identification of the NDM-1-encoding gene and concurrently to screen, by a tailor-designed biomolecular logical gate, two genetic fragments encoding the active sites bound to carbapenem⁴²⁴. A vast array of literature has covered the field of information-processing systems at the nanoscale to yield *intelligent* signal-responsive hybrid systems with built-in boolean logic^{359 425 426}.

Synthetic biopolymers have also been designed to act as biochemical stimuli responsive devices. In this approach, interaction of responsive polymers with molecular signals relies on the conjugation of polymers with biological molecules such as nucleic acids, enzymes, antibodies, and other proteins, or *de novo* molecularly imprinted polymers (reviewed in⁴²⁷) to yield diagnostic information or therapeutic activity *in vitro* or *in situ* upon systemic administration⁴²⁸. In this perspective, nanogels are likely to yield interesting diagnostic devices. Of polymeric nature, they can be tailored with a broad range of chemical modifications and entrap a large scope of biological molecules (nucleic acids, proteins and drugs). For instance, multi-functional core-shell nanogels combining magnetic regulation with biochemical sensing have been demonstrated⁴²⁹. Another approach relies on peptide-based or viral inspired self-assemblies for the design of hollow or solid peptidic nanostructures. For instance, Naskar *et al.* demonstrated how multivesicular structures built from self-assembling peptides, could display calcium ions sensitivity. Such intelligent stimuli responsive behavior could enable approaches of medically relevant biodetection⁴³⁰. Expanding peptide-based nanostructures by exploiting rationally engineered peptide functions, receptor or enzymatic activity, is likely to lead to novel nanomaterials with complex sensing functionalities. Finally, synthetic biology could provide interesting approach for the integration, the production and functionalization of metallic nanoparticles such as gold or quantum dots, which are of outstanding importance as diagnostic reagents. Synthetic biology is thus likely to provide ways to exploit new sensing and reporting mechanisms to create new tools by providing a biological interface to use metallic nanoparticles⁴³¹.

Similarly, the engineering of so called *biofuel cells* have received much attention to develop autonomous, self-powered biodetection devices. Biofuel cells emerged from the effort to engineer an interface between electronics and biology, which could benefit bioanalysis^{432 433 434 435}. They display properties that defines them as robust *in vivo* power sources for bioelectronics, and could greatly benefit the development of implantable diagnostics, such as glucose biosensors, or more complex *intelligent* devices^{435 436 437}. For example, Zhou *et al.* developed aptamer biosensors based on biofuel cells, where power release was triggered by biochemical signals processed according to the boolean logic operations, to generate self-powered medical diagnostics programmed into a biocomputing system²⁶¹. Other advances have showed the coupling of a self-

powered diagnostic operation with logic-activated drug release⁴³⁸. Combined with synthetic biology methods, such approaches could reveal valuable in producing novel tools.

Although still in its infancy, the opportunity to construct *de novo* increasingly complex processes and systems is emerging from the convergence of synthetic biology with new experimental and computational tools⁴³⁹. The ability to control the bottom-up design, synthesis and construction of synthetic systems by the direct assembly of synthetic nanoscale parts constantly increases, evolving towards cell-like complexity and capabilities for tailored biodetection. I envision that new approaches exploiting synthetic compartments encapsulating biosensing, biocomputing and diagnostic reagents are likely to generate innovative medical devices in the future, and hold enormous potential as nanostructured biomaterials for future *in vivo* drug delivery and diagnostic imaging applications⁴⁴⁰. For some of such systems, clinical trials are in progress, but extensive clinical evidence of significant patient benefit will be further required⁴⁴¹. The power of such systems can be realized with synthetic biology and bioengineering to generate functional devices for the clinics. Additionally, these approaches are likely to enhance our understanding and explore new ways of interfacing biological systems.

1.3 Synthetic biology and biological signal processing for diagnostics: to sense and to compute

Biosensing is indeed a mature application area of synthetic biology. IUPAC nomenclature defines a biosensor as “a device that uses specific biochemical reactions mediated by isolated enzymes, immunosystems, tissues, organelles or whole cells to detect chemical compounds”⁴⁴². Applied to medical diagnosis, these devices combine biological molecules as the recognition and transducing elements to provide quantitative or semiquantitative analytical data corresponding to the concentration of a specific biomarker. Interestingly, as we have seen biological systems are able to integrate various kinds of clinically relevant physical and chemical signals (nucleic acid, protein or lipid ligands, osmolarity, pH, temperature). This ability of biological systems to assess molecular pathophysiology by biorecognition of biomarker patterns is of great interest for the generation of diagnostic assays. Moreover, evolution has generated a vast natural repertoire that can be mined to retrieve useful biological functions, and synthetic biology provides tools and methods for their efficient re-engineering. In addition, biological systems have interesting characteristics for diagnostics, such as the ability to provide physiologically functional measurements, ability to perform ultrasensitive and specific response to input stimuli⁴⁴³, and integration of complex signal processing abilities. Additionally, biological systems are efficient problem-solving systems that use sensor and signal processing modules to analyze their environment relatively to their own state and compute phenotypic responses⁴⁴⁴. Indeed, the idea to engineer living organisms or their components as problem solving entities is not new^{126 445}, and molecular computers performing biological computation have been proposed for different purposes^{446 447 448}. The signal sensing event of biosensors can thus be associated to a computation process that can be engineered to integrate *compiled* medical knowledge in the form of a decision algorithm and computational versions of diagnosis using biological components have been proposed^{198 449 450}.

Information processing occurs naturally across hierarchical levels ranging from molecules to cells, tissues, organisms and even ecosystems. Computation on biological signals thus ubiquitously takes place in biological systems³⁹. Biological information is collected by sensing and signaling units, further processed and analyzed by organic matter, metabolites, proteins and gene circuits, and translated into specific molecular responses. Although biological processes are by nature noisy and use unreliable molecular devices interacting with analog and digital molecular signals, they manage to solve tasks precisely, in real time and energy-efficiently⁴⁵¹. While trading a simple for a more complex design would be counterproductive, modular oriented methodology with layered, standardized interface between sensing and reporter components can speed up the design, provide programmability and increase versatility and capabilities of engineered biosensing systems. The rationale behind such transmission devices, or signal processors, is to achieve signal integration from various sources, gain amplification, noise filtering, or logic operations⁴⁵² and to connect various input sensors to reporting platforms for output multiplexing. Synthetic biology enables the construction of tailored signal processing by means of modular plug-and-play, and thus the reprogramming of natural information processing systems either *in vivo* or *in vitro*, into autonomous nanomedical devices that perform diagnostic rules *in situ*.

In the context of diagnostics, biological circuitry needs to be easily reprogrammed to integrate varying clinical constraints, different medical agendas and a vast range of pathologies. Moreover, it needs to support the improvement in system robustness and overall medical service. Additionally, the kinetics of biological processes is to take into account to engineer clinically compliant signal processing systems for appropriate diagnostic devices, as transcriptional and translational circuits dynamics occur over timescales of minutes to hours while biochemical processes occur in seconds or less³⁰⁴. Noise propagation in synthetic systems is also to take into account to obtain reliable behavior, which is dependent on systems dynamics and scale of processing circuits. Consequently, keeping faster and simpler systems would have fewer mode of failure and overall great chances of clinical success.

The need for novel health monitoring systems has progressively opened a new domain that results from the fusion of sensors and signal processing in synthetic biological systems. Properties such as ultra-low-power information processing capabilities⁴⁵¹, self-powering, compactness from micro to nanoscale, data storage, real-time signal processing and multi-sensor communication are all important advantage for synthetic biological systems to implement integrated medical diagnostic devices. These properties enable the pre-processing and aggregating of low-level sensor physiological information to yield output signals intelligible by physicians, patients or researchers concerning diagnostically relevant events or biomarker patterns. Hence, I believe that the signal processing capabilities of synthetic systems can meet the challenge of developing portable autonomous health monitoring devices that can offer pragmatic solutions to achieve highest clinical impact, for developing countries or point of care, personalized medicine.

Critical parameters in the analytical performance of quantitative biosensing systems for diagnosis are the sensitivity (i.e. lowest analyte concentration that triggers a detectable response) and the dynamic range (DNR, range of analyte concentrations where it can be estimated precisely based on the output signal), while optimizing the signal to noise ratio (or response *fold change*). Quantitative systems provide analog signals which transfer function are ideally standardized response curve with wide DNR and low noise. Engineering biological analog detection can be performed using for example negative feedback loops. However, other qualitative or discrete, near-digital detection modalities are possible and can prove extremely valuable in specific context. For example, molecular ultrasensitive switches can provide digital behavior, providing an input detection threshold at which small changes in input biomarker concentration lead to large changes in output signal. Strategies involving positive feedback can be used to obtain digitization of signals. Cellular systems can also display fold-change detection, a response whose entire shape, including amplitude and duration, depends only on fold changes in input and not on absolute levels⁴⁵³. A wide class of mechanisms has shown to display this response, which could prove useful for biodetection. Another property to consider when designing signal processing devices for diagnostic application is robustness, that is, the ability of a system to tolerate exogenous perturbations while limiting modes of failure in biodetection. Achieving modulation of transfer function of synthetic systems is thus of particular importance for the clinics (**Figure 13**).

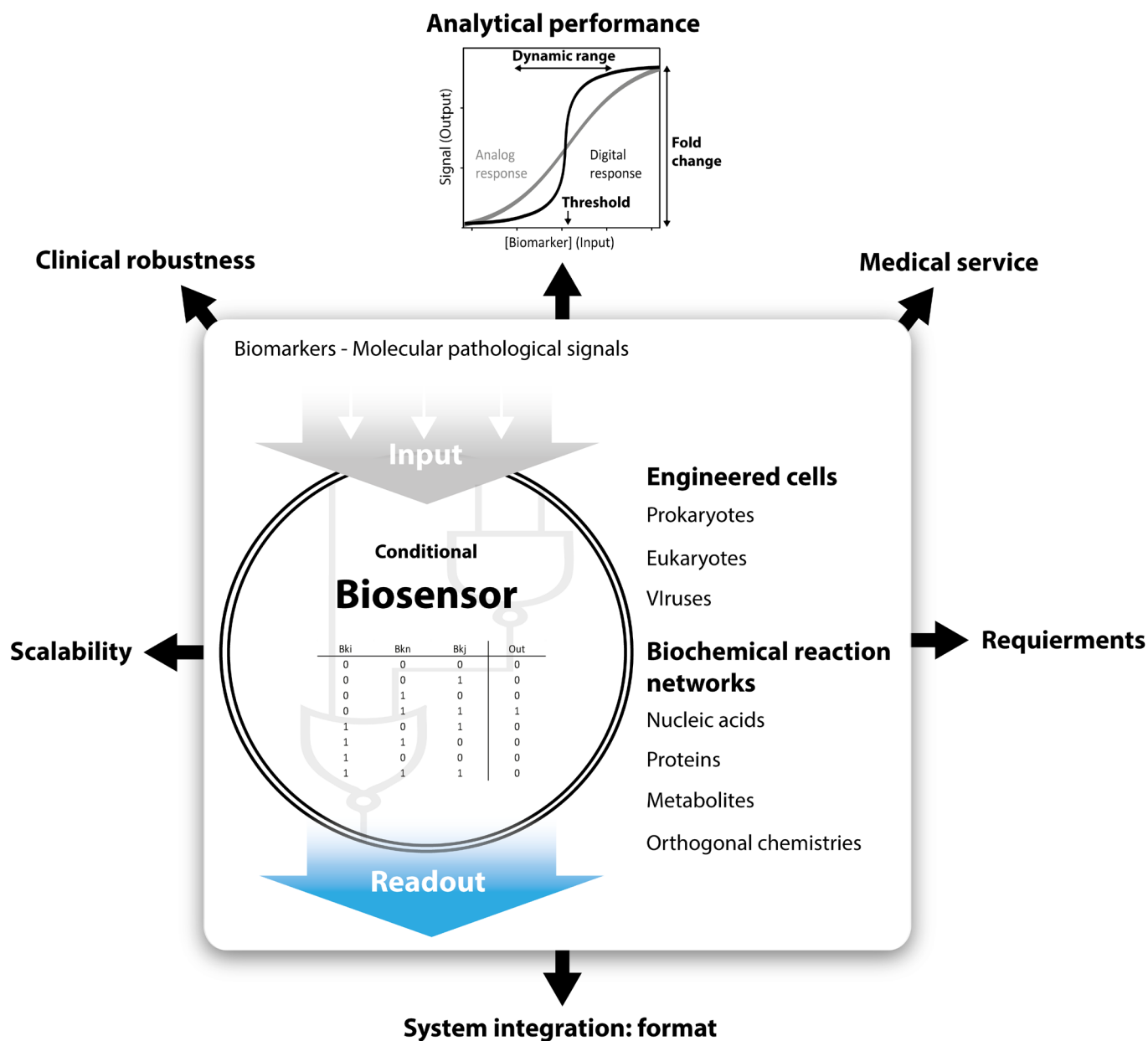


Figure 13: General considerations on constraints and architecture of synthetic biosensors for medical diagnosis. Biosensing devices conditionally generating a readable signal upon presence of specific patterns of pathological biomarkers can be synthesized using natural or synthetic components such as engineered cells, or biochemical reaction network. Such systems can be developed from the top-down perspective using modular biological parts assembled *in vivo* to generate useful synthetic phenotypes, or be assembled *ex-vivo* from a bottom-up perspectives. In order to obtain translational success, important constraints are to take into consideration in early design phases.

Crosstalk between biological and synthetic circuitry must be in most cases prevented, while some signals need to be combined, added or compared to enable decision making. This requires computation processes to be implemented in synthetic molecular devices. Biological information can be transformed through digital or analog processing or through a hybrid combination of both. A digital mode of operation has the advantage to enable the implementation of Boolean

logic based decision making circuits. In that perspective, synthetic biology attempts to apply the digital paradigm of electronic engineering to develop algorithmic processes with biological components. Many examples of synthetic biological signal processing have been achieved by manipulating digital information. Although more difficult, the rational design of analog based processing is appearing as extremely valuable to exploit the computational power of biology, as it could cope with more complex operations and larger sets of variable in smaller circuits, and is closer to the natural mechanism of biological systems⁴⁵⁴. Moreover, analog computation could for example enable pathological signal normalization, for instance directly computing ratios with physiological standard biomarkers like creatinine or albumin. I propose that an efficient and accurate signal processing approach to synthetic biological networks would integrate both analog and digital processing to achieve versatility, efficiency and reliability. Recent devices have been recently engineered in that direction, to perform analog to digital or digital to analog processing⁴⁵⁵

456 457 458 459

The aim of synthetic biology is to achieve systematic on purpose programming and tuning of these analytical characteristic for different biodetection agendas using modular signal processing circuit design. An increasing number of strategies have been developed for tuning the responses of biological systems. I believe it is on the way to enable the tuning of biological systems' transfer function in analogy to electrical signal processing.

1.3.1 Synthetic biological circuits operating *in vivo*

The first successful design and implementation of synthetic gene processing circuits were demonstrated with the genetic toggle switch and the repressilator^{16 17}. They proved that bioengineering-based methodology could enable the integration of user-defined information processing and computing capabilities in living cells. Following these studies, synthetic biologists have successfully established a repertoire of genetic components to engineer complex signal processing genetic circuits in living cells with a vast range of functionalities, such as switches, oscillators, timers, memory, filters, logic gates, cell-cell communicators, or buffers^{460 461}. Since, synthetic gene networks have been extensively used to reprogram cells for useful task such as decision making for cell-based biosensors¹¹⁵. Increasingly complex designer signal processing networks have been built in cellular systems to perform input-triggered genetic instructions with precision, robustness and computational logic. Moreover, the demonstration of the ability to rationally tune *in vivo* biological transfer functions in transcriptional, post transcriptional, and post-translational levels of regulation has been extensively described (review in⁴⁶²).

For future medical and diagnostic agendas, complex processing circuits operating robustly in living cells may require new types of orthogonal parts with increase in orthogonality to host physiology, that offer control on dynamic range, digital or analog signal modes, low crosstalk, and design versatility. Next-generation gene networks for biodetection could as such comprise tunable filters and noise controllers, analog-to-digital and digital-to-analog converters, or even adaptive learning networks⁴⁶¹. Moreover, the systematic design and quantification of genetic parts

in context⁴⁶³ is leading to a new era of well-characterized regulatory synthetic genetic devices, such as bicistronic RBS⁴⁶⁴, ribozyme parts insulators⁴⁶⁵, and synthetic terminators⁴⁶⁶. Parts mining and computational design, and directed evolution are further expanding the number of regulators that can be used together within one cell⁴⁶⁷.

Indeed, complexity of signal processing circuits *in vivo* is often limited to a few logic operations, because of unpredictable biochemical crosstalk occurring in the confined volume of the cell and the limited number of available parts, the size of signal processing circuits and composability has remained limited. Developing design strategy for the successful layering of orthogonal high performance parts or logic gates into large, integrated circuits in single cells remains a challenge. In a recent study, Moonet *al.* managed to overcome this challenge by applying part mining and directed evolution to build a set of orthogonal transcriptional AND gates in *E. coli* that could then be concatenated into complex programs, such as 4-input AND gate that consists of 3 circuits that integrate 4 inducible systems, thus requiring 11 regulatory proteins. Optimizing and refining the performance of individual gates was sufficient to predict the behavior of a complete program²².

New design concepts have recently taken a new step with the development of digital recombinase based circuits. For example, in the work carried on during this thesis (Courbet *et al.*⁹⁶, see **Chapter 2**) we found that promoters of clinical interest and control circuits that coordinate simple signal transduction showed inherent noisy and unpredictable responses with limited control over specificity and efficacy in host cells when operating in complex media. In fact, a known barrier to predictability in design is context⁴⁶⁸. Synthetic gene circuits are often easily perturbed and their behavior altered by the environment they are exposed to⁴⁶⁹ and the host they are integrated into. Heterologous pathways have not had the advantage of long periods of co-evolution with other cellular substrates. Thus, their function often suffers from uncontrolled/unpredicted interactions with the surrounding cellular context and environment. Lack of robustness has limited the utility of engineered gene circuits for further medical applications and hinders advances in synthetic biology. In our recent work, we proposed that context sensitivity can be reduced by incorporating synthetic genetic tools precedently developed, while keeping few components for fewer modes of failure and increased safety and likelihood of approval of cell-based biosensors in medical setting⁴⁷⁰. In order to buffer matrix effects and nonspecific environmental interferences, overcome variable part performance across changing complex media⁴⁷¹, and enable predictable and standardized translational coupling, we incorporated in our design (i) Expression Operating Unit (EOU)⁴⁶⁴, (ii) a ribozyme insulator part, RiboJ⁴⁶⁵ (iii) Digital gene switches and integrase logic gates⁴⁵⁸. Digitalizing along with amplifying and multiplexing input signals improves fidelity, sensitivity, mediate sharp response profiles and ensure robust biochemical processes. Bonnet *et al.* recently designed a new type of logic gates architectures which recapitulate all conventional logic functions using integrases Bxb1 and TP901^{458 459}. These integrase logic gates enable truly digital and discrete response, compared to previous systems that produces intermediate expression levels^{197 102 448}. This property makes them highly relevant for medical applications, and particularly in diagnosis as it is often threshold based. This system also embeds a built in memory capacity which enables the recording of weak or transient signals while giving a constant amplified output. Compared to transcriptional switches⁴¹, this *true* memory has non-existent metabolic cost and is stably written in either

chromosomal or plasmidic DNA, and could be addressed after extended periods of time and lysis of the bactosensor in clinical samples. These signal processing devices based on an engineered modular genetic logic gate have the advantage of high composability to be recombined for the programming of various medical algorithms. I suspect that these characteristics will be important to enable robust detection and computation in the context of intracellular and environmental fluctuations.

Taking synthetic parts improvement further, recombinases based systems have intrinsic properties that offer tremendous interest to develop cell-based biosensors: increase in scalability to larger networks by reducing their molecular payload, exert tight control and prevent cross-talk with off-target contrary to other DNA-binding proteins and control on genetic circuit in time-dependent fashion⁴⁷². Recently, Yang *et al.* extended the programmable memory capacity in a living cell to beyond 1 byte of information using 11 orthogonal integrases. A high number of events can thus be sensed, recorded and recalled at a later stage of the computation, thus increasing memory capacity could enable new type of biosensing to be performed in cells²¹.

Moreover, expanding the repertoire of available orthogonal genetic parts remains a challenge, particularly since digital logic requires many parts and will hinder the scalability of circuit design. Analog circuits constitute an attractive alternative as they can compute high order non-boolean functions such as amplification, addition, multiplication and integration, and could be regarded as a promising way for future designs for *in vivo* computations systems applied to diagnosis. Along with digital circuits, synthetic analog gene circuits have been engineered to execute complex computational functions in living cells have been recently examined theoretically and experimentally³⁹ and have recently demonstrated their value^{473 474 475 476}. Daniel *et al.* were capable of implementing analog circuits to straightforwardly compute arithmetic functions without necessitating layered digital logic gates. They demonstrated a wide dynamic range relying on positive feedback loops, which could perform or log-domain sensing, power law and addition or division of input molecular signals. Analog computation recently enabled the recording of sums of molecular events over a time period. Interestingly, ratiometric calculations are useful in diagnostic systems, because they enable the normalization of diagnostic threshold, comparisons between biomarker levels and complex decisions. This approach could provide wide dynamic range biosensors for quantitative measurements of biomarkers along with a binary, digital readout approach. Farzadfard *et al.* demonstrated that genomic DNA could be used as a rewritable and flexible substrate to memorize analog information, such as the magnitude of an input signal, as a proportion of cells in a population. This platform could enable long-term cellular recorders for diagnostic applications¹²¹.

Even though signal processing *in vivo* was first implemented with the use of synthetic gene circuits, fast kinetic events in biology are increasingly supported by protein-based signal processing systems. Beyond nucleic acid as a substrate for information in such circuits, protein-based synthetic systems have the potential to enable modular and efficient computation through post-translational mechanisms²⁰¹. Information processing can be supported by protein-protein interaction such as binding combined with activation or inhibition of catalytic activity like phosphorylation or proteolysis. Recently, intein splicing has received attention to construct synthetic protein circuits, as they support their own catalysis and subsequent excision followed by intein tagged protein fusion and function recovery. Interestingly, this event can be activated by

small molecule ligands or protein scaffolds, and allows for spatial control, implementation of Boolean logic, or signal amplification via synthetic cascading^{477 478 479}. Protease degradation has also been described as a tool to engineer control signal processing in synthetic protein circuits⁴⁸⁰. For example, Prindle *et al.* used protein degradation as a tool for rapid and tunable post-translational spatial and temporal control on gene expression⁴⁸¹. MAPK networks have also been successfully rationally engineered for synthetic cascading to generate modular, insulated, ultrasensitive and tunable signaling⁴⁸². Other approaches have made use of chimeric regulatory proteins in synthetic signaling, exploiting for example two-component systems of bacteria, to achieve novel customized signaling⁴⁸³. Moreover, due to the fact that genetic circuits and proteins operate on different time scales, developing hybrid synthetic networks could prove valuable. For example, the output of protein-based information process could then be stored in recombinase-based memory register, or integrated via CRISPR-Cas9 or inteins splicing protein such as TALEs or ZFN^{484 485 486 487 488}.

Although cellular context can be assumed disruptive, it may also play supportive roles in the functioning of synthetic circuits and provide relative robustness, performance and maintenance that can be valuable and exploited in specific contexts. However, while the engineering of orthogonal biological parts and signal processing frameworks *in vivo* have proven valuable for synthetic biologists, potential discrepancies remain, such as high context and chassis dependency.

1.3.2 Synthetic biological circuits operating *ex vivo*

In cell-free systems, synthetic parts are exempt of adaptation and evolution and as a result can benefit from relatively more tunable and reproducible behavior. Efforts to reproduce the response capabilities and complexity of cellular circuits from the bottom-up approach have been reported with the assembly of synthetic biochemical reaction networks³⁹. These synthetic systems involving biocatalytic reactions can be utilized for biosensing, information processing and biocomputing. Extensive research has been conducted on *ex vivo* systems, greatly motivated by applications in biodiagnostic. Advances in biomolecular computing systems mimicking electronic substrates, has resulted in the development of novel synthetic biological signal processing framework. For example different biomolecular tools, including proteins/enzymes, and nucleic acids have been used to implement layered Boolean logic gates. While further scaling up the complexity of biochemical information processing systems had remained a challenge, recent results showed promises in that direction.

As we have precedently seen, nucleic acids are modular chemical building blocks with structural, mechanical and catalytic capabilities. Nucleic acid enzyme-based or enzyme-free computation systems, aptamers, ribozymes, circuits, origamis, and gels offer a wide repertoire for the design of biological signal sensors and processors⁴⁵². DNA has been extensively and successfully used *in vitro* to implement networked logic operations, with an important scaling up in number of logic gates. Nucleic acids are capable of both carrying information and performing computations on that information. Circuits relying on nucleic acids as a substrate have few possible interactions

and points of control making their quantitative design, simulation and description manageable. For example, Kim *et al.* showed how a synthetic nucleic acid circuit could be systematically designed to perform pulse generation, adaptation, and fold-change detection. This study demonstrated the programmability and ability of such circuits to obtain predictive dynamical systems in a cell-free environment for biosensing applications⁴⁸⁹. Chen *et al.* also reported a DNA-based architecture for implementing *in vitro* computational programs using the formalism of DNA reaction networks as a universal *programming language* to implement any function that can be mathematically expressed. In this study, the formalization allows complex signal processing of intrinsically analogue biological and chemical inputs, and not only Boolean logic⁴⁹⁰.

Proteins have also been used to make Boolean logic gates *in vitro*. During the last decade, numerous studies have pioneered the engineering of enzyme-based logic gates concatenated in information processing systems⁴⁹¹. Biochemical reaction networks can implement multi-signal Boolean logic or arithmetic operations such as addition or subtraction^{492 493 494}. Biomolecular circuits are also capable of implementing dynamic behaviors including pulsing, adaptation and fold-change detection⁴⁹⁵. Novel cell-free biosensing concepts have capitalized on the idea of integrating multiple molecular inputs processed biochemically before transducing their output on *smart-material* interfaces such as functionalized electrodes or metallic nanoparticles, to give a hybrid bio/electronic signal processing. For instance, signal-responsive electrodes for signal readout have been coupled with biochemical logic gates^{359 496 497}. Moreover, taking technology further, future approaches could tend toward the full integration of biochemical and electronic processing⁴⁹⁸.

1.4 Thesis statement: How can synthetic biology serve the engineering of next generation diagnostics?

The engineering and refinement of standardized genetic and biochemical parts have constituted the central dogma of synthetic biology, which yielded increasingly efficient tools to be assembled into well-characterized circuits and systems. As we have discussed, a major thrust of synthetic biology has been to develop cell-scale devices achieving useful information processing operations with predictability and accuracy. This framework has proven of tremendous interest for biomedical applications, and we propose that synthetic biology could serve as a methodology to interface human physiology through biosensors development.

Indeed, diagnostics yield a great deal of information, which clinicians have to analyze and evaluate comprehensively in a short time. New diagnostic possibilities permitted by synthetic biology could improve the ability to assess pathological states and monitor diseases and their prognosis. The diagnostic process falls into the definition of computing, and synthetic biology provides a modular substrate for sensing, computation and interfacing.

However, although computational versions of diagnosis within biological components have been proposed, to date no biological computing system embedding diagnostic algorithm following medical knowledge has been approved as a medical problem solving systems. While the development of synthetic biosensors has increased in recent years, most of the potentially clinically relevant bioanalytical platforms discussed before were implemented in *clean* environments, their operation and optimization in *real* biological samples, such as serum or urine has not been addressed. To date, developing systematic and universal methodologies for synthetic biosensor engineering has remained elusive. Moreover, although models have been constructed *de novo* via a bottom-up approach, none have managed to develop a methodology to coordinate precise and predictive behaviors at the system-level. Therefore, novel methods to engineer synthetic biosensors are necessary if we wish to bring improvements in capabilities towards clinically compliant diagnostics.

In this work I explore how synthetic biology can be used to engineer new expert diagnostic systems, through the combination of biosensing with programmable signal processing, while keeping in mind that their effective use relies on bioengineering solutions ensuring robust and reliable behavior. For this purpose, I investigate how cells and protocells can be engineered into intelligent and accurate diagnostics through major engineering criteria:

***Full programmability:** achieve user-defined biosensing & biocomputing operations according to specifications: molecular input detection and decision making algorithms, to output bioactuation, should be able to be straightforwardly encoded within biological substrate and repurposed for different agendas.

***Scalability:** the platform should allow for increasingly multiplexed sensing and complex signal processing, and should support high yield fabrication.

***Autonomy and portability:** operation should be possible in delocalized set-ups, various contexts, complex environments and modalities.

***Robustness:** behavior should remain stable and predictable, while responding conditionally with spatiotemporal precision to clinically relevant patterns of molecular cues according to predefined specifications.

The two opposite approaches (top-down versus bottom-up) explored in this work and described below outlines methodologies that have been employed to engineer two different platforms: cell-based and cell-like synthetic biosensors, through two technological frameworks relying on synthetic genetic and biochemical circuits respectively.

In a long term vision, this work proposes unprecedented and universal approaches to program multifunctional biosensors through predictive composition of *de novo* genetic and biochemical circuits. The tools developed in this thesis are intended to be useful to both fundamental research and biomedical sciences, are expected to bring novel opportunities to interrogate and interface biology and serve the creation of a technological foundation to engineer next-generation biosensors for medical diagnosis.

1.5 Approach, summary of contributions and overview of the dissertation

This thesis focuses on fundamental synthetic biology research with an important translational interest, for which I benefited from a double medical and scientific background. My PhD and medical residency was dedicated to the creation of basic knowledge, molecular tools, experimental and computational methods to engineer autonomous and programmable biosensors, with a translational perspective in mind and diabetes as a testbed pathology. This permitted me to produce innovative applications and research with impact, while achieving the first methodological and technological bridge between synthetic biology and medical sciences to develop next-generation diagnostics.

In **Chapter 1**, I proposed to summarize and put into clinical perspective the most important medical advances of synthetic biology. Emerging as a mature field increasingly transitioning to the clinics, we discuss how synthetic biology can apply engineering principles to design and build biological systems with clinical specifications. In this section, we intended to capture the translational impact and medical importance of synthetic biology for the development of next-generation diagnostics. I discussed how key synthetic biology concepts (standards, modularity, programmability, biosensing, biocomputing) can prove extremely useful for integration into diagnostics to provide with novel capabilities (multiplexed, dynamic, real-time, *in situ* and programmable monitoring of pathologies). I discussed how synthetic biology is preparing the future of medicine, supporting and speeding up the development of innovative diagnostics to bring direct improvement in medical procedures from the clinical lab to the patient, while addressing healthcare evolution and global health concerns. I concluded that synthetic biology is

the most likely technological field on the edge of realizing the promises of precision medicine, personalized solutions to diagnosis and therapy and answer a need arising from healthcare evolution.

Then, in a first experimental approach discussed in **Chapter 2**, I investigated how to engineer next-generation cell-based biosensors as *intelligent* diagnostics. I engineered bacterial biosensors as programmable and autonomous diagnostic platforms that can be easily programmed for different agendas. I conducted genetic engineering *in vitro* to design synthetic gene circuits to perform specific tasks when embedded in bacteria. For this purpose, I exploited the modularity of genetically encoded digital amplifying genetic switches and Boolean Integrase Logic (BIL) gates, which rely on phage integrase enzymes (BxB1 and TP901) acting conditionally on bacterial DNA.

I first mined microbial organisms to identify bacterial sensor modules for human disease biomarkers, and developed an innovative directed evolution strategy to systematically tune the expression of synthetic integrase cassettes to match physiological biomarker concentration thresholds. This method enabled the fast prototyping and straightforward reprogramming of cell-based biosensors with improved capabilities to detect various clinically relevant molecular signals. The synthetic gene circuits I designed could multiplex the detection of biomarkers and perform signal digitization at selected thresholds, signal amplification, months stable data storage in DNA registers, human readable colorimetric reporting, and carry out computational processes to integrate medical decision algorithms.

In addition, I developed an analytical framework with which to quantify cell-based biosensor robustness of operation in human clinical samples, and identified synthetic genetic parts and standards to achieve signal normalization and improve biosensing robustness. I investigated the use of different bacterial chassis, and reported that engineered *E. coli* and *B. subtilis* were capable of supporting reliable operation in complex media with high signal-to-noise ratio. Then, I developed orthogonal polymer chemistries to integrate these bacterial biosensors within stable, optically clear, easy to handle format capable of classifying pathological states detected in human clinical samples according to user defined medical rules.

I conducted translational studies to assess how these novel diagnostic devices could perform in a clinical set up, and gathered diabetic patient cohorts and samples at the University Hospital of Montpellier. I showed that bacterial biosensors could operate robustly in human urine and serum, and detect pathological biomarkers with medical accuracy (i.e. onset of diabetes through detection of glycosuria and nitric oxide), while integrating Boolean decision algorithm to aid medical decision. I also investigated how bacterial quorum sensing systems could be repurposed and integrated into a bacterial biosensing platform to achieve nanomolar detection of sepsis biomarkers in pathological samples, while being capable of discriminating infectious aetiologies using genetically encoded algorithms. This new generation of bacterial biosensors constituted the first successful proof of concept for the diagnosis of different pathologies in real world clinical settings. These tools, biological standardized parts and methodologies developed are openly distributed amongst the community to enable the future modular engineering of new types of biosensors.

In a second approach described in **Chapter 3**, I developed systematic methods and concepts to engineer protocells from the bottom-up, capable of performing biosensing and biocomputing operations. I successfully engineered programmable microscale protocells capable of digitization of space and molecular signals through membrane compartmentalization and encapsulation of synthetic biochemical circuits respectively, in order to carry out specified tasks *in situ*. I proposed that protocells constituted by a phospholipid membrane encapsulating synthetic biochemical circuits (i.e. *de novo* networks of enzymes and metabolites) can perform biosensing of disease associated biomarkers and biocomputing operations, generate various types of outputs (e.g. fluorescence, colorimetric, enzymatic activity) and can be systematically generated with a robust framework to provide analytical solutions to specific clinical questions. I demonstrated that diagnostic processes formalized as Boolean functions (i.e. diagnosing specific pathologies through the biodetection of patterns of biomarkers in urine or blood) could be implemented *in vitro* with synthetic biochemical circuits.

First, in collaboration with computer scientists and mathematicians we developed and refined a computer assisted framework for systematic design of synthetic biosystems and biochemical circuits. We developed the first CAD softwares for bottom-up synthetic biology and evolved models to take into account biochemical and physicochemical parameters relevant to bottom-up design. These tools support user friendly design, biological parts mining, hybrid simulation, model checking, robustness and sensitivity analysis, and exploration of design space.

Secondly, developing protocells as robust biosensors required full control on physicochemical properties. Thus, methods for the encapsulation of synthetic circuits in protocells ensuring fine tuning on parameters was required, as well as well-defined size, lipid composition, enzyme content, catalytic performance and stability of encapsulated circuits. I thus set up a microfluidic platform, and then developed a microfluidic methodology that would simultaneously ensure (i) control on membrane biophysical properties, (ii) encapsulation efficiency and stoichiometry, (iii) monodisperse size, (iv) stability and (v) Integration of protein pores. I set up a microfluidic platform and designed PDMS chip architectures that permitted high-throughput fabrication with fine control on biophysical parameters.

I discuss how this strategy allows straightforward and accurate programming of protocells according to logic specifications, by designing synthetic biochemical circuits *in silico* and implementing them *in vitro*. This framework first relied on the isolation and identification of suitable biochemical parts such as enzymes and metabolites. This is achievable using the *in silico* tools we developed, that enabled us to mine for robust parts and circuits performing specific Boolean operations. I carried out simulations and model prediction in order to identify kinetically and functionally suitable enzymes, metabolic parts and circuits that would then be experimentally implemented *in vitro*. Then, I performed temporal logic model checking, robustness and sensitivity analysis in order to optimize experimental parameters and predict modes of failure. Using this framework, I successfully implemented protocells that could perform arbitrary biosensing tasks and recapitulate specified logic operations.

I demonstrated that the framework was successful at implementing synthetic biochemical reaction circuits inside protocells for biosensing and integration of decision algorithms. These diagnostic devices could perform multiplexed biodetection and were capable of classifying pathological states *in situ* in clinical samples according to specific patterns of biomarkers. Specifically, I engineered protocells integrating a full diagnostic algorithm that discriminates between all acute metabolic complications of diabetes in blood and urine, and achieve differential diagnosis between diabetic ketoacidosis, hyperosmolar hyperglycemic nonketotic syndrome, hypoglycemia and lactic acidosis, which are associated with high medical and socio-economic burden and an important mortality and morbidity. I further conducted a translational study on patient cohorts to provide with experimental evidence demonstrating the technological advantages, medical validity, and efficiency in clinical samples of this novel diagnostic approach for the diagnosis of human diseases.

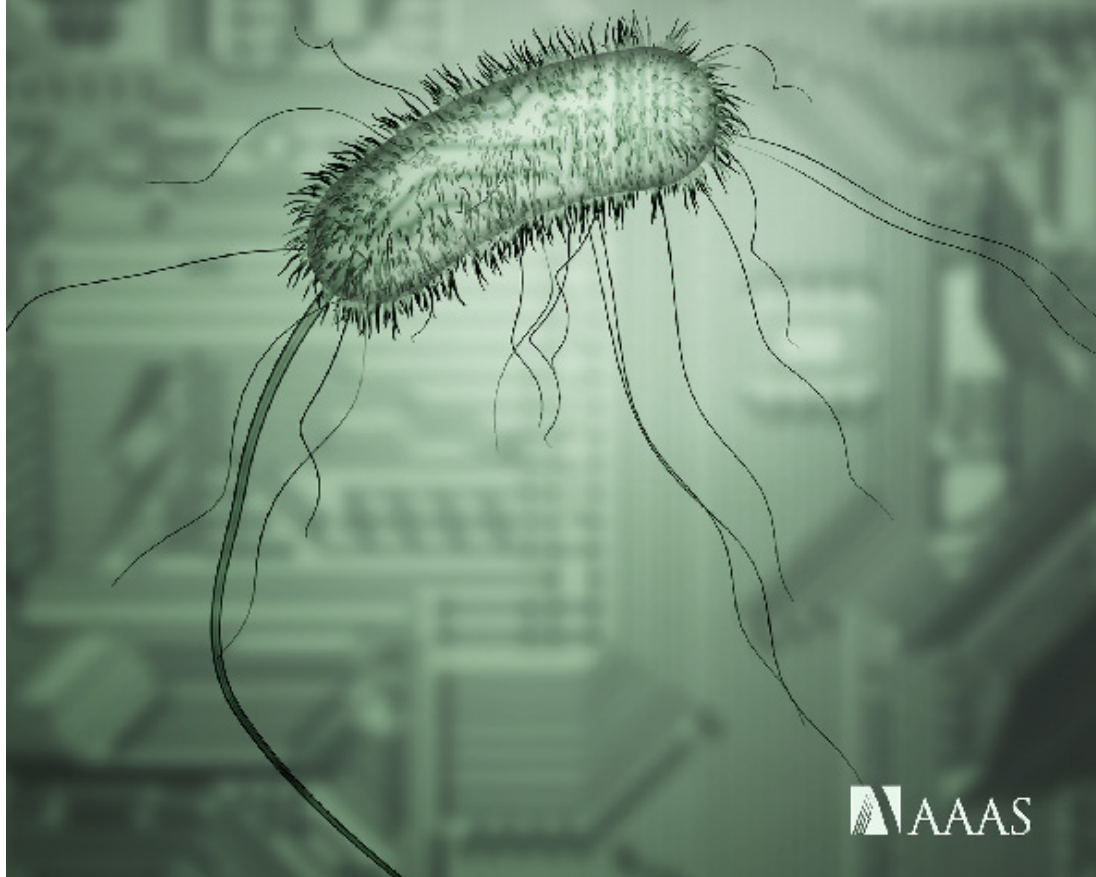
Rationally designed protocells appeared as highly promising tools for performing multiplexed *in vitro* diagnostic integrating medical algorithmic processes. This work demonstrated that this technology could be successfully applied to solve real clinical problems. These diagnostic devices are under patenting process, with a manuscript waiting for publication. The biotechnological tools, softwares and framework developed are under investigation for various applications, and are freely distributed.

Last, in **Chapter 4**, I show how the capabilities of these synthetic devices can be theoretically extended to solve various problems. We investigated the use of synthetic protocells as universal computing devices. We developed a theoretical framework for the use of populations of protocells as new kinds of computers, and explored the modalities of operation and advantages they could offer. I show how these assemblies of standardized protocells can constitute ultra-low power, massively parallel computers at very high densities ($\sim 10^{12}$ machines/ml), and address complex problems while being biologically interfaceable. We provided mathematical and biological evidence to support the engineering of protocellular computers and analyzed specific examples of implementation to solve NP-complete problems in fast and non-heuristic ways.

Science

27 May 2015

Translational Medicine



Bug-Based Biosensors. *Shown is a bacterium as imagined by a synthetic biologist—as a living tool with insides akin to a printed circuit board, which can be rewired to perform a variety of desired functions. With their natural ability to detect biomolecules, process signals, and respond, bacteria are clear candidates for biosensing devices...*

A little help from our (little) friends

It's only logical: Translation of diagnostics to home health care or remote setting requires simple methods for measuring markers in complex clinical samples. And living cells—with their ability to detect biomolecules, process the signal, and respond—are logical choices as biosensing devices. The recent buzz on human microbiota has expanded our view of bacteria beyond infectious enemies to metabolic buddies. Now, Courbet et al. refine that view further by engineering bacteria to serve as whole-cell diagnostic biosensors in human biological samples. Although whole-cell biosensors have been shown to serve as analytical tools, their quirky operation and low signal to noise ratio in complex clinical samples have limited their use as diagnostic devices in the clinic. The authors engineered bacterial biosensors capable of signal digitization and amplification, multiplexed signal processing (with the use of Boolean logic gates), and months-long data storage. As a proof of concept, the “bactosensors” detected pathological levels of glucose in urine from diabetic patients, providing a framework for the design of sensor modules that detect diverse biomarkers for diagnostics.

Editor's Summary⁹⁶

The DNA size probably forbids us forever to modify the genome.

Jacques Monod, *Le Hasard et la Nécessité*, 1970

Chapter 2

Engineering bacteria as a programmable and autonomous biosensing platform through integrase based synthetic gene circuits

Portions of the text below were adapted from *Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates* by Courbet *et al.*, *Science Trans. Med.* (2015)⁹⁶

Abstract

Cell-based biosensors have several advantages for the detection of biological substances and have proven to be useful analytical tools. However, several hurdles have limited cell-based biosensor application in the clinic, primarily their unreliable operation in complex media and low signal-to-noise ratio. Here I report that bacterial biosensors with genetically encoded digital amplifying genetic switches can detect clinically relevant biomarkers in human urine and serum. These biosensors perform signal digitization and amplification, multiplexed signal processing with the use of Boolean logic gates, and data storage. In addition, we provide a framework with which to quantify cell-based biosensor robustness in clinical samples together with a method for easily reprogramming the sensor module for distinct medical detection agendas. We demonstrate that biosensors can be used to detect pathological glycosuria in urine from diabetic patients. Last, we discuss potential application and preliminary results for sepsis differential diagnosis. These next-generation cell-based biosensors with improved computing and amplification capacity could meet clinical requirements and should enable new approaches for medical diagnosis.

2.1 Introduction

2.1.1 Bacterial biosensing

As we have precendtly discussed, *in vitro* diagnostics are growing in importance in the global health arena because of their non invasive nature and resulting ease of use and scale^{86 499}. However, conventional detection methods are often expensive and complex, and thus difficult to implement in resource limited settings⁹⁰. In response to these challenges, bioengineers have developed attractive methodologies that rely for instance on synthetic nanoprob^{91 500 501} or microfluidics^{502 503}. Yet, there remains a need for easy-to-use, portable biosensor devices that can be used by nonspecialists to make clinical measurements at home or in remote locations^{91 504 505}. Among biosensing devices, cell-based biosensors mainly based on bacteria have proven to be applicable for the detection and quantification of a wide range of analytes^{117 506}. Living cells have many attractive properties when it comes to diagnostics development. Cells detect biomolecules with high sensitivity and specificity and are capable of integrated and complex signal processing. Cells also provide a self-manufacturing platform via autonomous replication¹¹⁰, and the production of laboratory prototypes can be scaled using existing industrial frameworks⁵⁰⁷. Spores from cell-based biosensors can remain functional for extended periods of time, increasing the shelf life of a diagnostic product in harsh storage conditions⁵⁰⁸. Last, cell-based biosensors are highly versatile and can be used as stand-alone devices or interfaced with other technologies such as electronics, microfluidics, or micropatterning^{134 137 509}. All of these advantages have prompted the development of cell-based biosensors that measure a variety of clinical parameters^{139 141 142 162 510 511}.

However, cell-based biosensing systems have not yet been applied for the monitoring of medically relevant parameters in a clinical context. Many challenges have limited their translation to the clinic: *(i)* unreliable operation and low signal-to-noise ratio in complex and heterogeneous clinical samples; *(ii)* the inability to engineer ligand-tailored sensors; *(iii)* limited signal-processing capability, which precludes the integration of several biomarker signals for accurate diagnosis; *(iv)* lack of consistent frameworks for the assessment of robustness in challenging clinical conditions; *(v)* response times that are not compatible for diagnosis that require fast delivery of results; and *(vi)* compliance to clinical formats. Therefore, here we aim at employing the emerging field of synthetic-biology to streamline the rational engineering of biological biosensing systems, to bring new hopes for compelling translational medicine applications.

2.1.2 Synthetic biology & cell-based biocomputing: integrating algorithms for decision making

The reader is here kindly referred to **Section 1.3** of the introduction, where the subject is covered in detail.

As we have seen, researchers have developed proof-of-concepts where they embedded medical algorithms within living cells for diagnosis, disease classification, and treatment^{102 189 194 512 513}. However, the use of synthetic biology tools and concepts to improve *in vitro* diagnostics technologies has remained limited. As we have previously presented, synthetic biology focuses on parts and systems standardization, the engineering of modular components, and systematic strategies for the engineering of biological systems and new biological functions with reliable and predictable behaviors. Molecular modules such as sensors, reporters, or switches could ultimately be assembled at a systems level to perform specific tasks. Genetic devices that support *in vivo* computation were developed recently and enable living cells to perform sophisticated signal-processing operations such as Boolean logic, edge detection, or cellular profiling^{24 39}. Therefore, synthetic biology could presumably support the design of cell-based biosensors that meet medical specifications and help to translate cell-based biosensors to clinical applications. However, despite interesting advances, synthetic gene circuit for cell-based biocomputing remained difficult to scale in order to reach useful biosensing application, mainly due to low composability in individual cells.

2.1.3 Integrase based digital amplifying gene switches and Boolean logic gates for medical biosensing

Recently, synthetic genetic devices relying on the use of phage integrases have been introduced. Integrases are viral enzymes catalyzing strand exchange (i.e. recombination) between specific DNA sequences. These mechanisms are naturally exploited by bacteriophages when integrating their genomes within their host. Recombinatory mechanisms have been presented as unique avenues to perform biocomputing and signal processing within biological systems⁵¹⁴. In fact, integrases allow distinct binary states to be encoded into a DNA sequence, thereby bypassing limitations arising from layering of biological logic gate and memory systems. More specifically, among a vast family, serine recombinases have the advantage to operate autonomously in complete orthogonality across all kingdoms of life. It allows for precise manipulation of DNA *in vitro* and *in vivo* depending on the relative location or orientation of short recombination sites (~25 bp) it provokes either integration, excision or inversion⁵¹⁵ (**Figure 2.1A and 2.1B**).

Exploiting these mechanisms, Bonnet *et al.* recently managed to build a device architecture in which a minimal number of biomolecular entities could be reused to implement and concatenate all existant logic gates within a single logic register⁴⁵⁸. Interestingly, decoupling input and output allows for straightforward engineering of Boolean logic biological devices, analogous to transistor architecture with standardized switching and amplifying behavior. Importantly, the biological substrate supporting the device is encoded in DNA, a common signal carrier ensuring biological

universality of operation and programmability. Similarly to transistors, these three terminal devices or *transcriptor*, are constituted by independent analog input control signals that govern *current*, which in this case is analogous to transcriptional flow (RNAPol along DNA, **Figure 2.1C**). The *Gate* is constituted by asymmetric logic elements written in DNA, such as transcription terminators. These gate elements are flanked by integrase recognition sites, which they can unidirectionally invert or delete to allow transcriptional current to proceed.

Composing with specific transcriptors enables to write complex yet portable sequences of operation (i.e. multiple integrase logic gates) to be performed by a genetic circuit, which do not require layering. Logic registers can be built via straightforward and rational DNA encoding of genetic elements (recombinase targets, terminators and other modules), enabling simple user-defined programming of logic and memory⁴⁵⁷. Along with logic, transcriptors integrate memory capabilities that prove precious to perform complex and persistent computing such as sequential logic. In addition, DNA written memory is a powerful method for long-term storage, since it is propagated through generation, and persists even after cell death. Although this technology was first developed using robust integrases from bacteriophages TP901-1 and Bxb1⁵¹⁶, an increasing number of orthogonal integrases are being mined from organisms and standardized. This strategies will allow for increasing biocomputing and biosensing capabilities, with now 1.375 bytes of information to be encoded in this way (i.e. 2^{11} distinct states)²¹.

These integrase logic gates enable truly digital and discrete operations and thus system output responses, compared to previous systems that produces intermediate expression levels^{102 197 448}. Integrase control signals can be wired to a biosensing module, thereby providing digitization, amplification and multiplexing of biological signals. This in turn improves robustness, sensitivity and mediates sharp responses. These properties make them highly relevant for medical applications and particularly for diagnostics as it is often threshold based. The built in memory capacity also enables the recording of weak or transient signals while giving a constant amplified output. Compared to transcriptional switches⁴¹, it has no metabolic cost and is stably written in either chromosomic or plasmidic DNA, and can be addressed after extended periods of time and lysis of the bactosensor in clinical samples. These signal processing devices have the advantage of high composability for the programming of various medical algorithms. We suspect that these characteristics will be important to enable robust biosensing and computation in the context of intracellular and environmental fluctuations.

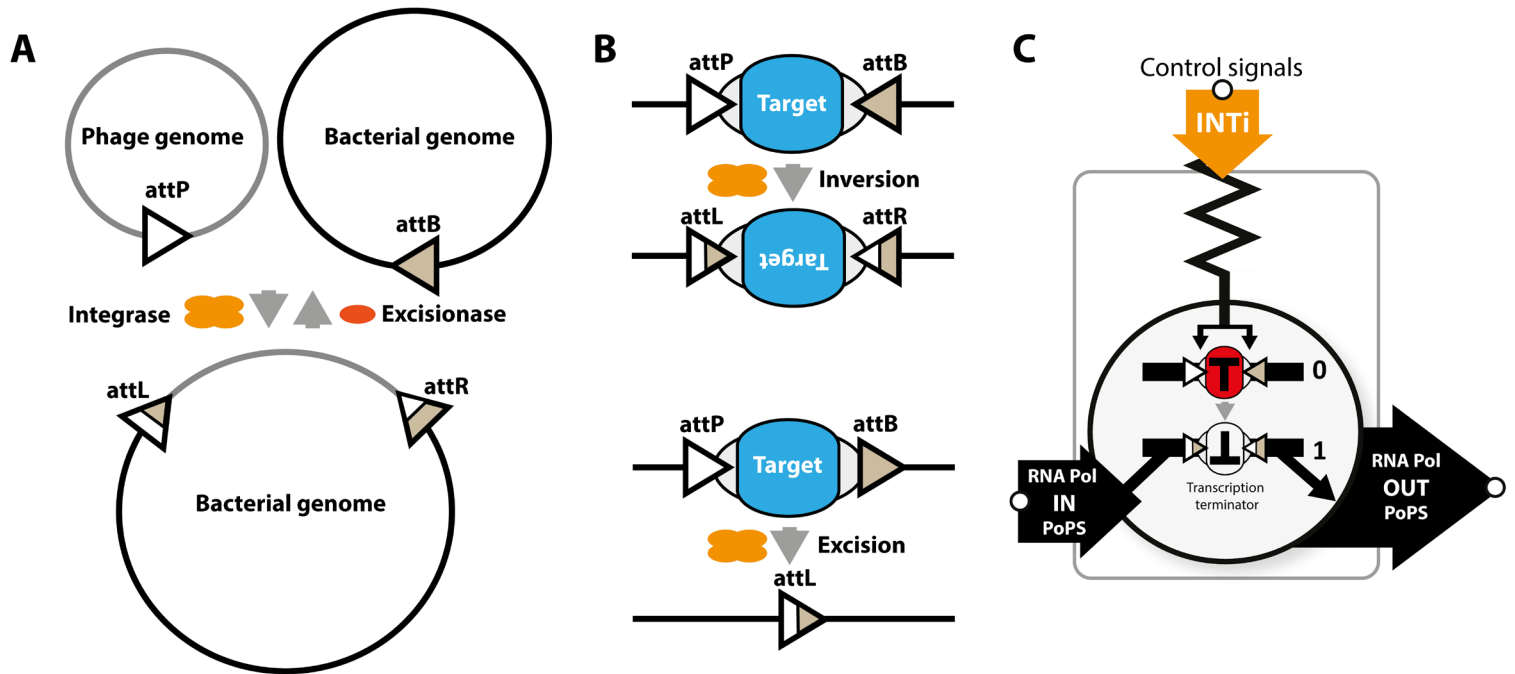


Figure 2.1: Phage integrases reactions are exploited to build Transcriptor devices (A) Natural mechanisms of integration/excision of Phage DNA into bacterial host. In the presence of integrase enzyme, recombination proceeds and the the viral DNA is integrated at the bacterial attachment site *attB* (i.e. lysogeny). The protein excisionase can reverse the reaction and produce excision of viral DNA (i.e. lytic induction). (B) Engineered version of the natural system, were the integrase is used to either flip a target DNA sequence flanked by *att* sites, or excise it, thereby controlling specific genetic registers. (C) Three terminal transcriptor device uses integrases triggered by control signals to control RNA polymerase flow between a separate input and output.

Here, we investigate the use of these recently developed digital amplifying genetic switches and logic gates⁴⁵⁸ to bring the performance of cell-based biosensor closer to clinical requirements. These genetic devices enabled bacteria to perform, in human clinical samples, reliable detection of clinically relevant biomarkers, multiplexing logic, and data storage. We also provide a framework for quantifying cell-based biosensor robustness in clinical samples together with a method for straightforward reprogramming of the sensor module for distinct medical detection agendas. Hence, our platform architecture is highly modular and could be repurposed for various applications. We anticipate that such engineered bacterial biosensors, we named *Bactosensors*, that are capable of *in vivo* computation could be tailored according to medical knowledge and used as expert biosensing devices for medical diagnosis (Figure 2.2).

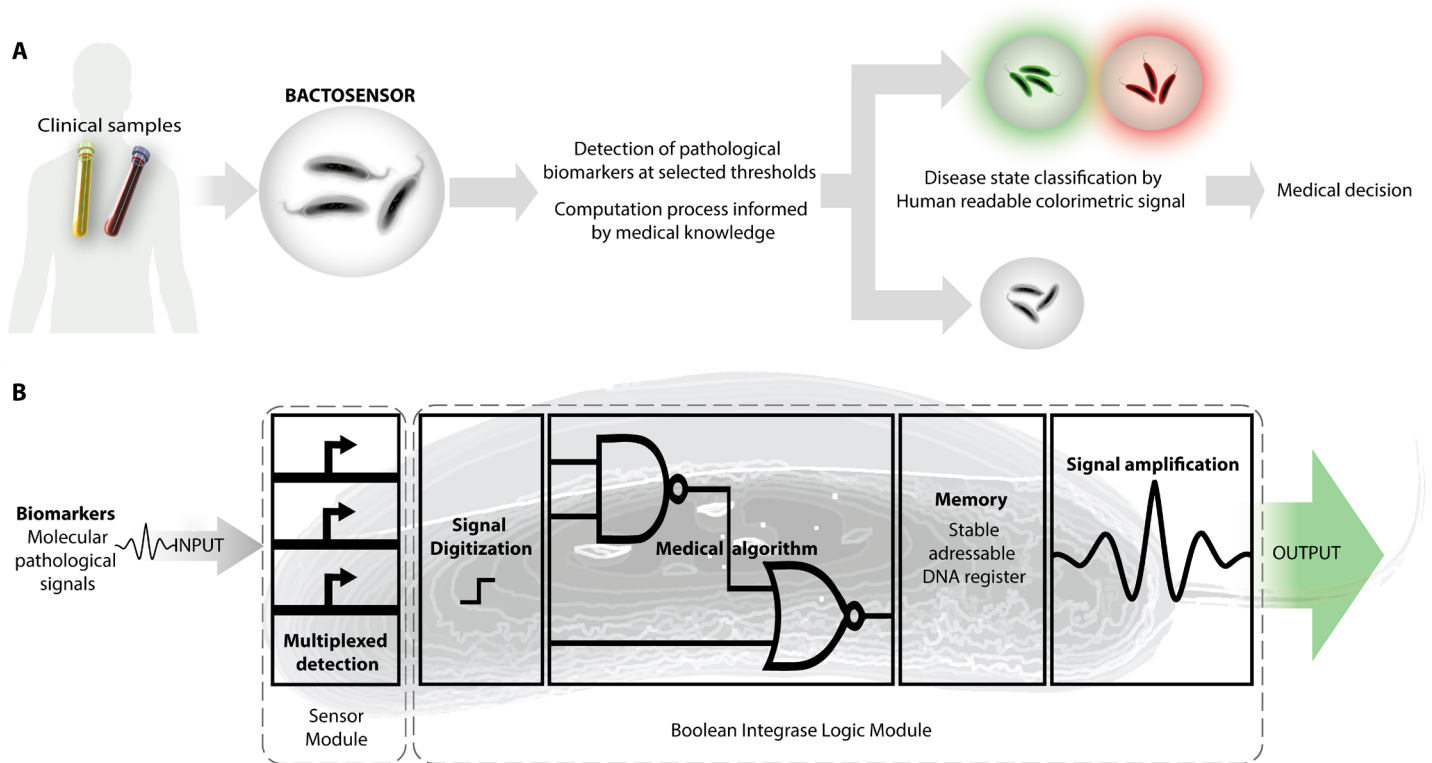


Figure 2.2: Operational principle and architecture of batosensors to perform medical diagnosis. (A) Medical diagnosis is a computational process that can be formalized using Boolean logic in vivo and embedded into a batosensor. The batosensor is capable of detecting a pattern of specific biomarkers in human clinical samples at selected thresholds and integrates these signals using an in vivo computational process. If a pathological pattern of biomarker is detected, the batosensor generates a colorimetric output. (B) Schematic architecture of a batosensor. A sensor module enables multiplexed detection of pathological biomarkers. These control signals drive a Boolean integrase logic gate module, which is the biological support for a userdefined digital medical algorithm. Boolean integrase logic gates also enable signal digitization and amplification along with storage of the diagnosis test's outcome in a stable DNA register that can be interrogated a posteriori.

2.2 Operational principles, design and architecture of bactorsensors

2.2.1 Behavior and robustness of bacterial chassis in human clinical samples

Our first goal was to determine the operational characteristics of bacterial chassis of interest in terms of growth, viability, and functionality of synthetic gene circuit in human body fluids of clinical relevance: human urine and blood serum. We chose to evaluate the robustness of Gram-negative and Gram-positive bacterial models (i.e. *Escherichia coli* and *Bacillus subtilis*, respectively) that have been used in previous cell-based biosensors designs. To this end, we collected urine and serum from healthy volunteers, pooled the samples to average molecule concentrations to account for possible variations among individuals, and prepared dilutions with a defined culture medium (see **Materials and Methods**). We then inoculated various clinical sample dilutions with cells from stationary cultures of *E. coli* DH5 α Z1 or *B. subtilis* 168, grew these cultures for 18 hours at 25°, 30°, or 37°C, and measured their optical densities. For both cell types, we observed cell growth across the complete range of sample dilutions and at all three temperatures (**Figure 2.3**). However, growth was strongly inhibited at 100% urine or serum concentrations, probably because of the lack of nutrient provided by the diluted culture medium. Growth of both cell types decreased with increasing urine or serum concentration, but cell death was insignificant (<2% for all samples; **Figure 2.4**). These results demonstrate that both Gram-positive and Gram-negative bacteria can survive and proliferate in human clinical samples for several hours. Because of the larger number of tools available for the reliable control of gene expression^{464 465}, we chose *E. coli* for further engineering of a prototype bactorsensor.

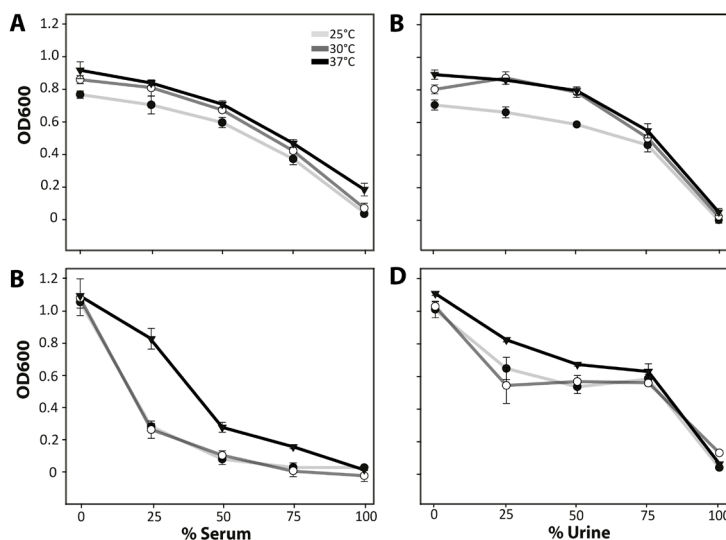
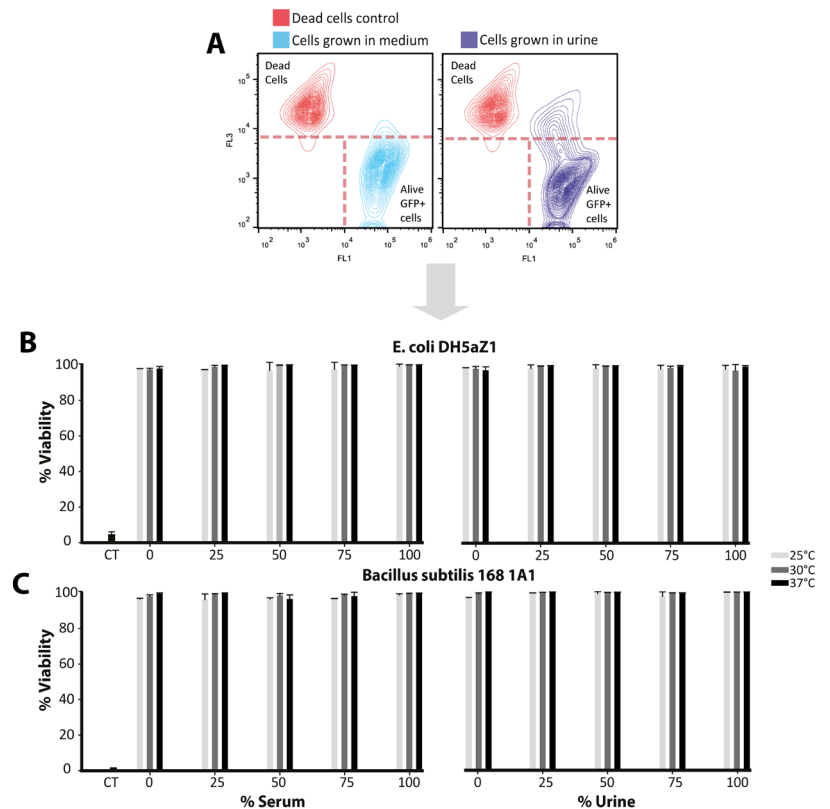


Figure 2.3: Bacterial chassis growth characteristic in human urine and serum. Optical densities measurements of bacterial cultures after growth in various concentrations of human urine and serum and at different temperatures. *E. coli* (A, B) and *B. subtilis* (C, D) cultures were grown at different dilution (0, 25, 50, 75, 100%) of human serum (A, C) or urine (B, D), and at different temperature of incubation (25, 30, 37°C). Bacterial colonies were grown in LB for 6 hours at 37°C, diluted 1:100 in various dilutions of serum or urine and grown at different temperatures for 18 Hours. OD600 values of the culture medium or with specific dilution in clinical sample were subtracted from each measurement.

Figure 2.4: Single cell measurements of bacterial chassis viability in urine and serum.

(A) After measuring optical density as shown in figure S3, cells were stained with propidium iodide and analyzed by flow cytometry recording propidium iodide and constitutive GFP fluorescence. This enables to discriminate living cells (propidium iodide negative, GFP positive) from dead or compromised cells (propidium iodide positive, GFP positive) as depicted. The percentage of viable cells can then be expressed as the ratio of the number of cells under a *dead* threshold on the total of cells. See⁵¹⁰ for more information. *E. coli* (B) transformed with a reference promoter carrying plasmid (BBa_J23101) controlling expression of sfGFP and *B. subtilis* (C) cultures were grown at different dilution (0, 25, 50, 75, 100%) of human serum or urine, and at different temperature of incubation (25, 30, 37°C). Bacterial colonies were grown in LB for 6 hours at 37°C, diluted 1:100 in various dilutions of serum or urine and grown at different temperatures for 18 Hours.



We next assessed the capacity of *E. coli* cells growing in clinical samples to respond in a reliable way to exogenously added molecular signals using the model transcriptional promoters pTET (i.e. responding to anhydrotetracycline (aTc) induction) and pBAD (i.e. responding to arabinose (ara) induction), both driving expression of a reporter gene that encodes the green fluorescent protein (GFP). Both promoters were functional at all concentrations of clinical samples (Figure 2.7A and 2.7B), but cells that were induced in 100% serum failed to produce GFP, which indicates that serum has an inhibitory effect on bacterial gene expression. Increasing sample concentrations produced variations in autofluorescence, which could be corrected using a reference promoter⁵¹⁷. Such reference promoters could be used as internal standards to increase measurement reliability in a clinical setting (Figure 2.6). These results demonstrate that synthetic gene circuits can remain functional in clinical samples.

2.2.2 Design consideration of cell-based biosensors & rational improvement in robustness for the clinic

To improve the performance of synthetic circuits for the clinics, we incorporated in our design recently developed standardized regulatory genetic elements. Lack of robustness (e.g. matrix effects and nonspecific environmental interferences and variable parts performance across changing complex media⁵¹⁸) has limited the utility of engineered gene circuits for further medical applications and hindered translational advances in synthetic biology. We reasoned that using a minimal circuit design (for example, minimal possible number of parts), along with a genetic design relying on standardized genetic elements, could limit failure modes, increase safety and likelihood of approval of biosensors in a medical setting⁴⁷¹. In this regard, the single-layer architecture of Boolean Integrase Logic gate offers a significant advantage⁴⁵⁸. We also suspected that gene circuits' context sensitivity could be reduced by incorporating synthetic genetic part recently developed to decrease context dependency and improve predictability and reliability of gene expression⁴⁶³. Therefore, we incorporated in our design the following standardized genetic parts:

(1) Bicistronic Expression Operating Unit (EOU)

Engineering experience demonstrated that even simple genetic elements such as prokaryotic Ribosome Binding Sites (RBS) that initiate translation of a gene are known to behave differently in different genetic settings⁵¹⁹. In particular, this could be explained by the appearance of differential mRNA secondary structures at the 5'UTR region, which impacts the binding and translational activity rates of the ribosome⁵²⁰, and thus introduce unpredictability.

Therefore, we integrated a genetic device in the transcriptor design, which was previously developed by the BioFAB group to overcome this source of translational variability, the Bicistronic Design (BCD2, BioFAB #apFAB682). This genetic device has been shown to provide more predictability and standardization in engineering gene expression. The BCD relies on two sequential Shine-Dalgarno motifs, or two *cistrons*, where the RBS for the gene of interest is entirely embedded in the coding sequence of an upstream short synthetic peptide, while translation of the second cistron of interest is thus be coupled to translation of the first short cistron. This design exploits the helicase activity of ribosomes to extinguish secondary RNA structures that hamper translational initiation. Mutalik *et al.* demonstrated how it led to reliable and simply modeled gene expression⁵²¹. Therefore, we expected this device to improve reliability by ensuring a higher "ON" state of the transcriptor and thus a higher colorimetric signal compatible with human reading, while reducing context sensitivity. We experimentally validated this operational improvement (**Figure 2.5**). While moderately increasing the basal output from the transcriptor (~2 times higher), and leaving unchanged the fold change of the transition, the bicistronic design enables to achieve a more important discrimination between "OFF" and "ON" states by an increase in swing (e.g. the absolute change in fluorescence intensity between non-induced and induced states). The BCD is also capable of decreasing sensitivity due to the matrix effects by lowering unspecific activation of the transcriptor when used in clinical samples.

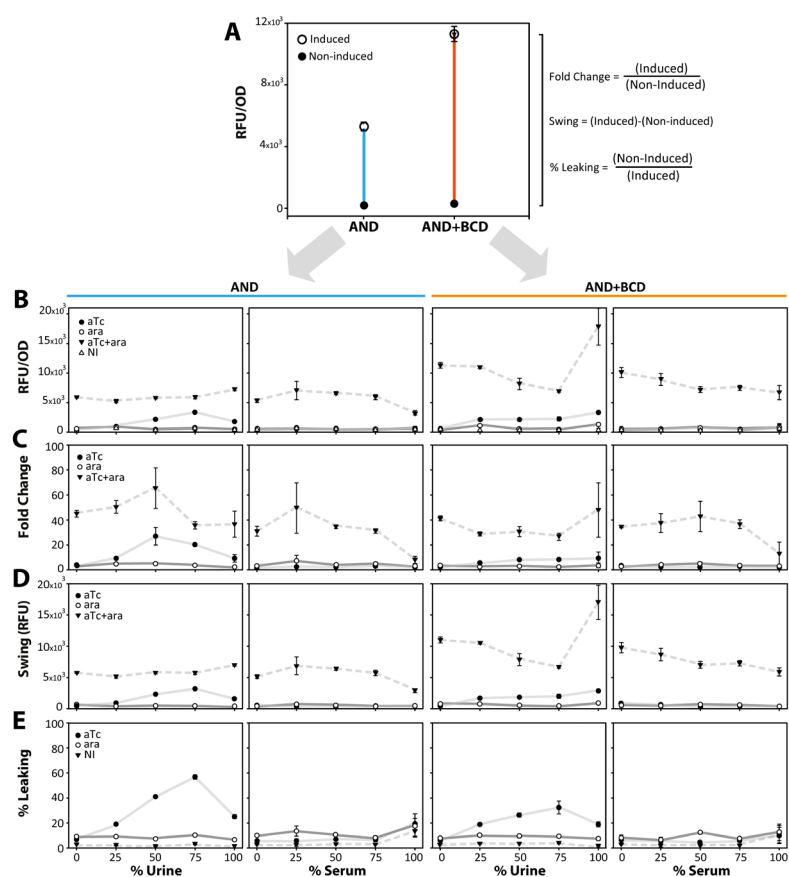


Figure 2.5: Comparison of the operational characteristic of AND logic gates with and without the Biconditional Device (BCD) for their use in clinical samples. (A) Comparison of response before and after induction between AND gate without (left) and with (right) BCD in culture media, at 25°C **(B)** Comparison of fluorescence intensity relative to optical density across various dilution of urine or serum **(C)** Comparison of Fold Change between different dilution of urine or serum. Fold change appears slightly reduced with BCD **(D)** Comparison of Swing (defined as the difference in fluorescence intensity between induced (anhydrotetracycline and arabinose) and non-induced state). Swing appears higher with the BCD. **(E)** Comparison of percentage of logic leaking (% Leaking, defined as the ratio of the response without or with one inducer to the response after induction with both inducers). Leaking increases with increasing concentration of urine, in the presence of only anhydrotetracycline, which could be explained by the non-specific induction by an arabinose analog present in urines. This non-specific effect tends to be lowered by the use of the BCD.

(ii) A ribozyme insulator part, RiboJ

Likewise, the nature of the region between the 5'UTR and the Shine-Delagarno in prokaryotes is known to generate context dependant fluctuation of the translational rates. We therefore integrated a self-cleaving ribozyme in the 5'UTRs of integrase controllers and transcriptors in order to get rid of this source of unreliability. We used an autocatalytic ribozyme device previously developed, called RiboJ⁴⁶⁵. Although we didn't validate the utility of RiboJ in our system, we believe that this genetic element, which was previously shown to efficiently buffer context effect of synthetic gene circuits⁴⁶⁵, will be useful to connect logic gates to various transcriptional sources.

(iii) A reference promoter used as an internal standard, pREF

Moreover, as we observed variations in the output fluorescence intensity measured depending on clinical samples concentration and type (**Figure 2.6**), we reasoned that the experimental design had to be improved to satisfy clinical requirements. We thus normalized the measured fluorescent output signal to an internal, *in vivo* reference signal (reference standard promoter, BBa_J23101) accounting for the analytical performance of the bactosensor across changing contexts. We thus use Relative Promoter Units (RPU) as an *in vivo* clinical reference standard for bactosensor operation and signal⁵¹⁷. Such internal standardization procedure based on reference objects has proven useful in several field of engineering and biomedical analysis and we believe will facilitate standardization of clinical assays using whole-cell biosensors.

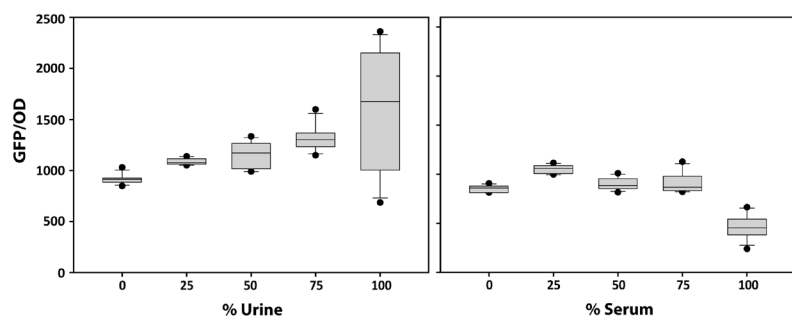


Figure 2.6: Influence of clinical media (urine and serum) on GFP fluorescence output generation and measurement. *E. coli* transformed with a reference promoter carrying plasmid (BBa_J23101) controlling expression of sfGFP were grown overnight in LB, and back diluted in fresh Azure medium with urine or serum as previously described. The cells were then cultivated at 25°C for 18 hours, and GFP and OD were measured, and their values corrected from the blank.

2.2.3 Multiplexing logic and memory in human clinical samples

As we have previously discussed, multiplexed biomarker assays are known to improve the performance and robustness of diagnostic tests⁹⁰. Signal processing allows an assay to integrate the detection of multiple inputs and to perform complex analytical tasks such as diagnostics algorithms informed by medical knowledge. Performing such integrated multiplexed detection and analysis within living cells thus requires some form of engineered biomolecular computation. We thus exploited the recently designed integrase based transcriptors, to control the flow of RNA polymerase along DNA via unidirectional inversion of an asymmetric transcriptional terminator^{458 459}. Transcriptors are digital amplifying switches that operate as analog-to-digital converters, are capable of signal amplification, can perform data storage and record transient signals, and can be composed to produce a variety of Boolean integrase logic gates (**Figure 2.7C**).

We first wanted to assess whether Boolean integrase logic gates could enable cell-based biosensors, operating in clinical samples, to execute complex signal-processing algorithms. We first evaluated the functionality and robustness of an AND gate that responded to the molecular signals ara and aTc in clinical samples and found that the logic gate operated reliably at room temperature in 100% urine and serum (**Figure 2.7D**, AND gate architecture depicted in **Figure 2.7C** and **2.9B**). We obtained similar results using NAND and NOR gates (**Figure S2.1**). Moreover, after gate switching, cells stored at 4°C could be regrown and the fluorescent output measured after up to 3 weeks of storage time (**Figure 2.7E**). Moreover, the signal stored within the DNA register could be recovered from bacterial cells that had been dead for 8 months using polymerase chain reaction (PCR) or Sanger DNA sequencing (**Figure S2.2**). Together, these results showed that living cells with embedded Boolean integrase logic gates can perform programmable, multiplexed signal processing in clinical samples. The ability to perform stable data storage over extended periods of times provides new opportunities for delayed readout in clinical environments.

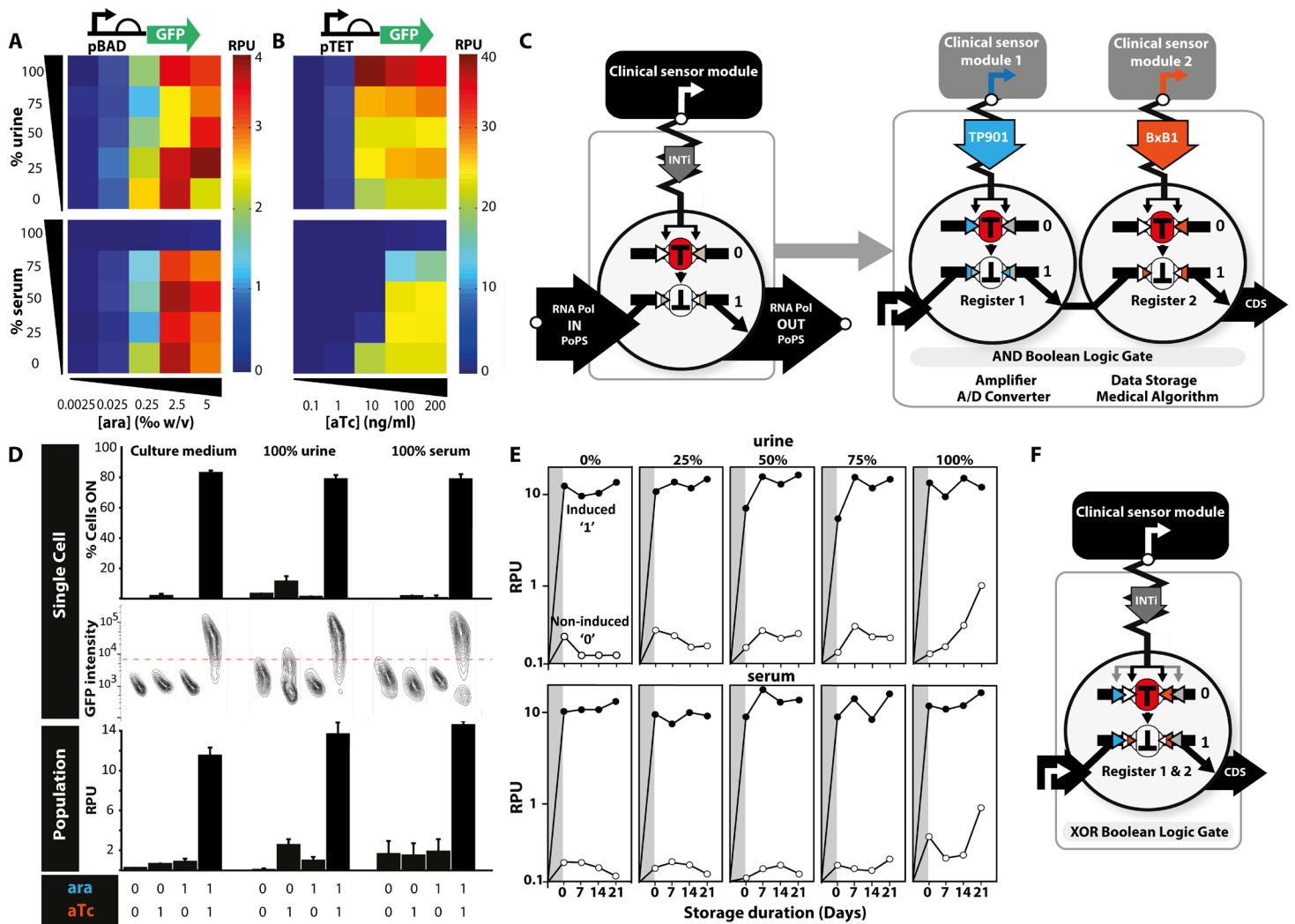


Figure 2.7: Control of gene expression, multiplexed signal processing, and long-term data storage within bacteria operating in human clinical samples. (A) Response of cells that contain the pBAD-GFP plasmid to increasing concentrations of ara in various dilutions of urine (top) or serum (bottom). Arabinose concentration is shown in parts per thousand (%; w/v). (B) Response of cells containing the pTET-GFP plasmid to increasing concentration of aTc. Fluorescence intensities were normalized to relative promoter units (RPU). The rainbow color key from blue to red depicts increasing signal intensities measured in RPU. (C) Architecture and functional composition of transcriptor-based digital amplifying genetic switches. The clinical sensor promoter drives integrase expression, which inverts a transcriptor module that controls the flow of RNA polymerase (RNA pol) along the DNA. Two transcriptors that respond to different signals can be composed in a series to produce an AND gate. A/D, analog to digital. DNA register 1 and DNA register 2 are placed between a constitutive promoter and the coding sequence (CDS) of a fluorescent reporter (sfGFP or mKate2 in this study). In this study, logic registers and sensor-integrase modules are encoded in two distinct low copy plasmids. (D) Operation of an AND gate at 25°C, at various dilutions (0, 100%) of human urine and serum in response to ara (0.5% w/v) and aTc (200 ng/ml). The 0 or 1 values symbolize absence or presence, respectively, of a particular inducer. Population (bottom, RPU) and single-cell (top) fluorescence intensity measurements are shown. The middle row shows raw flow cytometry data (x axis: side scatter). Error bars indicate SDs from three independent experiments, each performed in triplicate. (E) Stability of functional memory in various dilutions of urine (top row) or serum (bottom row) in living cells. The AND gate was switched with 0.5% (w/v) ara and aTc (200 ng/ml). Cells were then kept at 4°C for 7, 14, or 21 days and then grown overnight in fresh medium. For each medium concentration, GFP fluorescence in RPU is represented for non-induced cells (open circles) and induced cells (filled circles). The gray-shaded regions depict the duration of the period in which cells were exposed to the inducing signal. (F) Architecture of the XOR gate used in this study, as a single input TP901 and BxB1 addressable digital amplifying gene switch.

2.2.4 Systematic methodology to achieve tunable biosensing coupled to programmable signal processing

We then aimed at performing an analytical evaluation of bactorsensors for the detection of biological parameters of interest in clinical samples (i.e. urine and serum). We first chose nitrogen oxides (NO_x), which is a biomarker for various pathologies involving inflammation⁵²². For this purpose we mined for microbial molecular machineries that could be derivated for biosensing. We identified a nitrite/nitrate-sensitive transcriptional promoter⁵²³, pYeaR, from the year-yaog operon of *E.coli*. This promoter is regulated in *E coli* via the action of NsR, a member of the Rrf2 family of transcription factors that naturally regulate the response of *E coli* to nitric oxides (i.e. NO_x sensitive repressor). Here we thus rewired the natural cellular metabolic machinery to provide useful biosensing application. Using a GFP reporter, we measured the transfer function of pYeaR promoter with increasing concentrations of NO_x at various urine and serum dilutions (**Figure 2.8A**). The pYeaR activation threshold decreased with increasing concentrations of urine or serum and was activated in 100% urine without the addition of NO_x, probably due to the presence of endogenous NO_x. Moreover, pYeaR was totally inhibited in 100% serum. These results highlight the traditional sensitivity of cell-based biosensors to context perturbations that need to be overcome for successful medical applications.

In fact, direct transcriptional coupling using traditional promoter-reporter architectures often show weak, leaky activity that lead to nonspecific expression and overall unreliable biosensing. Therefore, our main goal was to develop a systematic method to wire an arbitrary transcriptional signal to transcriptor operation, by that mean obtaining a control on concentration switching threshold, programmable signal processing, digitization and multiplexing. We developed an engineering solution consisting of a directed evolution strategy⁵²⁴ capable of coupling a sensing module (i.e. promoter of interest) to transcriptor operation (i.e. integrase of interest expression) in a tunable way.

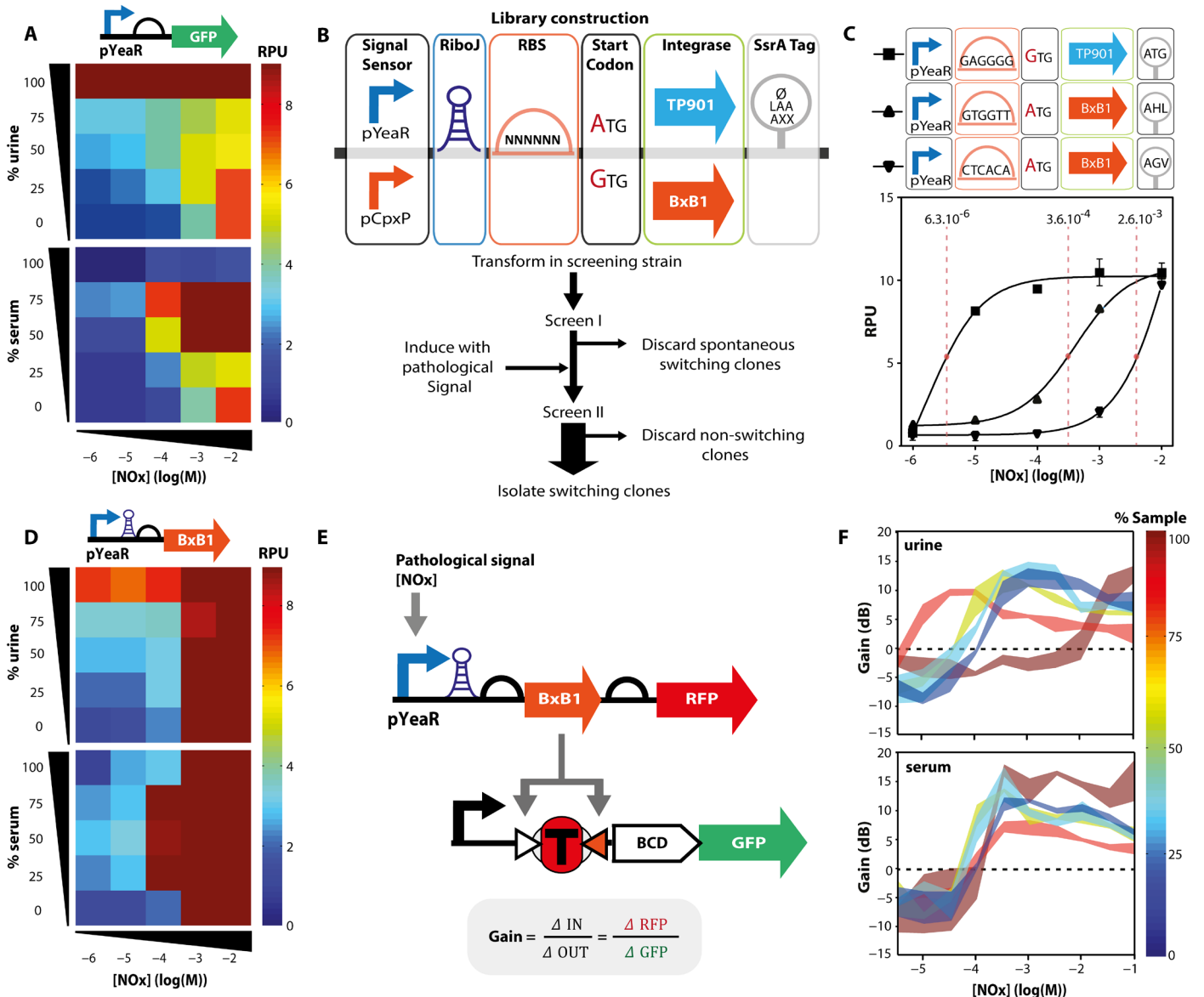


Figure 2.8: Thresholding, digitization, and amplification of biologically relevant signals in clinical samples using amplifying genetic switches. (A) Characterization of the NO_x-responsive promoter pYeaR driving expression of GFP in various dilutions of serum and urine. The rainbow color key from blue to red depicts increasing signal intensities measured in RPU. (B) Workflow for connecting biological signal–responding promoters to amplifying digital switches. An integrase expression cassette library driven by promoters of interest was built by introducing combinatorial diversity in RBSs, start codons, and C-terminal SsrA degradation tags. Libraries were transformed in a screening strain, spontaneously switching clones were eliminated, and the remaining cells were induced with the biological signal of interest. Switching clones were identified on a plate reader or using a fluorescence-activated cell sorter and isolated (see Materials and Methods for details). (C) Multiple switching thresholds for biomarker detection. Clones were isolated from the various pYeaR libraries and characterized. Midpoint switching values are indicated. Variation in sequences among the isolated clones is depicted in the upper panel, along with the correspondence between graphs symbols and a particular switch sequence. (D) Digitization and amplification of the NO_x signal in urine and serum using amplifying digital switches. Cells cotransformed with pYeaR switch and exclusive OR (XOR)–GFP gates (42) were induced with NO_x, and bulk fluorescence was measured on a plate reader. (E) The plasmid to measure amplification of NO_x input consists of bicistronic BxB1-RFP cassette driven by pYeaR. This construct was cotransformed with a XOR-GFP gate to enable the precise simultaneous measurement of fold change in input control signal (RFP) and output signal (GFP) after induction with NO_x. (F) In vivo molecular pathological signal amplification in clinical samples. Gain in decibels was calculated as the 10log of the RFP/GFP ratio. The line thickness represents the SD over three independent experiments.

Our goal was to design and build a transcriptor controller driven by biomarker responsive promoters. In order to obtain functional synthetic circuits, we needed to finely tune transcriptor operation so that translation levels of integrases match relevant clinical dynamic ranges for our applications. To this aim, we developed a directed evolution approach to introduce combinatorial diversity in synthetic integrase expression cassettes. The strategy we developed relied on the combination of randomized regions that regulate integrase gene expression and enzymatic kinetics, coupled with flow cytometry automated bacterial library screening (**Figure 2.8B**).

For prototyping, we first used pYeaR promoter which we cloned upstream of the Bxb1 or TP901-1 integrase genes. We build combinatorial libraries of pYeaR expression constructs driving expression of integrase BxB1 and TP901 respectively, by randomizing the RBS, the initiation codon (i.e. ATG/GTG) and a C-terminal SsrA tag^{459 525} (AXX) (thereby producing 1179648 variants). SsrA are protein degradation tags, which are short peptide sequences naturally used by the cell's machinery to trigger protein degradation (i.e. through ClpXP or ClpAP protease recognition). In fact, SsrA tags modulate protein functionality half-life. The rationale behind the use of SsrA tags is to introduce a tunable parameter related to integrase concentration and activity kinetics within the cell. Here we used synthetic variants of SsrA degradation tags that have been engineered⁵²⁶ (i.e. AAV, ASV, LVA, LAA, or AXX peptides).

The library was transformed in a screening strain containing the boolean integrase based Transcriptor device. In all our single inputs experiments, we used the XOR logic gate that behaves as the simplest TP901 and BxB1 adressable digital amplifying gene switch. Its architecture is depicted in **Figure 2.7F** and **2.9A**. The bacteria were then induced at selected biomarker concentration. Positive and negative screen steps were carried out to enable the isolation of suitable clones (**Figure 2.8B**). Additionally, we developed a more general workflow capable of accommodating user defined thresholds driving transcriptor operation (**Figure 2.10**). In this approach, one can precisely tune the switching thresholds of an integrase based transcriptor, which is defined by the biomarker concentration region ranging from low and high limits (**Figure 2.10**).

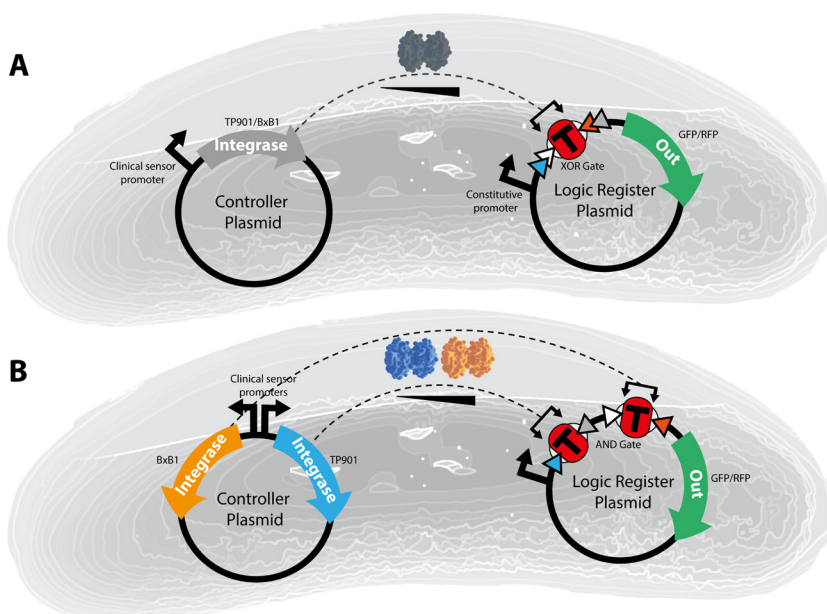


Figure 2.9: Detail of the molecular mechanisms of bactorsensors operation described in this study. (A) Single input digital amplifying gene switch relying on a XOR boolean integrase logic gate plasmid controlled by BxB1 or TP901 integrases according to an analogic signal coming from a clinical sensor promoter of interest on a controller plasmid. (B) Integrating two relevant molecular inputs into a digital amplifying AND boolean integrase logic gate. Detail on molecular biology can be found in **Materials and Methods**.

This strategy relying both on transcriptional and post-translational control permitted us to isolate integrase expression cassettes suitable satisfying given transfer functions thereby allowing their use for specific applications. To our knowledge, this approach constitutes the first successful attempt and description of rational tuning of biosensing threshold in engineered living cells with this precision. Although the use of SsrA tags, randomized RBS and initiation codon have been precedently used in various biotechnological applications to modulate protein degradation and expression levels respectively^{459 462 527}, the combination of these bioengineering tools had never been proposed elsewhere. Here, we show that this methodology enables to achieve unprecedented combinatorial diversity, allowing improved exploration of the design space. The novelty of this approach lies in the capacity to generate and sample a library efficiently to isolate genetic circuits according to precise dynamic range specifications. In addition, our results suggest that this strategy could be systematically used to reliably connect multiple control signals to integrase based synthetic gene circuits, which would be of significant importance if one considers building complex multi-input circuits.

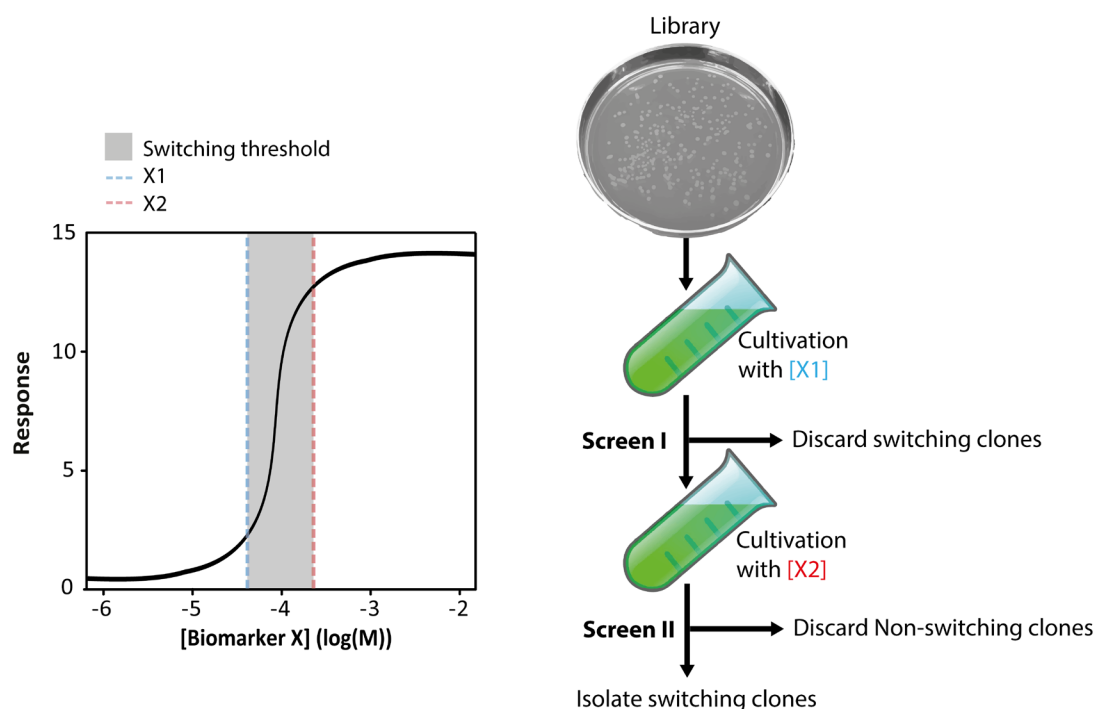


Figure 2.10: Workflow for engineering switches responding to biological signals over user-defined thresholds. In order to tune a switch to be activated at a specific signal concentration, cells from library need first to be grown at a given concentration [X1] below the desired switching threshold to eliminate all devices from the library that respond to lower concentrations of signal. In a second step, cells are grown at the desired switching concentration [X2] and switching clones are isolated. This approach differs from the one presented in Figure 2.8 in which we grew cells without any inducer, discarded spontaneous switching clones, then induced at a given concentration of inducer resulting in the obtention of switches activated along the whole inducer concentration range between 0 and the given concentration.

2.2.5 Thresholding, digitization, and amplification of biologically relevant molecular signals using digital amplifying genetic switches

We then tested whether transcriptor-based digital amplifying switches could improve the detection of signals of clinical interest. From this library, we selected and characterized three switches that contained variations in the RBS, start codon, integrase type, and SsrA proteolysis tag^{459 525}. Switches were activated at different NOx thresholds that spanned several orders of magnitude (**Figure 2.8C**). These data suggest that digital amplifying genetic switches could be systematically tailored to detect a specific biomarker over defined pathological thresholds that meet clinical requirements (**Figure 2.10**).

We mapped the transfer function of one of the switches at various sample dilutions (**Figure 2.8D**) and observed signal digitization and marked improvement of the signal-to-noise ratio compared to the pYeaR-GFP construct. Interestingly, the inhibitory effect of 100% serum on NOx detection was overcome, although signal interference was still observed in 100% urine. Using an amplification reporter system, we quantified pYeaR switch-mediated signal amplification across a range of signals and sample concentrations (**Figure 2.8E**) and measured maximum gain values between 10 and 15 dB. These results show that digital amplifying switches increase the robustness of cell-based biosensing systems and could thus enable the development of clinically compliant biosensors.

2.3 Developing polymer chemistries to immobilize batosensors within portable formats

Different strategies have been proposed as a way to stabilize and ease the use of microbial biosensors: encapsulation, covalent binding, adsorption, and cross-linking on various substrates. Although interesting formats have been proposed like immobilization on paper strips¹⁷³, we reasoned that the encapsulation of batosensors in hydrogel beads could increase the robustness and preserve viability and response characteristics of sensing cells under the harsh environmental conditions they are exposed to (i.e. urine, serum), could prevent their spread, and could enable the combination of computing operation in different population of beads at the same time. This strategy could improve the analytical performance, increase the cell density, handling, storage and preservation of microbial biosensors without the need of continuous cultivation, and make them suitable for integration into deployable and *ready to use* devices for unskilled personnel¹⁷².

We thus sought to develop a bacterial encapsulation system that could evolve towards a clinically compliant format. For this purpose, spherical polymeric matrices appeared as promising for their chemical versatility and large range of functionalities. However, the identification of an appropriate matrix was of utmost importance for successful applications. Specifically, we sought to identify polymer chemistries that would ensure high durability in aqueous solvent, high cell viability, small molecule and gas permeability, and appropriate optical properties (i.e. optically clear to support colorimetric and fluorescence readout). Both natural (agar, agarose, alginate or chitosan) and synthetic (polyacrylamide, polyurethane or polyethylene glycol) polymers have been exploited for cell immobilization⁵²⁸⁻⁵²⁹. Importantly, the first materials usually show poor mechanical resistance and durability, while the later often exhibit cell toxicity⁵²⁹. Amongst synthetic polymeric hydrogel, we identified polyvinyl alcohol (PVA) has a promising non-toxic, optically clear, cheap, and highly durable material. Spherical beads of this polymer could be fabricated by crosslinking PVA droplets in a saturated solution of Boric Acid (i.e. Hydroxyl groups of PVA are crosslinked by borate ions). However, this method suffered from low kinetics of polymerization, which affected bactsensor viability (i.e. pH=4 solutions have time-dependent lethal effect on *E. coli*) while agglomeration of beads complexified the fabrication process. In addition, PVA polymers suffer from poor gas and small molecule permeability, which would affect both mass transfer of biomarkers to the bactsensors and cells viability. We reasoned that a composite material would compensate for disadvantages, and showed we could overcome these drawbacks via the addition of alginate combined with PVA.

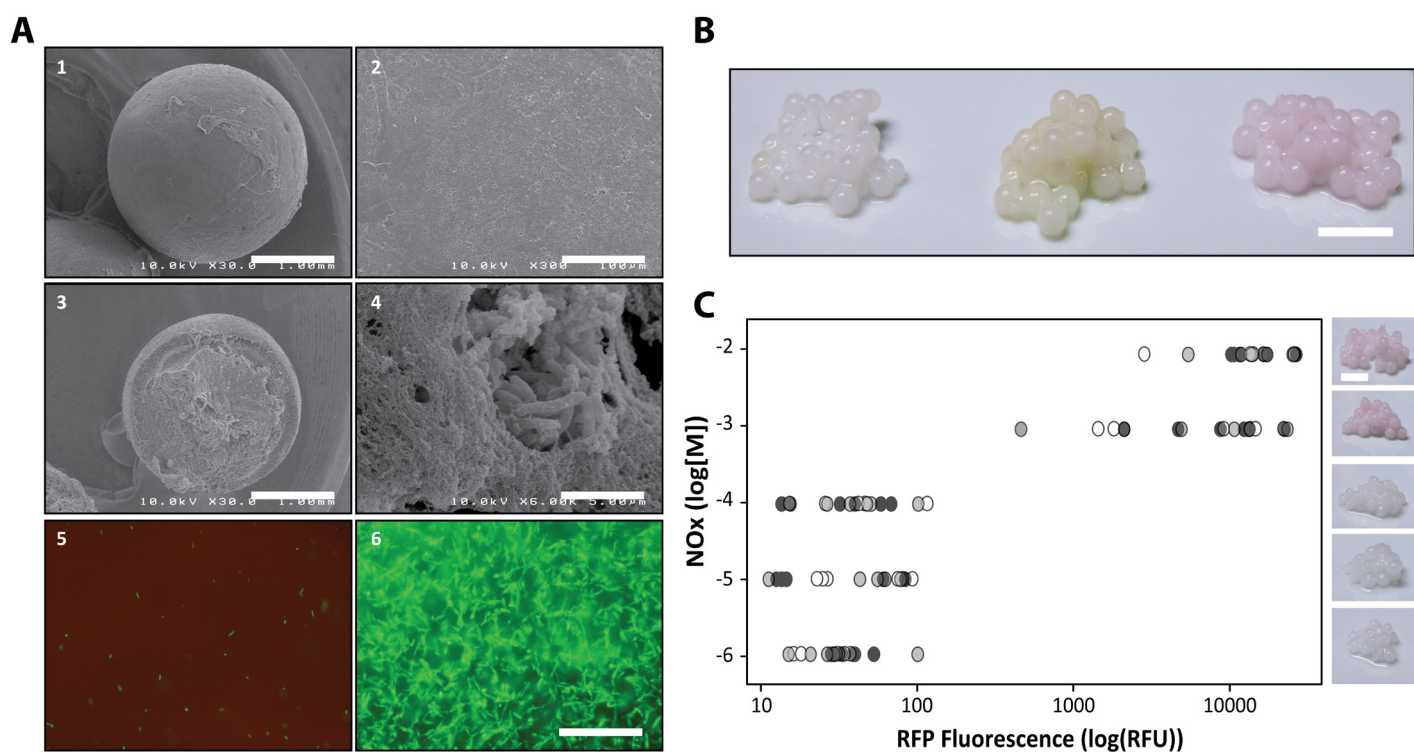


Figure 2.11: Insulation of Bactsensor in hydrogel beads, and pathological signal detection. (A) SEM microphotographs of PVA-Alginate uncut (1, 2) and cut (3, 4) beads entrapping bactsensors. Fluorescence imaging of uncut beads (5, 6): Cells were co-transformed with pYeaR-BxB1 controller 1 and XOR-BCD-GFP gate and encapsulated as described. Beads were then incubated overnight in Azure media (5) or Azure media with 0.01M NO_x (6). Beads are optically clear, their surface appears porous and permits mass transfer of inducers solutes to the inside. (B) Left: Uninduced beads. Middle: NO_x-induced beads with GFP reporter. Right: NO_x-induced beads with RFP reporter. Scale bar, 0.5 cm. (C) Transfer function of beads entrapping bactsensors: XOR-BCD-RFP gate was co-transformed with pYeaR-BxB1 controller 1, and cells were encapsulated as described. Beads were then incubated overnight in Azure media with different concentration of NO_x, and RFP fluorescence was measured after 24 hours. RFU, relative fluorescence unit. Detection thresholds for urinary dipsticks and for bactsensor are indicated. Pictures of the beads at various inducer concentrations are shown. Scale bar=0.5 cm.

We thus proposed to immobilize bactosensors within spherical polyvinyl-alcohol/alginate hybrid beads. The beads were produced by crosslinking droplets of PVA/alginate with boric acid solution (for a short time to prevent damage to the immobilized bactosensor), followed by esterification of the PVA with sulfate for further strengthening. To change the poor gas permeability of PVA gels, sodium alginate was thus added to the PVA solution and the saturated boric solution was supplemented with calcium chloride to produce polymerization of the alginate. Later, a treatment with a sulfate solution induces PVA crystallite formation permitting to obtain a more porous structure allowing mass transfer of solutes inside the bead, along with a more stable polymer with increased strength. Briefly, this mechanism relies on sulfate ions that destabilize the hydrogen bonds between PVA and H₂O, thereby enhancing the formation of hydrogen bonds within the PVA polymer, inducing a PVA crystallite⁵³⁰. From the PVA/alginate mixture, we generated size calibrated droplets using a controlled microfluidic flow. This technique enabled the generation of monodispersed, clear and stable beads with an average diameter of ~2 μm. Furthermore, we obtained and validated the full operability and viability of the bactosensors immobilized in this way (**Figure 2.11**).

2.4 Analytical evaluation of bactosensors for the detection of biomarkers in clinical samples

2.4.1 Detection of a metabolized biological signal [glucose] in clinical samples

Glucose is a biomarker of clinical interest whose blood levels can be used for the monitoring of diabetes (high blood glucose or glycemia) and whose presence in urine (glycosuria) marks the onset of or presence of uncontrolled diabetes. Point-of-care technologies that enable clinicians to detect glycosuria or continuously monitor glycemia remotely can greatly improve and simplify care of diabetic patients⁵³¹. The fact that glucose is one of the primary carbon sources metabolized by bacterial cells makes it a challenging molecular signal to monitor. To perform glucose detection, we identified the pCpxP promoter from the CpxAR regulon from *E. coli* metabolism, which I naturally used to integrate environmental information. We derivated this Cpx pathway as a driver of target gene expression, which is activated in the presence of glucose, pyruvate, or acetate⁵³². pCpxP showed a high basal level of expression in bacterial growth medium and human serum and a low signal to-noise ratio (**Figure 2.12A**, maximum fold change ~1.5 in medium, ~2.2 in urine, and ~1.7 in serum).

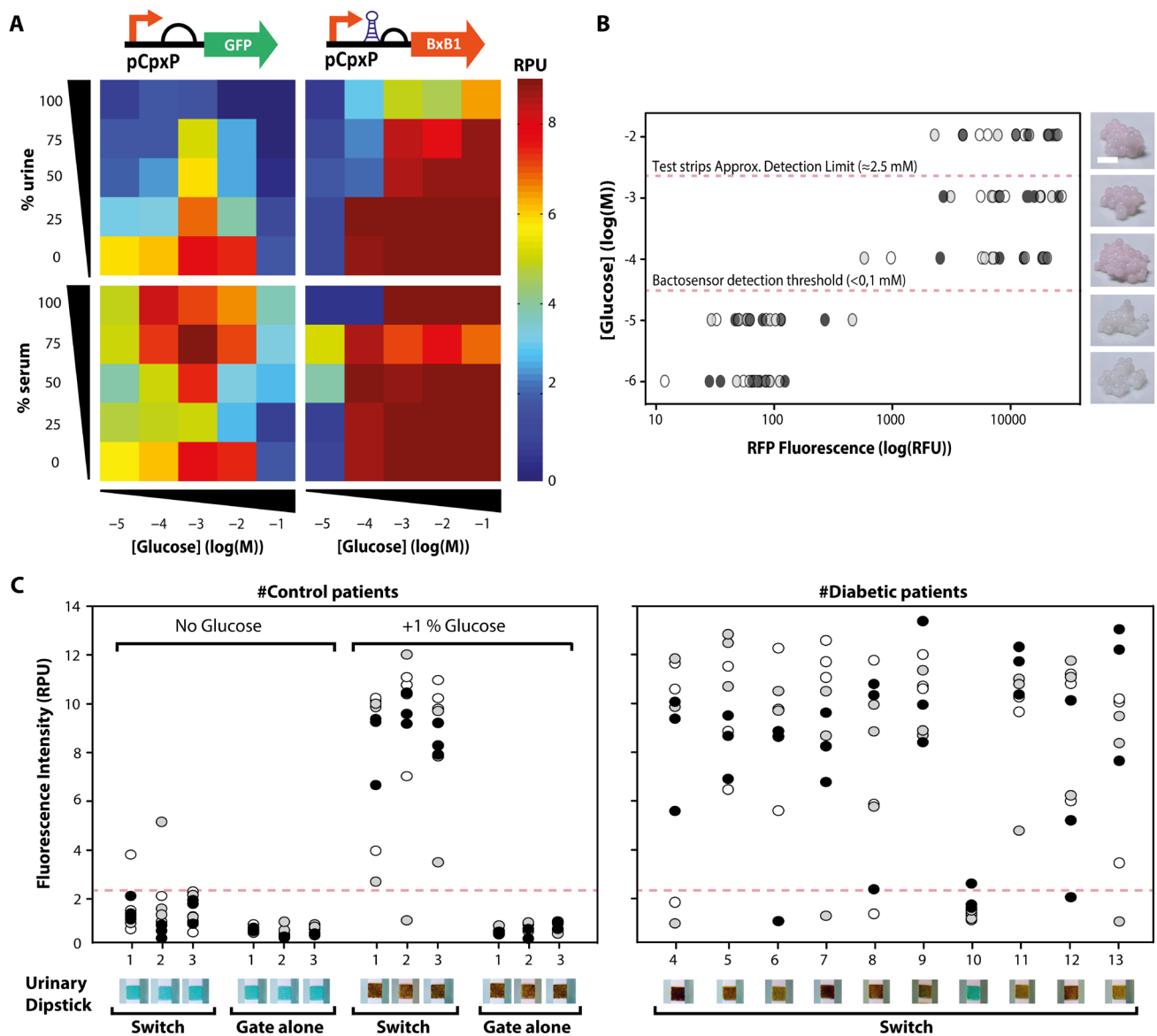


Figure 2.12: Bactosensor-mediated detection of pathological glucose levels in clinical samples from diabetic patients. (A) Comparison of the response of pCpxP-GFP (left) and pCpxP-Bxb1 (right) constructs to increasing concentrations of glucose in various concentrations of urine (upper panel) or serum (lower panel). (B) Operability of amplifying digital switches encapsulated in PVA/alginate beads. Beads that contained cells cotransformed with the XOR-RFP gate and pCpxP-TP901 were incubated in culture medium supplemented with various concentrations of glucose, and RFP fluorescence was measured after 24 hours. RFU, relative fluorescence unit. Detection thresholds for urinary dipsticks and for bactosensor are indicated. Pictures of the beads at various inducer concentrations are shown. Scale bar, 0.5 cm. (C) Bactosensor-mediated detection of abnormal glycosuria levels in clinical samples from diabetic patients. PVA/alginate beads encapsulating cells transformed with both the pCpxP-TP901 controller and the XOR-GFP gate, XOR-GFP gate alone, or the reference construct J23100-GFP were incubated in urine samples. Left panel: three glucose-negative samples from independent individuals and three positive controls [same samples supplemented with 1% (w/v) glucose]. Right panel: Urine samples from 10 nonstabilized individual diabetic patients. GFP fluorescence was measured after 24 hours. Response to glucose was compared with standard urinary dipsticks (lower panels). The lower panels show the glucose reactive band of the urinary dipstick. In the absence of glucose, the band is blue. When glucose is present, the band turns brown. Data from three different experiments performed on different days are depicted by black, gray, and white circles. Each circle corresponds to one replicate (three replicates for each experiment).

Moreover, pCpxP was inhibited by increasing urine concentrations and glucose concentration greater than 10^{-2} M. For the latter case, we confirmed, by kinetic assays, the time-dependent inhibition of pCpxP putatively as a result of a glucose-induced drop in the pH of the medium⁵³³ (**Figure 2.13A**). We next built a pCpxP switch and again observed a marked improvement in the signal-to-noise ratio (Fig. 4A, maximum fold changes ~ 12.4 in medium, ~ 12.6 in urine, and ~ 20.6 in serum) and a near-digital switching (that is, the system responded in a nearly all-or-none fashion). Response of the pCpxP switch to glucose was detectable up to 100% urines, indicating that a low signal produced by the promoter in these conditions was detected, amplified, and stored by the switch. The transient pCpxP activity at high glucose concentrations was also detected and stored by the pCpxP switch (**Figure 2.12A** and **Figure 2.13B**). Therefore, the detection of multiple clinically relevant signals can be systematically improved by digital amplifying switches. Finally, as a proof of concept, we performed dual detection and multiplexed signal processing of clinical biomarkers by building two-input logic gates controlled by NOx and glucose and performing arbitrary computation processes (**Figure 2.13D**).

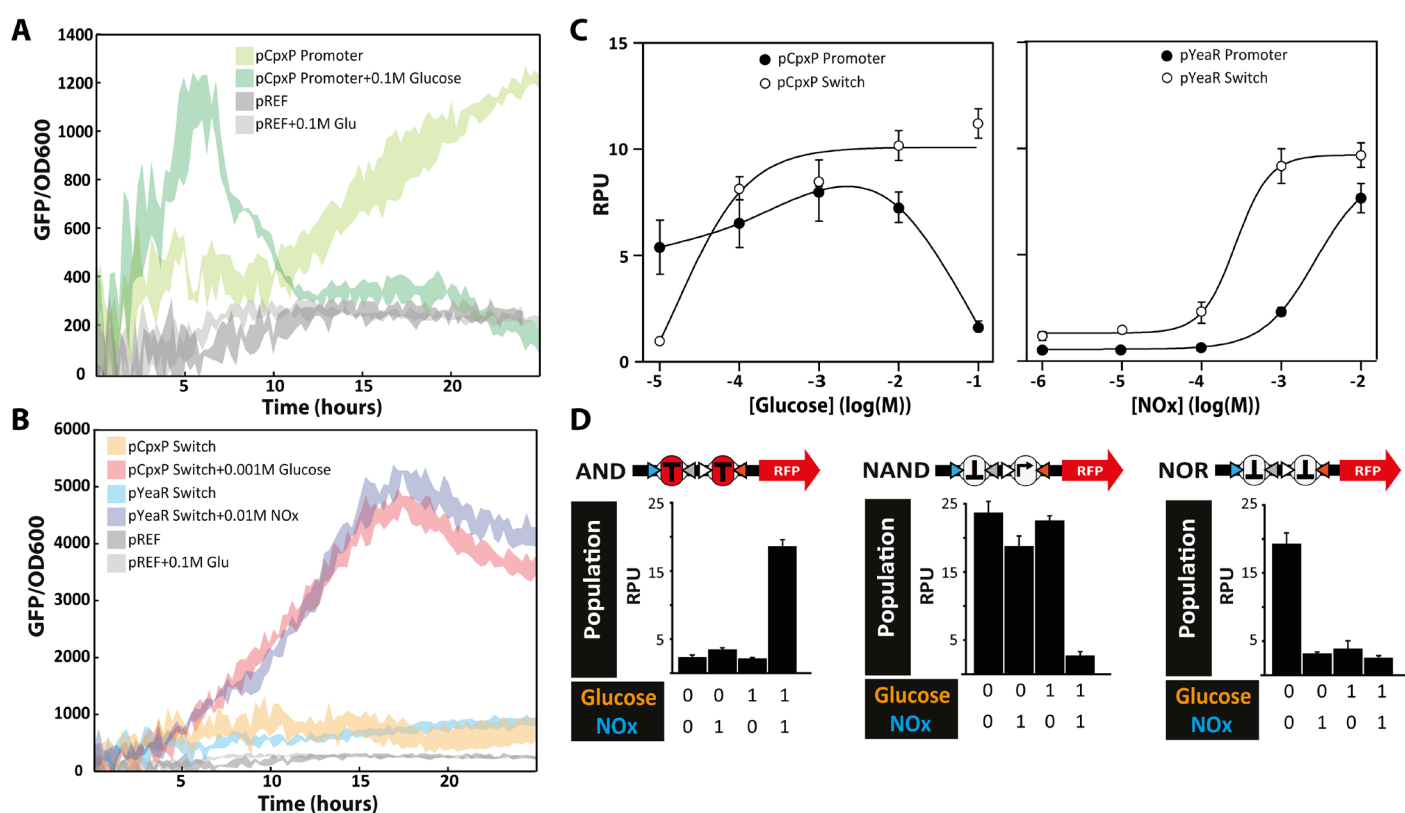


Figure 2.13: Kinetic measurements and transfer functions of promoters and switches. (A) Plate reader fluorescence measurements of pCpxP promoter and reference promoter (pREF), as indicated. At an inducer concentration of 0.1M Glucose, the pCpxP promoter is repressed after 5 hours. Signal from pCpxP promoter is known to be affected by different host/environment perturbations: alkaline pH, surface attachment or accumulation of misfolded periplasmic proteins are reported to inactivate transcription^{532 533 534}. (B) Plate reader fluorescence measurements of pCpxP-BxB1 controller and pYeaR-BxB1 controller 1 co-transformed with XOR-BCD-GFP gate, and induced with 0.001M Glucose and 0.01M NOx respectively, as indicated. An increased in output fluorescence is detectable around 5 hours. (C) Transfer function of pCpxP and pYeaR promoters and switches. A slight change in control signal (promoter response to inducer) triggers a digital response of the switch. (D) Multiplexing detection of glucose and NOx, with AND-BCD-RFP, NAND-BCD-RFP and NOR-BCD-RFP Boolean integrase logic gates. Here logic gates operation is controlled by the pYeaR-TP901/pCpxP-BxB1 dual controller plasmid (NOx/Glucose inducible). We measured RFP fluorescence at the population level after overnight induction performed with glucose (1mM) and NOx (1mM) in LB medium.

2.4.2 Bactosensor mediated detection of pathological glycosuria in diabetic patients samples

To assess the relevance of digital amplifying switches for disease detection in a clinical assay, we sought to develop a proof-of-concept biosensor that detects endogenous levels of a pathological biomarker in clinical samples from patients. As a preliminary validation, we aimed at detecting glycosuria using the pCpxP digital amplifying gene switch. To do so, we used the prototype clinical format we previously developed and encapsulated viable bactosensors in PVA/alginate hydrogel beads as described (**Figure 2.12B**). First, using the red fluorescent protein (RFP) as a reporter, we tested the response of beads that encapsulated the pCpxP or pYeaR switch to increasing inducer concentrations in culture medium and observed digital switching detectable with the naked eye (**Figure 2.12B**). The pCpxP switch was activated at a threshold concentration under 0.1 mM glucose, outperforming the detection limit of urinary dipsticks, the gold standard point-of-care test for glycosuria, by an order of magnitude. We then tested the beads in urine samples. The pCpxP switch beads produced a robust and specific response in nonpathological urine samples exogenously supplemented with glucose (**Figure 2.12C, left panel**). Finally, we tested the pCpxP switch beads in individual urine samples from 13 patients diagnosed with diabetes but not yet stabilized (**Figure 2.12C, right panel**). The assay reliably detected glycosuria in samples from diabetic patients, with a sensitivity of 88.9% and a specificity of 96.3% (**Figure S2.6**). We observed some variability in beads response, which we attributed to our bead fabrication process. Improvements in the encapsulation process should increase the reliability of our assay and reduce bead-to-bead variability. Nevertheless, our system was capable of reliably detecting the presence of endogenous glucose in urine from 12 different diabetic patients, suggesting that bactosensors are relatively robust when faced with interindividual variations in sample composition. Together, our data demonstrate that digital amplifying genetic switches can enable bactosensors to perform clinical assays and detect endogenous biomarkers of disease in patient samples.

2.4.3 Developing standards to quantify bactosensor robustness in clinical samples

We then aimed at establishing a quantitative framework with which to evaluate the robustness of the bactosensor response against clinical sample-induced perturbations (**Figure 2.14**). At each inducer concentration and for each clinical sample dilution, we quantified the change in signal relative to cells grown in culture medium. We proposed that the relative changes in signal values could be averaged to obtain a global robustness score (RS), which was inversely proportional to the robustness of the biosensor against sample-induced perturbations. For pCpxP, use of the switch reduced RS values from 0.6 to 0.3 in urine and from 0.27 to 0.20 in serum. For pYeaR, RS values decreased from 2.1 to 0.44 in urine and from 1.15 to 0.69 in serum. Using digital amplifying switches thus systematically improved the robustness of the bactosensor response

against sample perturbations. Part of this improvement in robustness also resulted from the use of standard parts for the translational control of the switch output⁴⁶⁴ (**Figure 2.5**).

We also used an innovative approach to quantify the improvement in signal digitization conferred by the digital amplifying switches. We measured the digitization error rate (DER), which is defined as the combined probability of scoring a false-positive or a false-negative when using a digital classifier like the bactosensor⁴⁵⁸. For both pYeaR and pCpxP, we measured promoter-only constructs and promoter-switch constructs in the presence of minimal and maximal inducer concentrations (**Figure S2.3, S2.4, S2.5**). We found that sensors that incorporated digital amplifying switches generally displayed a reduced DER, demonstrating the improvement in signal digitization provided by the switches.

Together, these results demonstrate that digital amplifying switches can markedly improve the reliability of the detection of clinically relevant signals in clinical samples.

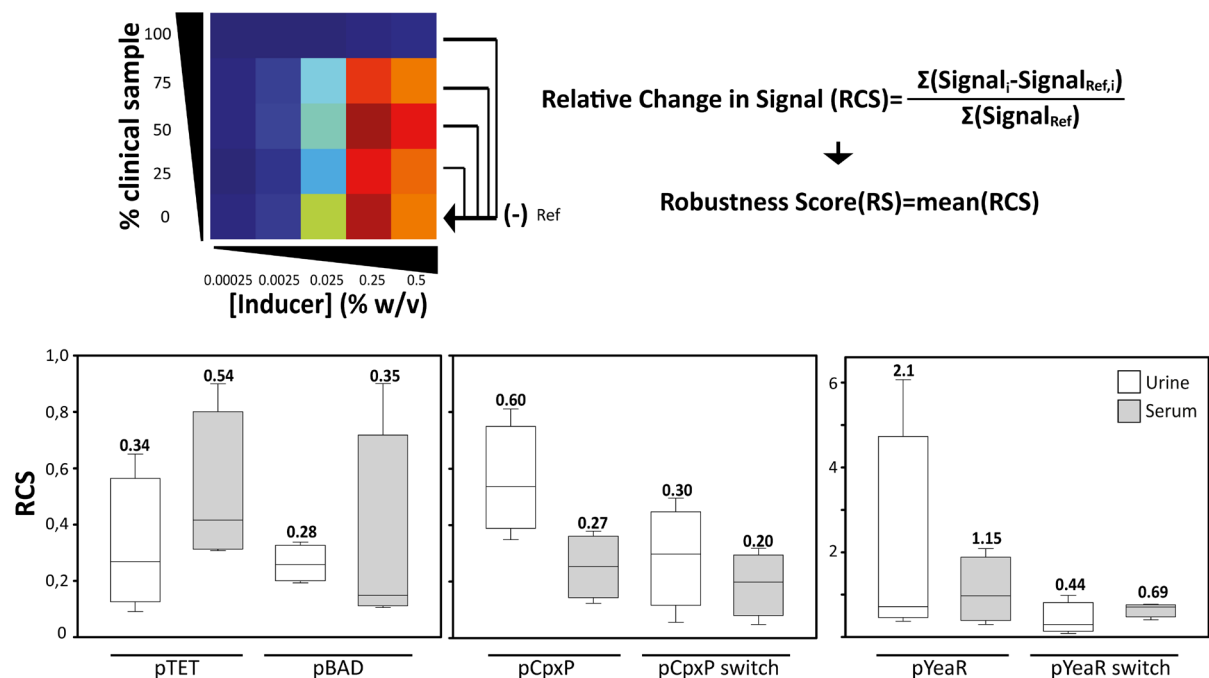


Figure 2.14: Evaluation of the robustness of inducible systems against clinical media induced perturbation (urine and serum). Transfer functions of inducible systems (pTET/pBAD promoters, pCpxP/pYeaR promoters, pCpxP/pYeaR switches) were measured after overnight induction in different dilution of clinical sample (urine or serum) in culture media. At each sample dilution (25, 50, 75, 100% clinical sample) we calculated the Relative Change in Signals (RCS) by summing the difference in output signal between perturbed condition (introduction of clinical sample) and reference conditions (only culture media) with identical inducer concentration. We then defined a robustness score (RS), as the mean of the 4 RCS obtained for one heat map. The RCS are plotted for each inducible system, with the RS value specified in bold. Digital amplifying genetic switches have lower RS, and enable more robust response in clinical context.

2.5 Engineering bactosensor quorum sensing for intelligent sepsis diagnosis

In the last sections, I presented the engineering of the bactosensor platform, and provided a first proof-of-concept prototype for disease detection in clinical samples. Now, I propose to provide a novel diagnostic perspective for the timely diagnosis of infectious diseases, which constitutes an important clinical challenge (where 4-5 hours is a major improvement of current technologies). Here, I discuss ongoing work to program bactosensors for sepsis diagnosis.

2.5.1 Sepsis: clinical significance and diagnostic challenge

Sepsis is a localized or systemic microbiological infection associated with a systemic inflammatory response syndrome, which can evolve towards multi-organ dysfunction in case of severe sepsis⁵³⁴. Sepsis is an urgent clinical situation, and is among the most common causes of death in hospitalized patients worldwide with a continuous increase in incidence of approximately 5–10% per year, while the mortality rate is approaching 50%^{535 536}. As microbiological investigations are commonly negative and late, sepsis still mainly remain a clinical diagnosis. After each hour of delay in medical support mortality increases by 10%, the rapid diagnosis of sepsis is thus of outstanding importance⁵³⁷. The current *gold standard* of sepsis diagnosis is based on blood cultures (i.e. the detection of viable microorganisms present in blood). However, this technique gives positive results only in 20% of cases and even if positive, results are obtained too late to influence decision making. Empirical antibiotic therapies are started without a definitive microbial result based on symptomatology. Factors such as antibiotics initiated before blood sampling or the presence of fastidious pathogens may have a negative impact on the diagnostic yield of blood cultures even when a bloodstream infection is strongly suspected⁵³⁸. Furthermore, other factors can influence blood culture sensitivity: blood volume, time from sampling to incubation, fastidious uncultivable pathogens and antimicrobial therapy, turnaround time to definitive identification, and low sensitivity⁵³⁹.

Other strategies to sepsis diagnosis have been proposed to tackle these limitations. Nucleic acid-based diagnostic technologies such as PCR have been proposed to shorten the time required for pathogen detection, but this technology suffers from several drawbacks (i.e. mainly contamination, high costs, inhibitors in the blood sample)^{540 541}. Proteomic technologies have also been proposed as they are able to identify pathogens allowing a definitive identification, but have extremely high resource requirements and low sensitivity. Phage-based assays have also been proven to have the potential to lead to interesting tools, even though they require bacterial amplification to reach the high detection threshold²⁰⁸.

Therefore, current technologies suffer from major drawbacks such as high cost and cumbersome procedures. These technical limitations cannot be eliminated by improvements of current

diagnostic techniques, and other approaches are needed. I propose that innovative cell-based diagnostic strategies could provide non-culture-based techniques for the diagnosis of sepsis. Valuable alternatives to standard approaches for sepsis diagnosis may be offered by synthetic biology, particularly using cell-based biosensing strategies, as it allows a thorough engineering of *smart* cells able to perform sense and compute operations. I propose that bactorsensors could be programmed to perform biodetection of sepsis associated biomarkers, thereby bringing a cost-effective, rapid, disposable and *intelligent* diagnostic tool.

2.5.2 Bactorsensor mediated sepsis detection

Very few sepsis biomarkers have been used for clinical diagnosis. Although PCT and CRP have been the most described, they have limited ability to discriminate between sepsis and other inflammatory conditions, and have insufficient specificity or sensitivity⁵⁴². However, detecting a direct product from pathogens in blood is possible and has already been described. Indeed, bacteria and fungi pathogens are known to communicate using chemical signaling molecules known as quorum sensing (QS) molecules, whose presence is known to be significantly correlated with infectious pathologies. QS systems are widespread among most common bacterial pathogens capable of inducing sepsis, particularly gram-positive *S. aureus*, *S. pneumoniae*, and gram-negative *E. coli*⁵⁴³. Moreover, QS system is activated at an early phase of infection, potentially permitting an early diagnosis of sepsis. Importantly, QS molecules are species specific: their presence enables the identification of the pathogen, which is the most clinical important parameter for appropriate therapeutic management. The QS molecules in gram-negative bacteria are constituted by a large class of N-acyl homoserine lactones (AHLs), while gram-positive bacteria use a wide range of species specific oligopeptides. However, a third class of molecules, autoinducer-2 (AI-2) is a near-universal QS signal between both gram-positive and gram-negative bacteria⁵⁴⁴. Last, all pathogenic fungi such as *Candida* species, secrete farnesol as their main QS molecule. Therefore, targetting three main QS molecules as molecular biomarkers: AHLs, AI-2, and Farnesol provide infectious aetiology with an appropriate decision algorithm (i.e. Gram+ bacteria, Gram- bacteria, or Fungi induced sepsis) (**Figure 2.15A**).

These signaling molecules involved in bacterial communication can serve as biomarkers for the diagnosis and management of bacteria-related infections, and their detection in clinical samples by employing analytical techniques is clinically relevant. Common techniques to detect quorum sensing molecules comprise different bulky, low throughput and tedious analytical chemistry approaches (i.e. GC-MS, HPLC-UV, LC-MS-MS...) that require time and personnel and high resources environments. Therefore, their detection in blood is a problem seeking for technological solutions.

Interestingly, cell-based biosensors have already been described and employed as cheap, portable yet accurate QS molecule detection systems^{545 546}. However, they have remained at an engineering state that cannot accommodate clinical requirements, portability, ease of use and robust and reliable biosensing. Therefore, this project proposes to engineer autonomous, expert and integrated bacterial biosensing platform for early sepsis diagnosis. It aims at rewiring the QS

machinery to our previously developed bactosensor platform, to provide an opportunity to use QS molecules as biomarkers of sepsis. In addition, we propose to program bactosensors to integrate an appropriate differential diagnosis algorithm capable of discriminating between all possible infectious aetiologies.

In the first part of our project, we aimed at repurposing natural QS machinery from different species in an engineered bacterial chassis (i.e. *E. coli*), and evaluate the feasibility and relevance of the project in terms of (i) QS molecules biosensing capabilities in clinical samples, and (ii) availability of robust parts and prototyping design.

2.5.3 Preliminary results

To evaluate the medical analytical capabilities of engineered QS systems, we first focused on Gram- bacteria biosensing systems, which are the most common infectious cause of sepsis. In other words, as a proof-of-concept preliminary study, we asked if it were possible to use bactosensors to detect relevant QS molecule in clinical samples. For this purpose, we reasoned that achieving a sensing system with largest AHLs specificities would enable to detect the various chemical ranges of AHLs, since Gram- bacteria are known to secrete molecules with various acyl lengths (i.e. C4 to C12, along with other chemical modifications such as *oxo* modifications).

From a molecular perspective, quorum sensing systems comprise a *two-gene* (i.e. an autoinducer synthase protein (“I”) coupled to a receptor protein (“R”). In other words, the receptor protein coding gene, which is a QS molecule conditional transcriptional activator of a specific promoter, can be cloned into a chassis strain to achieve QS conditional expression of a synthetic construct. This works for freely membrane penetrating QS molecules, such as AHLs and Farnesol, but also for AI-2 which is actively transported via a specific transporter in *E. coli*.

This required manipulating different QS machineries from different bacterial species, which had to be transferred and cloned in *E. coli* vectors. Moreover, it necessitated using in an *E. coli* chassis strain which is QS silent, bearing a deletion in the luxS system used to synthesize QS molecules (JW2662-1: Δ luxS768)⁵⁴⁷. For reporting, we used luminescence to achieve best detection limits via the luxCDABE operon from *V. fischeri*. These constructions, pSB401, pSB536 and pSB1142, beared respectively: LuxR from *V. fischeri* (best sensitivity for C6-C12 AHLs), AhyR from *A. hydrophila* (short chains AHLs), and LasR from *P. aeruginosa* (long chain AHLs)^{548 549 550}. Additionally, constructions bearing the QsR system from *P. aeruginosa*, and the LsR system from *E. coli*, which are natural Farnesol and AI-2 sensing systems, respectively, are under construction.

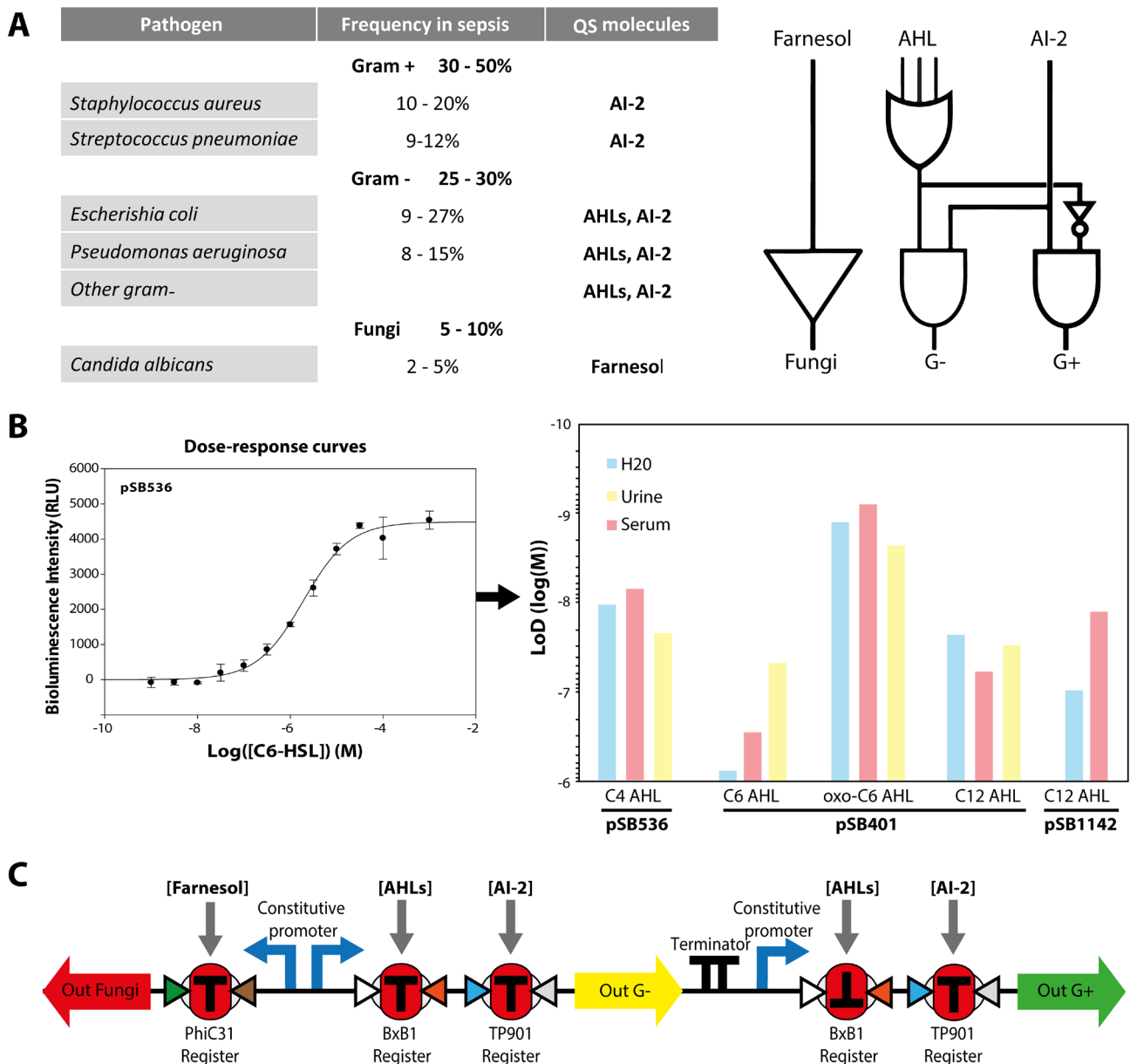


Figure 2.15: Sepsis diagnosis through bactorsensors mediated QS biosensing. (A) Left: Sepsis aetiology related to bacterial species, their clinical frequency and QS molecules secreted in each case. Right: Abstract Boolean representation of a classification algorithm giving access to infectious aetiology from 3 input QS molecules (B) Analytical limits of detection (LoD) were calculated using triplicate dose response measurements for each plasmid construction in each context (QS molecule spiker water, urine or seum), according to definition from *Armbruster et al.* ($LoD = \text{mean}_{\text{blank}} + 1.645(SD_{\text{blank}}) + 1.645(SD_{\text{low concentration sample}})$). AHLs were dissolved in acetonitrile before appropriate dilution in 100 μ l Azure medium with log phase culture of engineered bacteria. We incubated for 2 hours at 25°C under agitation before measurements (C) Design prototype of integrase based recombinatorial logic to implement the decision algorithm. As discussed in previous sections, Farnesol, AHLs and AI-2 sensing modules can be engineered to accurately couple to expression of PhiC31, BxB1 and TP901 integrases, which will exert their action on the logic register before reporting.

Using these constructs, we performed quantitative assays in AHL spiked water, urine or serum, and measured dose response curves as depicted in **Figure 2.15B**. We then calculated limits of detection obtained for each QS systems, molecules and sample type. We found that we could obtain a wide range, near nanomolar detection of all types of AHL in clinical samples. These sensibilitites are indeed well adapted for clinical biosensing since pathological concentration of AHLs have been previously described from the nanomolar to micromolar range⁵⁵¹. These AHLs biosensing systems are thus valid to be rewired to the bactosensor platform. In addition, we obtained very interesting reponse kinetics, with measurable outputs already present after 30 minutes. In the perspective of sepsis diagnosis, this is extremely interesting since current gold standard methods require cultivation times reaching 24 to 48 hours, which constitutes one of the most important drawback affecting patient management.

Next, we now aim at measuring AI-2 and Farnesol biosensing systems to evaluate their analytical capabilities. Unfortunately, the limits of this thesis did not allow for AI-2 and Farnesol biosensing modules to be built and tested. For this purpose, the same experimental plan as for AHLs will be carried out.

Then, reasoning with the same bactosensors architecture as before (i.e. supported by digital amplifying gene switches and logic gates), we are aiming at building prototype plasmid constructions recapitulating the diagnostic algorithm depicted in **Figure 2.15A**. Briefly, we will build sensing modules, namely AHL controller-BxB1 integrase (concatenated AhyR, LuxR, LasR systems), AI-2 Controller-TP901 integrase (LsR system), and a Farnesol controller-PhiC31 integrase-RFP (QsR system) (See Annexes for plasmid maps). These sensing modules will be engineered as previously described to exert an action on coupled logic registers and reporting modules (**Figure 2.15C**).

If successfull, we envision that we could evolve the sepsis biosensors towards a multiplexed Boolean integrase Logic device, capable of coordinating the full decision algorithm, as depicted in **Figure 2.15**. To these aims, plasmid constructions are underway and should yield results in a foreseeable future. As precendently, we have yet to evaluate robustness and reliability towards a clinical format.

2.6 Conclusion and discussion

The past decade has witnessed the development of innovative biodiagnostic technologies and biosensor approaches, promising a new era of fast, versatile, easy-to-use, and reliable point-of-care diagnostic devices⁹¹. However, within the biosensing device family, cell-based biosensors (despite their potential) have not been able to be translated into real-world clinical applications. Here, we bring bacterial biosensors closer to medical application by addressing some of the limitations that have hindered their translation to the clinics.

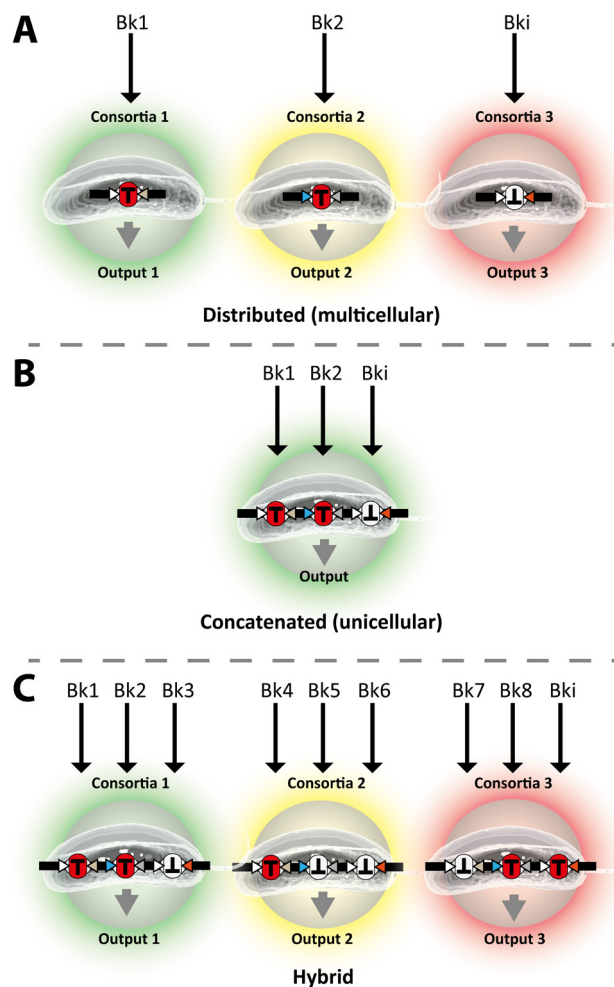
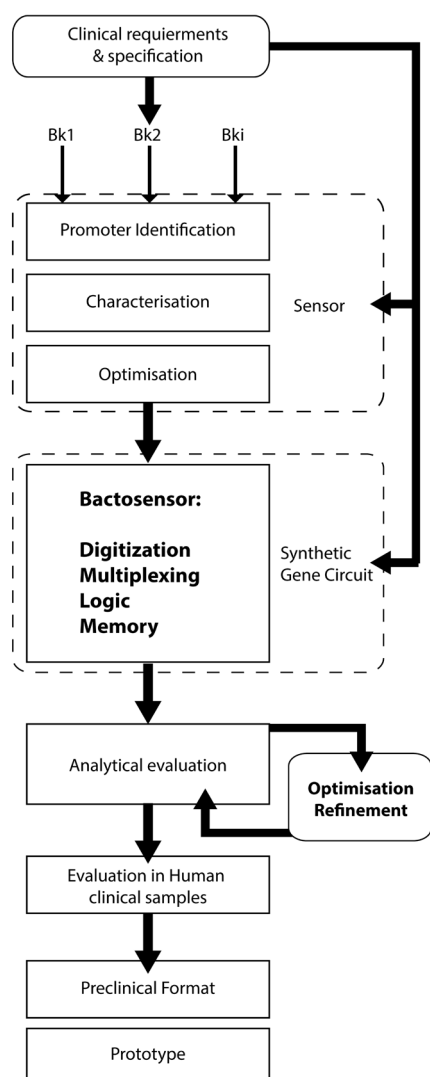


Figure 2.16: Potential modalities of bactosensor based diagnosis and composability of integrase based logic for the development of decision-making tests. (A) Performing a diagnostic test using a medical algorithm based on synthetic cellular consortia where the detection, computation process and outputs are distributed over multiple engineered cells. Different consortia of bactosensors are encapsulated in beads each with a specific sensor module and a transcriptor producing a specific output. One drawback of this approach would be the limitation in number of outputs that can be discriminated. (B) Performing the same test using a medical algorithm based on concatenated Boolean Integrase Logic gates where the detection, computation process and the output generation is confined within one type of engineered cells (unicellular). This approach is limited by the number of available integrases. (C) Multiplexing with different medical algorithms using a hybrid approach: multicellular consortia of concatenated BIL gates. This approach enables the construction of more complex algorithms for different detection purposes. (Bk: biomarker)

As a prototype, our system demonstrates that digital amplifying switches and logic gates can overcome several typical problems faced in the clinical application of traditional (i.e. simple promoter-reporter circuits) cell-based biosensors: low signal-to-noise ratios, partial inhibition of the sensor by the samples, and logic processing of multiple biomarkers signals. Digitalizing along with amplifying and multiplexing signal detection improved sensitivity, mediated sharp response profiles, and offered an all-or-none response based on a pathological biomarker threshold (The strategy we describe here should thus support the engineering of fold change detection⁴⁵³ in a synthetic gene circuit). This strategy also enabled the implementation of user-defined decision algorithms, which can be used to design expert living diagnostics. Moreover, digital switching provided constant outputs and dynamic ranges irrelevant of the control signal and greatly facilitated clinical assay standardization and high-throughput measurements. In addition, transient signals (glucose) that are undetectable in endpoint assays using conventional cell-based biosensors were detected and stored using our system. This long-term data storage property enables diagnostics tests to be performed and results stored for several

months under harsh conditions. Last, we have provided a quantitative framework to evaluate the function and robustness of cell-based biosensors in clinical samples. Furthermore, from the success of this work, I envision that bactosensors could support a systematic programming framework for various biodetection agendas (**Figure 2.17**).



However, several hurdles still need to be overcome to achieve full translation of cell-based biosensors into clinical applications. First, methods to engineer new sensing modules tailored to detect ligands of interest are lacking. Current research efforts focus on mining databases for transcriptional regulators that respond to various biological signals and engineering tailored ligand-responsive RNA switches or transmembrane receptors^{109 201 552 553}. Recent encouraging successes suggest that we will witness significant progress in this field in the near future. Multiple sensors that specifically respond to clinically relevant biomarkers could then be connected to Boolean integrase logic gates to perform multiplexed biomarker detection and analysis in clinical samples. Using logic in this setting could not only provide with delocalized expert decision systems, but also increase the scale of measurement compared to a set of parallel single input bactosensors (**Figure 2.16**). Second, response times of bactosensors can be too long for clinical assays. For example, here we measured signal after 18-hour incubations and were able to detect an interpretable output response after 4 to 5 hours (**Figure 2.13**). Given the current response time of our assay, it is thus unlikely to compete with test strips for the detection of glycosuria. Further work should thus be devoted to obtaining the shorter response times needed for diagnostic tests, for example, by interfacing our sensors with electronic devices or engineering circuits that rely on post-translational signaling (such as protein phosphorylation).

Figure 2.17: Conceptual workflow for the systematic development of medical bactosensors. Clinical requirements and specifications for medical diagnosis consist of a set of selected biomarkers at concentration threshold specific of a pathology. Biomarkers detection requires the identification and characterization of specific sensor modules, in our case specific transcription regulators driving target promoters. Medical knowledge informs the choice of the medical algorithm embedded within a synthetic gene circuit also responsible for signal digitization, multiplexing of inputs, and memory. The bactosensor is then constructed, and analytical performance is evaluated and can be refined and optimized before being tested in a clinical context. The process leads to a preclinical prototype.

Nevertheless, whereas long measurement times are not compatible with timely diagnosis for certain applications (e.g. toxicological emergencies), our system would still be relevant for addressing certain medical questions that require less urgent results such as large-scale population screening, monitoring of chronic disease evolution, or companion diagnostics.

From a broader perspective, our work could be a stepping stone toward future applications that use living cells to perform *in vivo* diagnostics coupled with *in situ* synthesis and delivery of therapeutic molecules. Interestingly, it could offer new avenue in the emerging field of intelligent theranostics and probiotics. For instance, bactosensors could offer novel biosensing platforms to perform monitoring of pathologies or coordinate therapeutic interventions *in situ* in the human gut, in the form of programmable probiotics. While this approach has already been described, a transcriptor based platform would provide with drastically new capabilities compared to previous studies relying on high metabolic cost gene circuits, or on transcriptional or post-translational memory⁴¹.

Last, although this study addresses robustness and standardization issues essential for commercialization approval, regulatory and safety concerns regarding the use of engineered living organisms in the clinics remain, and societal and ethical questions must be addressed before such agents can be effectively used in the clinic⁵⁵⁴ (**Figure 2.17**). All switches, logic gates, and uses demonstrated or disclosed herein have been contributed to the public domain via the BioBrick public agreement (<https://biobricks.org/bpa>).

2.7 Materials and methods

Study design

Our goal in this study was to investigate the use of recently developed digital amplifying genetic switches and logic gates to bring the performance of cell-based biosensor closer to clinical requirements. In particular, we wanted to assess whether digital amplifying switches could overcome typical problems faced in the clinical application of wholecell biosensors, such as low signal-to-noise ratios, partial inhibition of the sensor by the samples, and logic processing of multiple biomarker signals. Using glycosuria as a model system, we aimed at demonstrating the detection of an endogenous clinically relevant biomarker in a clinical setup using samples from diabetic patients. We developed a technological platform that was used to build several biosensors capable of detecting various biomolecules. To this aim, we use synthetic biology principles (including standardization and modularity) and provided a method to couple new detection sources to our system. The gates are fully modular (that is, the logic can be easily altered by changing the target DNA sequence for the recombinases) (See **Figure S2.1**). To evaluate the robustness of our system and its functionality in clinical samples, we used serum and urine pools from healthy individuals as well as urine samples from healthy individuals and diabetic patients. Regarding collection of clinical samples, nonpathological (control) and glycosuric (diabetic) urine samples were obtained from the Department of Endocrinology of the Lapeyronie Hospital, Montpellier, France, under the supervision of E. Renard. Individual informed consents were obtained from the patients and control individuals. Glycosuric urine samples were collected from 10 newly discovered, non-stabilized, were obtained from the Etablissement Français du Sang, Montpellier, France. Serum was heat-inactivated by incubation in a 56°C bath for 30 min. Serum and urine samples were stored at -80°C before use.

Molecular biology

Constructs used in this study were cloned using standard molecular biology procedures or one-step isothermal assembly⁵¹⁸. All enzymes were purchased from New England BioLabs (NEB). PCRs were performed using the Q5 PCR mastermix (NEB, 1-min extension time per kilobase). Most primers were generated using J5 (j5 DNA Assembly Design Automation Software⁵⁵⁵; <http://j5.jbei.org>). Primers were purchased from Eurofins Genomics and IDT (Carlsbad, USA). Detailed information and plasmid maps, primers, and Gblocks sequences can be found in the Supplementary Materials. Boolean integrase logic GFP Boolean integrase logic gates and pBAD/pTET plasmid constructs used in this study come from previous work⁴⁵⁸. XOR, AND, NOR, and NAND gates were then modified by Gibson assembly to replace GFP with mKate2.

Library design, construction, and screening

To obtain functional synthetic networks, we needed to finely tune gates operation so that translation levels of integrases match relevant clinical dynamic ranges for our application. To introduce diversity within integrase expression cassettes, we built combinatorial libraries of pCpxP and pYeaR promoters driving expression of integrase TP901 and BxB1, by using primers (JB587, JB588, with G1005) and randomizing (i) RBS (4096 variants), (ii) RBS and initiation

codon (8192 variants), and (iii) RBS, initiation codon, and SsrA tag (AXX) (1,179,648 variants). Effective randomization at specific positions was achieved using degenerated primers, and amplified fragments were cloned in a medium copy plasmid (J64100, chloramphenicol resistance). The library was then electroporated in DH10B electrocompetent *E. coli* (Life Technologies) and plated on chloramphenicol plates. After overnight growth, ~8000 colonies per library were counted. The libraries were grown overnight at 30°C in 10 ml of LB with chloramphenicol and mini-prepped. The libraries were then transformed into a chemically competent screening strain containing an episomal XOR-BCD-RFP logic gate. To isolate NO_x-responsive switching clones, the pYeaR library cells were plated, and 600 clones were picked and induced overnight in 400 ml of LB with chloramphenicol with 10 mM NO_x. Clones were then screened using a plate reader by measuring RFP fluorescence levels. Different clones switching after induction were kept for further investigation, yielding controller 1 and controller 2. To obtain controller 3, the TP901 fragment library was cloned in pYeaR_J64100, the library was cotransformed with XOR-RFP gate and induced with 10 mM NO_x, and 400 clones were screened using a plate reader. To isolate glucose-responsive switching clones, the pCpxP-BxB1 library was cotransformed with XOR-RFP gate and sorted using a FACSAria (BD Biosciences): on a first sort step, constitutively switching cells were discarded and the remaining clones were kept and then induced in LB medium containing 0.5% (w/v) glucose for 6 hours. After induction, cells were washed and grown in fresh LB medium overnight at 30°C. On a second sort step, switching cells (~1000 clones) were kept, and nonswitching cells were discarded. One pCpxP-BxB1 controller clone was finally kept for use.

Beads assay

To test the operability of bactosensors in PVA/alginate beads, beads were inoculated in 300 ml of culture medium with or without inducer, or urine from patient diluted at a ratio of 1:4 in culture medium for a total volume of 300 ml in a 96-well plate. After 24 hours of incubation, fluorescence was read using a synergy H1 plate reader (more details on encapsulation of bacteria in beads can be found in the Supplementary Materials). We concomitantly tested these urines from nonstabilized diabetic patients using the Siemens Multistix 8 SG reagent strip according to the manufacturer's protocol.

Cell culture and data collection

We used *E. coli* DH5aZ1 and *B. subtilis* 168 1A1 for all measurements. Cells were cultivated with shaking at 400 rpm and grown for 18 hours at 25°, 30°, or 37°C, in either Azure Media (Teknova) or LB phosphate buffer adjusted to pH 7. Antibiotics used were carbenicillin (25 mg/ml), kanamycin (30 mg/ml), and chloramphenicol (25 mg/ml) (Sigma). D-Glucose, nitrate, L-ara, and aTc (Sigma) were used at final concentrations of 0.5% (w/v), 0.1 M, 0.5% (w/v), and 200 ng/ml, respectively. Cells were streaked from a glycerol stock, and then one clone was inoculated in 5 ml of LB with carbenicillin and/or chloramphenicol and grown overnight at 30°C. The cells were then diluted at a ratio of 1:200 and grown for 6 hours at 30°C until an optical density (OD) of ~0.5. The cells were then back-diluted at a ratio of 1:100 into 1 ml of azure medium (Teknova) and diluted with urine or serum, induced with 0.5% (w/v) ara, aTc (200 ng/ml), 0.5% (w/v) glucose, or 10 mM NO₃⁻, and grown for 18 hours at 25°C in 96 DeepWell plates. The next morning, the cells were put on ice, and we measured RFP/GFP fluorescence levels (588ex/633em, 485ex/528em, respectively) and OD₆₀₀ using a synergy H1 plate reader

(BioTek) and a Beckman Coulter FC 500 flow cytometer recording 50,000 events per sample. Events were gated on forward and side scatter to exclude debris, dead cells, and doublets. The overnight growth, back dilution, and measurement procedure were performed three times on separate days in triplicates. Measurements for each data point were normalized using a reference promoter (BBa_J23101) driving expression of sfGFP (low-copy plasmid pSC101 origin with chloramphenicol resistance). For functional and genetic memory experiments, cells were cotransformed with pTET/pBAD dual controller plasmid and AND-BCD logic gate plasmid, and induced overnight at 25°C with 0.5% (w/v) ara and aTc (200 ng/ml), in 300 ml of urine, serum, or Azure medium in p96 plates. The plates were kept for 8 months at 4°C. Plasmid DNA was then recovered by scrapping and dissolving the dry cellular residues of cells in phosphate-buffered saline, and extracted using QIAamp (Qiagen) kit. We used specifically designed primers (attL/attR Bxb1 or TP901) to PCR-amplify the recombined targets. For Sanger sequencing experiments, the gate plasmid DNA was amplified using primers G1004 and G1005, and the PCR product was sent for sequencing.

Data analysis and statistics

Experimental values are reported as means \pm SD. All experiments were performed at least three times on different days and in triplicate. Data, statistics, graphs, and tables were processed and generated using MATLAB (MathWorks), SigmaPlot (Systat Software Inc.), and the R with ggplot2 package. Flow cytometry was performed using an FC 500 (Beckman Coulter Inc.), and data were analyzed using FlowJo and Flowing Software (Turku Centre for Biotechnology). We used RPUs to integrate into clinical measurements an in vivo internal standard for bactosensor operation and signal generation (45). For signal amplification experiments, amplification was quantified by the gain defined as the $10\log$ ratio between the fractional change in the output signal GFP and the fractional change in the input signal RFP. For receiver operating characteristic analysis, a set of 27 measurements performed in nonpathological urine were compared to 27 measurements performed in urine containing 1% glucose. See the Supplementary Materials for details on calculations.

2.8. Supplementary Materials

2.8.1. Molecular biology

pYeaR/pCpxP measurement plasmids

pYeaR-GFP and pCpxP-GFP measurement plasmids were built from Gblocks (containing pYeaR/pCpxP promoters and RiboJ flanked with BioBrick suffix and prefix) purchased from IDT ((Carlsbad, USA), and PCR amplified using primers G1004 and G1005. These fragments were then cloned in j64100 plasmids (bearing chloramphenicol resistance and ColE1 origin of replication) using BioBrick assembly standards ([BBF RFC 10](#)). Then, these constructs were PCR amplified using primers AC_00A/AC_00B and ligated by Gibson Assembly with previously amplified GFP fragment using primers AC_00C/AC_00D, to generate pYeaR-GFP and pCpxP-GFP measurement plasmids.

Boolean Integrase logic Gates plasmids

XOR, AND, NOR and NAND gate⁴⁵⁸ were cloned upstream mKate instead of GFP. The gates were amplified using primers AC_00E/AC_00F, and a mKate2 fragment was amplified with primers AC_00G/AC_00H. The amplified fragments were then ligated by Gibson assembly.

Signal amplification measurement plasmid: pYeaR-BxB1_CT1-RFP

Controller 1 pYeaR-TP901 switching clone was PCR amplified using primers AC_001/AC_002, and a TP901_RFP fragment was amplified using #534_dual controller_RFP as a template⁴⁵⁸. These two fragments were then ligated by Gibson Assembly.

Dual controller plasmid: pYeaR-TP901/pCpxP-BxB1

Controller pCpxP-BxB1 plasmid was PCR amplified using primers AC_005/AC_006, and a pYeaR-TP901 fragment was amplified using AC_007/AC_008 and pYeaR-TP901 controller as a template. These two fragments were then ligated by Gibson Assembly.

Primers used in this study

AC_00A	TGTACAAATGATGATACTAGTAGCGGCCGCTGCAG
AC_00B	CGCATATGTATATCTCCTTCTTAAAAGATCTTTAAACAAAATTATTTGTAGA GGCTGTTTCGTCCTC
AC_00C	TTGTTTAAAGATCTTTTAAAGAAGGAGATATACATATGCGTAAAGGCGAAGA GCTGTTCCTACTGG
AC_00D	CGGCCGCTACTAGTATCATCATTTGTACAGTTCATCCATACCATGCGTG
AC_00E	GGACATCGTTGATAATACTAGAGCCAGGCATCAAATAAAACG
AC_00F	CTTTAATTAATTCTGACATTAGAAAACCTCCTTAGCATGATTAAGATG
AC_00G	TGCTAAGGAGGTTTTCTAATGTCAGAATTAATTAAGAAAATATGCACATG

AC_00H TGCCTGGCTCTAGTATTATCAACGATGTCCTAATTTTCGACGG

AC_001 gtcgaaattaggacatcggtgataaTACTAGTAGCGGCCGCTGCAGtcc

AC_002 aatatatacctcttaatttttactagtaCGACATCCCGGTGTGTAGCC

AC_003 CACCGGGATGTCGtaactagtaaaaattaagaggtatatattaatgtcag

AC_004 cggaCTGCAGCGGCCGCTACTAGTAttatcaacgatgtcctaatttcga

AC_005 GGGAGGATTATAGATGGGAAAGGCAGAAATTACGTCATCAGACG

AC_006 ttcCAACTCGCTACCGGTTAACTCTAGAAGCGGCCGCGAATTC

AC_007 CTGATGACGTAATTTCTGCCTTTCCCATCTATAATCCTCCCTGATTC

AC_008 CTTCTAGAGTTAACCGGTAGCGAGTTGgaattt

AttL BxB1 TCGACGACGGCGGTCTCAGT

AttR BxB1 TCAACCACCGCGGTCTCCGT

AttL TP901 CATCTCAATTAAGGTA ACTA

AttR TP901 CGTTTATTTCAATCAAGGTA

JB_587 GAATTCGCGGCCGCTTCTAGAGAAGCTTgcggtttcacacNNNNNNgctagcRTGAG
AGCCCTGGTAGTCATCCG

JB_588 GAATTCGCGGCCGCTTCTAGAGAAGCTTgcggtttcacacNNNNNNgctagcRTGacta
agaaagtagcaatctatacagagtatcc

G-blocks used in this study

pCpxP GAATTCGCGGCCGCTTCTAGAGaggcagaaattacgtcatcagacgtcctaatacctgactttacgttggttttacccccct
gacgcatgtttgagcctgaatcgtaaactctctatcggtgaatcgcgacagaaagattTGCATGCATGCAAGCTGTCACCG
GATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTT
TGTTTTACTAGTAGCGGCCGCTGCAG

pYeaR GAATTCGCGGCCGCTTCTAGAGTTCCCATCTATAATCCTCCCTGATTCTTCGCTGATA
TGGTGCTAAAAAGTAACCAATAAATGGTATTTAAAAATGCAAATTATCAGGCGTACCC
TGAAACGTGCATGCATGCAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCC
GTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAACCAAGCTTGGGTACTAGT
AGCGGCCGCTGCAG

Maps of plasmids used in this study

The reader is kindly referred to the annexes for graphical maps.

2.8.2. Relative promoter units, signal gain amplification, and thresholding calculation.

We use Relative Promoter Units (RPU) to integrate in clinical measurements an *in vivo* internal standard for bactosensor operation and signal generation⁵¹⁷. The output signal is normalized to a reference signal accounting for the analytical performance of the bactosensor across changing contexts with a reference standard promoter (BBa_J23101). RPU can be calculated as the ratio OD600 corrected fluorescence intensity of the measurement on the OD600 corrected fluorescence intensity of the reference promoter in the same conditions.

$$RPU = \frac{\left(\frac{RFU_{\text{sample}} - RFU_{\text{blank}}}{OD_{\text{sample}} - OD_{\text{blank}}}\right)}{\left(\frac{RFU_{\text{pREF}} - RFU_{\text{blank}}}{OD_{\text{pREF}} - OD_{\text{blank}}}\right)}$$

Gain amplification calculation can be evaluated by changing the input signal intensity $RPURFP$ to $RPURFP + \Delta RPURFP$ and measuring the response $\Delta RPUGFP$ in the output signal $RPUGFP$ from its stationary value $RPUGFP_0$. The amplification can be quantified by the gain defined as the $10\log$ ratio between the fractional change in the output signal GFP and the fractional change in the input signal RFP.

$$\text{Gain (dB)} = 10 \log_{10} \left(\frac{\frac{\Delta RPUGFP}{RPUGFP_0}}{\frac{\Delta RPURFP}{RPURFP_0}} \right)$$

Threshold calculation for Figure 3C was achieved by fitting experimental points with a four parameter logistic equation. Calculated threshold were: $6.291 \cdot 10^{-6} \text{M}$ [$3.29 \cdot 10^{-6}$; $1.203 \cdot 10^{-5}$] $R^2=0.9745$, $3.617 \cdot 10^{-4} \text{M}$ [$3.402 \cdot 10^{-4}$; $3.846 \cdot 10^{-4}$] $R^2=0.9991$, $2.616 \cdot 10^{-3} \text{M}$ [$2.954 \cdot 10^{-5}$; 0.2316] $R^2=0.9975$.

2.8.3. Bacterial growth and viability measurements.

B. subtilis 168 1A1 and *E. coli* DH5aZ1 transformed with reference promoter carrying plasmid (BBa_J23101) controlling expression of sfGFP (low-copy plasmid pSC101 origin with chloramphenicol resistance) were plated from a previous glycerol stock. One clone was inoculated in 5 ml of LB with carbenicillin and grown overnight at 30°C. The cells were then diluted at a ratio of 1:200 and grown for 6 hours at 30°C until reaching an OD of ~0.5. The cells were then back-diluted at a ratio of 1:100 into 1 ml of LB with antibiotics diluted with urine or serum and grown for 18 hours at 25, 30 or 37°C in p96 deepwell plates. The next morning, the cells were put on ice and we measured GFP fluorescence levels and OD600 using a synergy H1 plate reader (Biotek). Then, 100µl of overnight culture were added to 100µl of PI solution (100µg/ml) and incubated for 5 minutes at room temperature prior to flow cytometry analysis. A death positive control was prepared by adding 900µl ice-cold ethanol (80%) to 100µl control bacteria, incubated 60 minutes at 4°C and then resuspended in 1 ml LB. We used a Beckman Coulter FC500 flow cytometer to measure GFP and PI fluorescence recording 50,000 events per samples. Events were gated on forward and side scatter to exclude debris, dead cells, and doublets. The overnight growth, back-dilution, and measurement procedure were performed 3 times on separate days in triplicates.

2.8.4. Kinetics experiments

One clone of a previously streaked glycerol stock was inoculated in 5 ml LB with antibiotics, and grown overnight at 30°C. The cells were then diluted 1:200 and grown 6 hours at 30°C, and diluted again 1:100 in 300 µl Azure medium with antibiotics with or without inducers (Glucose 0,1M, NOx 0,01M) in a p96 microwell plate, and incubated at 30°C in a Synergy H1 plate reader with 250 cpm double orbital agitation. OD600 and GFP fluorescence were read every 15 minutes for 25 hours.

2.8.5. Bacterial cell encapsulation in hydrogel beads.

Sodium alginate, polyvinyl-alcohol, boric acid, sodium sulfate and calcium chloride were purchased from Sigma-Aldrich, USA. An overnight culture of bactosensors was diluted at a ratio of 1:200 in 20 ml LB and grown for 6 hours at 30°C until an OD of ~0.5. The culture was then centrifuged (3000g, 5min) and the pellet diluted in 25 ml of a mixed solution of PVA (10% w/v) and sodium alginate (1% w/v) in Azure medium with antibiotics and agitated 30 min at 30°C at 400 rpm. Spherical PVA/alginate beads are produced by crosslinking with boric acid solution (for a short time to prevent damage to the immobilized bactosensor) and calcium chloride, followed by esterification of the PVA with sulfate for further strengthening. To change the poor gas permeability of PVA gels, sodium alginate is added to the PVA solution and the saturated boric solution includes calcium chloride for further crosslinking of the alginate. Later, this is removed by using the sulfate solution, to obtain a more porous structure allowing mass transfer of solutes inside the bead. The mixture was then loaded into a 20 ml glass syringe controlled by neMESYS syringe pumps (Cetoni, Germany), and dropped to form beads at a flow rate of 40µl/s into 200 mL of boric acid (0.8M) and calcium chloride (0.2M) solution using a syringe needle (21 G) and

stirred for 30min. The beads were then removed and stirred in 200ml of sodium sulfate solution (0.5M) for 90 min. Beads were then washed in PBS and stored at 4°C in PBS before use. The average diameter of the beads obtained was ~2 μm. Fluorescent imaging of induced /non-induced beads was performed using a Leica DMIL inverted microscope with a 40x lens with standard three filter cube. Beads appeared optically clear. We used a scanning electron microscope (HITACHI S4000 with 8nm resolution, Hitachi High-Technologies, Tokyo, Japan) at 10kV to visualize the external and internal structure of the beads, previously fixed with ethanol. Photographs of beads were taken using a 5D MarkII camera with a Sigma 24mm/f1.8 Macro Lens.

2.8.6. Population and single-cell measurements of multiplexing Boolean Integrase Logic gates operation in urine and serum

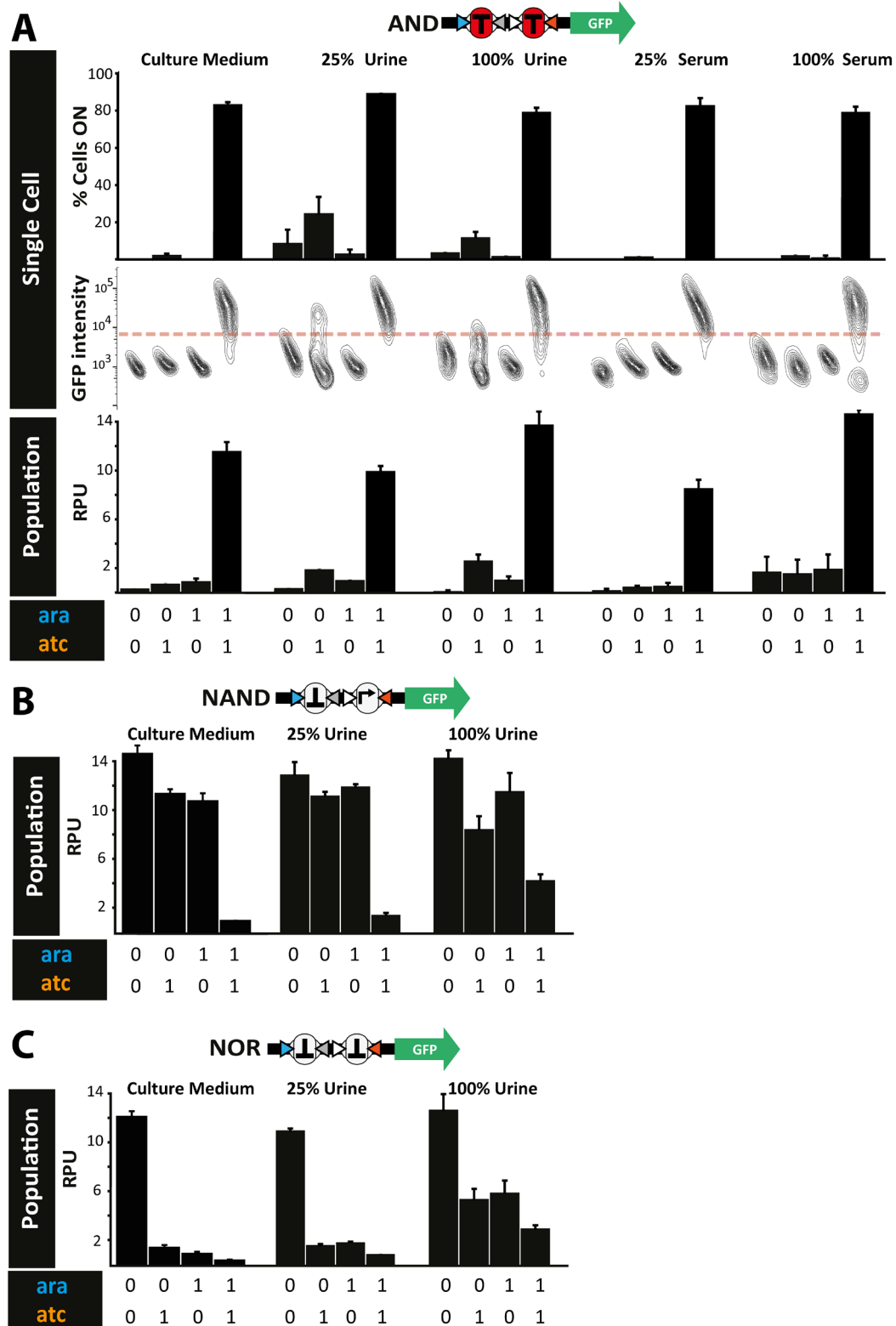


Figure S2.1: Population and single-cell measurements of multiplexing Boolean Integrase Logic gates operation in urine and serum. Operation of AND-BCD-GFP (A), NAND-BCD-GFP (B) and NOR-BCD-GFP (C) logic gates as controlled by the original dual controller plasmid at 25°C, at different dilution (0, 25, 100%) of human urine and serum in azure culture medium. We measured GFP fluorescence after overnight induction performed with arabinose (0.5% w/v, Bk1) and anhydrotetracycline (200ng/ml, Bk2). BCD: BiCistrionic Design.

2.8.7. Stability of genetic memory in clinical samples

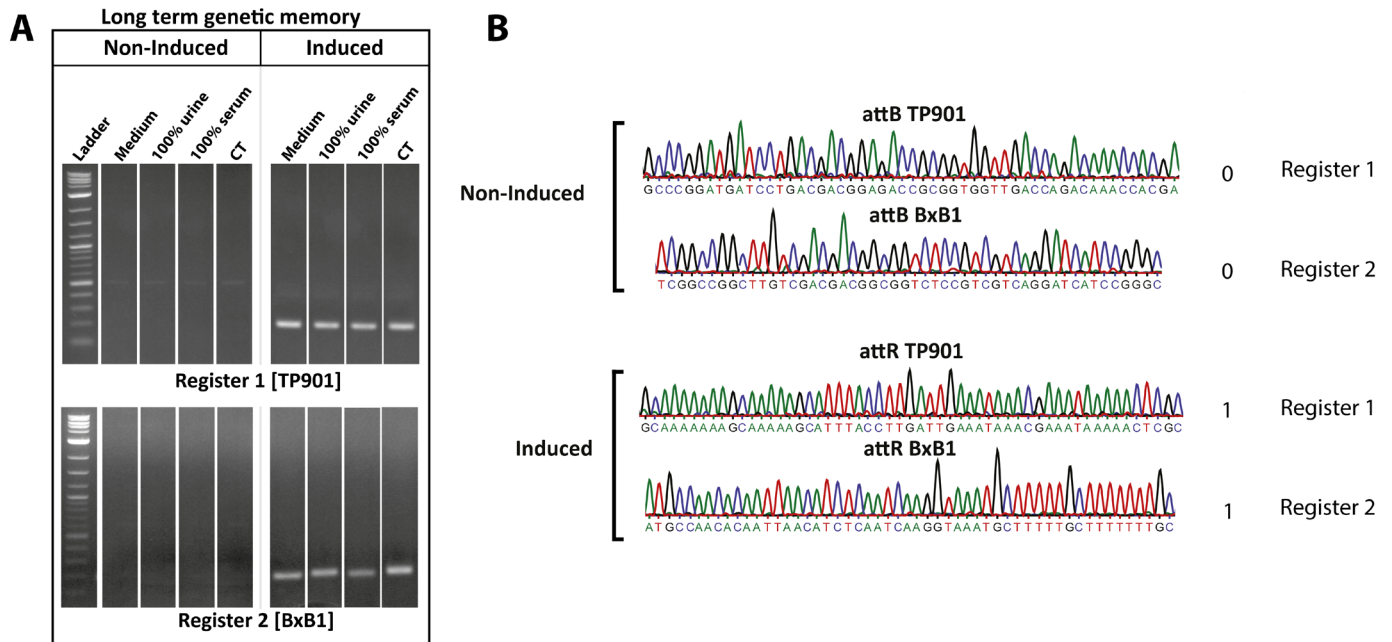


Figure S2.2: Stability of genetic memory in clinical samples (A) Long-term data storage within DNA and recovery using Polymerase Chain Reaction. Cells from were allowed to dry at 4 C° and kept for 8 months, after what they were analyzed by PCR with specific primers **(B)** Addressing the DNA register with Sanger sequencing. Dual controller and AND-BCD-GFP gate were co-transformed and induced with 0.5% w/v arabinose and 200ng/ml anhydrotetracycline and cultivated overnight in urine. Cells were then kept at 4°C. After 8 months, DNA was extracted, PCR amplified and sequenced. The outcome of the test is stored digitally in two distinct DNA registers. The non-induced control, compared to induced cells, doesn't show recombination of integrase target sites attB in attR.

2.8.8. Detailed single cell measurements and analysis.

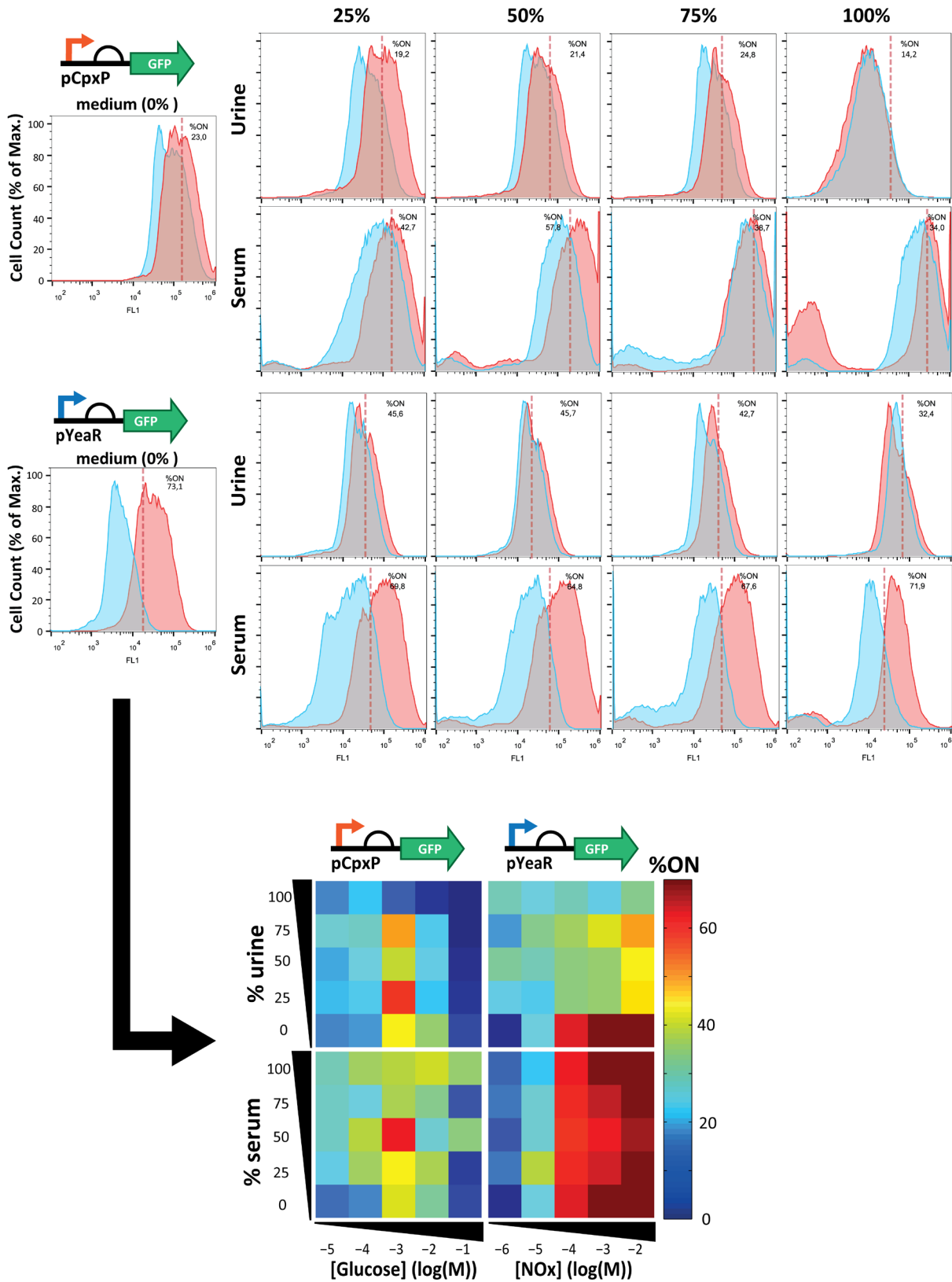


Figure S2.3: Single cell measurement of promoters transfer functions in clinical samples.

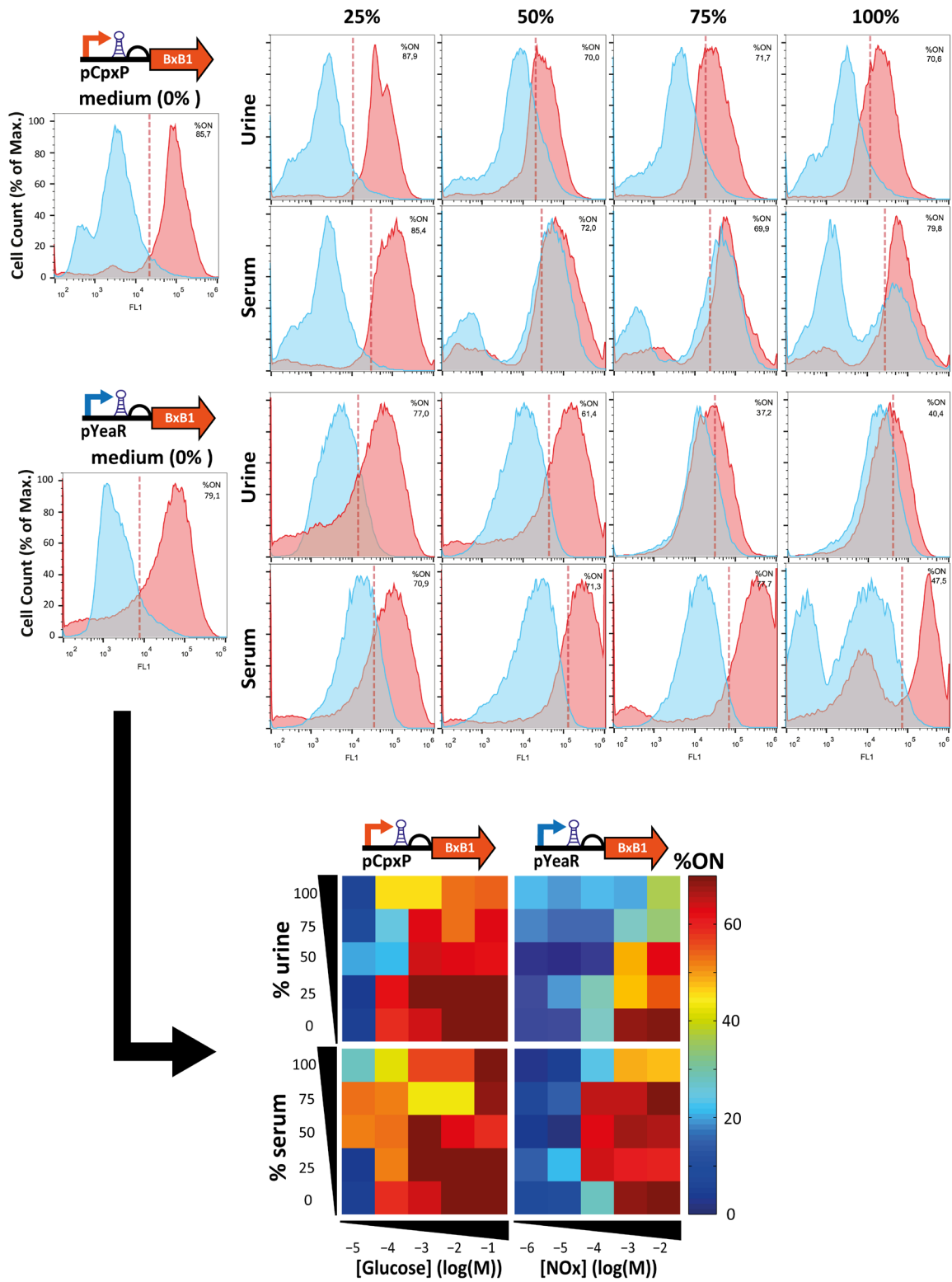


Figure S2.4: Single cell measurement of switches transfer functions in clinical samples.

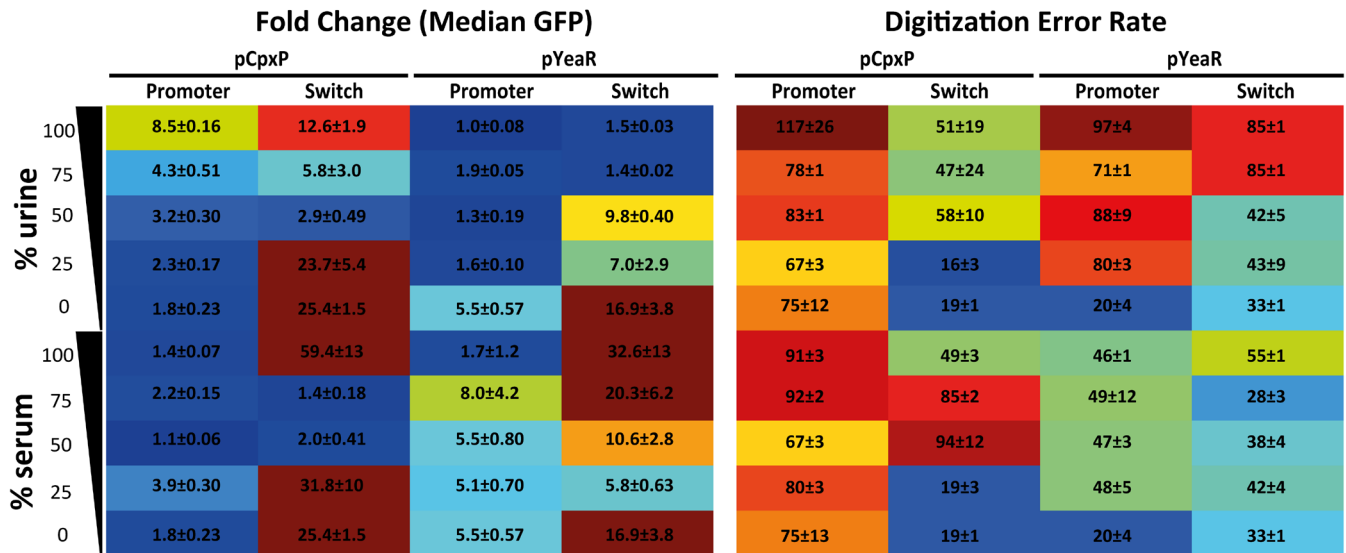


Figure S2.5: Single cell measurements of Fold Change and Digitization Error Rate for promoters and switches. After measuring transfer functions for pYeaR and pCpxP promoters and switches at the population level using a plate reader, cells were analysed at the single cell level by flow cytometry. For each dilution of urine or serum, we applied a threshold maximizing the discrimination between induced and non-induced states, in terms of percentage of cells being GFP-positive (% ON). This enables us to plot the heat maps shown below the histograms. These flow cytometry data were used to determine the Digital Error Rate (defined as the sum of the number of cells in a false positive state and false negative state) as well as the fold change of median GFP fluorescence for switches and promoters. Note the augmentation of discrimination power in terms of fold change and error rate enabled by amplifying genetic switches.

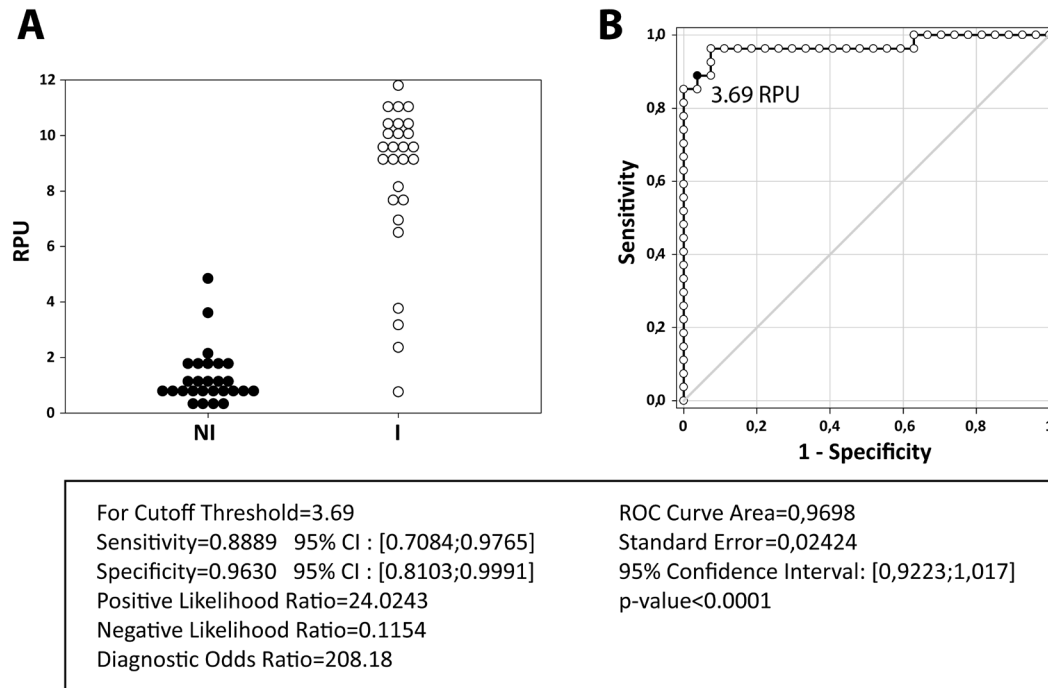


Figure S2.6: Receiver Operating Characteristic (ROC) analysis for Bactosensor mediated detection of glucose in urine. (A) Dot histogram of data used for ROC analysis. (B) ROC curve depicting sensitivity versus (1-specificity). A set of 27 measurement performed in non-pathological urine were compared to 27 measurements performed in urine containing 1% w/v glucose. Beads used for this assay consisted of bactosensors co-transformed with pCpxP-BxB1 controller and XOR-BCD-GFP. Cells were encapsulated as described. Beads were then incubated for 24 hours in urines. GFP fluorescence was read the next day using a plate read and normalized with beads entrapping cells bearing the reference promoter. (B) Using a cutoff threshold of 3.69 RPU, the assay can be interpreted as an excellent test to discriminate between glycosuric and non-glycosuric urines.

Οὐδὲν γὰρ ἄρρημα γινεται οὐδέ ἀπόλλυται ἀλλ' ἀπὸ ἐόντων χρημάτων συμμίσγεται τε καὶ διακρίνεται.

For nothing comes into being nor perishes, but is rather compounded or dissolved from things that are.

Anaxagoras, in *Simplicius' Commentary on Aristotle's Physics* ~500 BC

Chapter 3

Programming autonomous protocell biosensors via integrated synthetic biochemical circuits

Some of the section below was adapted from a publication in preparation and a patent under application process, by Courbet and Molina.

Abstract

Biological systems have evolved powerful modules to sense molecular signals in their environment and inform their phenotypes accordingly. The field of synthetic biology has extensively exploited these capabilities through the genetic engineering of living organisms in order to process information and energy in useful ways. However, *de novo* design of high order functional assemblies of biological components such as synthetic protocells, can achieve comparable signal processing tasks, and thus interface biological information at the microscale. Here, we propose to rationally program protocellular biosensor systems we termed *protosensors*, to perform biodetection of clinically relevant biological signals. We develop an *in silico* and *in vitro* framework to systematically design and engineer robust protosensors that can be programmed to solve different problems. I report that protocells embedding synthetic biochemical reaction circuits are capable of digitization of space and molecular signals through membrane compartmentalization and enzymatic Boolean logic gates respectively. I engineer diagnostic protosensors integrating algorithms informed by medical knowledge to classify pathological states in clinical samples according to the presence of specific patterns of biomarkers. I then demonstrate that protosensors can be used to detect pathological glycosuria in the urines of diabetic patients. These next-generation biosensors with improved biocomputing capacity could enable novel approaches to medical diagnosis, and pave the way for a new generation of autonomous micromachines for the precise interfacing of complex biological systems, such as the human body or ecosystem environments.

3.1 Introduction

Biological systems, ranging from molecular machines to high-order organism level, have evolved to sense, solve information processing problems, and act on their environment in the most optimized way. Translating these biological organizational principles into synthetic systems combining on-purpose biosensing and biocomputing, is of intense interest to advance both basic and applied *in vitro* synthetic biology and would provide with unprecedented tools for solving clinical problems⁵⁵⁶. Synthetic biology attempts to organize biological matter to process information and energy according to user-defined, useful specifications⁵⁵⁷. However, in the process of unveiling and exploiting organizational principles of biological structures, we are still much better at repurposing natural systems than rationally building *de novo*. In this perspective, top-down synthetic biology has focused on the systematic engineering of synthetic gene networks from standardized and composable genetic parts that are then assembled in living organisms⁵⁵⁸. This approach has provided many useful devices and proven very valuable in the biomedical field, for example for drug production⁵⁵⁹, to develop smart therapeutics⁵⁶⁰ or innovative bioanalytic devices⁹⁶. Considering the later, we previously discussed how cell-based synthetic biosensors can be engineered and employed as versatile, scalable and programmable tools to detect molecular signals for environmental or diagnostic purposes. Even though they proved as valuable devices capable of integrating medical expertise and providing smart analytical solutions, they still rely on the use of genetically modified living organisms, which poses ethical, evolutionary, ecological, and industrial challenges⁵⁵⁴.

While I previously engineered bacterial biosensors with improved biosensing capabilities using integrase based digital amplifying gene circuits and logic gates, I believe that the bottom-up design framework as permitted by recent advances of synthetic biology is another key approach to the development of synthetic biological systems with useful properties. Recently, the bottom-up engineering of cell-free molecular biological circuits has witnessed a dramatic growth in achievable complexity, modularity, and programmability⁵⁶¹. Numerous studies have exploited nucleic acids or proteins to design molecular logic gates and biocomputing system either *in vitro* or *in vivo*³⁹, and it has been reported that *de novo* designed synthetic peptide networks can mimic some of the basic Boolean logic functions of biological networks^{342 367}. Although biochemical information processing has been extensively explored, advancing understanding about how to engineer *in vitro* robustness and predictability has remained a critical challenge. Biochemical circuits comprise intrinsic complexity and have thus been used less frequently than gene circuits to build synthetic systems. To date, no clear engineering principles and methodologies exist to design and build cell-free synthetic systems according to specifications.

Living systems share the cellular organization as a common unit structure, which is analog to a complex bioreactor enclosed by phospholipid bilayer membranes. This structural anisotropy through compartmentalization of biochemical reactions is required to achieve complex behaviors and perform catalytic and information processing operation in order to exploit a specific medium. Alternatively, phospholipid vesicles have been widely used to serve as synthetic compartments, and their properties can be exploited to build robust and stable cells mimics, or protocells⁵⁶². In addition, biochemical reactions constituted of enzymes and metabolites have

been fruitfully used to build circuits that mimic Boolean logic gates and achieve biosensing and complex signal processing⁴⁴⁹. Moreover, *de novo* synthetic metabolic pathways can be reassembled *ex vivo* from robust and well characterized components that can be isolated from naturally occurring biomolecular systems, reverse engineered or build with complete orthogonality. Further, integration of biological parts and modules can be accelerated through high-throughput experimentation and computer assisted design^{563 564}. These circuits can be encapsulated within synthetic cell mimics to achieve space digitization and achieve efficient and multiplexed treatment of biological information. Here, we propose that protocells can be programmed to perform biodetection of disease associated biomarkers and biocomputing operations, and can be systematically generated with a robust framework to provide analytical solutions to specific clinical questions. The methodology relies on *in silico* design and accurate system prediction, as well as experimental production using robust microfluidic assembly methods. Here I demonstrate the feasibility of the approach by implementing a full diagnostic algorithm that discriminates between all acute metabolic complications of diabetes and achieves differential diagnosis. I provide experimental evidence demonstrating the technological validity, and the advantages and efficiency in clinical samples of this novel diagnostic approach for the diagnosis of human pathologies.

3.1.1 Protocells and bottom-up engineering of biological systems

Since Oparin's or Schrödinger's precious consideration on the nature of the living^{565 566}, a vast scientific landscape of ever more detailed description of biological systems has emerged. If our exploration scenarios on the origins and nature of life have greatly improved, the formal understanding of nature, however, still remains incapable of facilitating a way to *ab initio* construction of life⁵⁶⁷. In fact, we are still much better at repurposing natural systems than rationally building *de novo*⁵⁶⁸.

Later coined under the term *protocell*, the challenging approach consisting of fabricating artificial cells in the lab, has only been accepted as experimentally feasible in the past decade. Since the first extravagant description of protocellular systems⁵⁶⁹, various approaches have investigated the construction of synthetic cells in the lab, and although significant efforts have been made towards engineering information, metabolism and self-organization within synthetic vesicular cell mimics, it has failed to sustain autopoietic systems undergoing Darwinian evolution, which is a major hallmarks of life. Apart from achieving autopoiesis, which stands as a goal in itself, the synthesis of cell-like systems that could display context adaptation and computation capabilities is of tremendous interest from both fundamental and applied sciences.

Cells are complex biochemical systems that constitute the building blocks of the living. They have evolved as autonomous, intelligent, adaptable and modular microscale machines, capable of managing energy and information to interface their environment and inform their phenotypes and genotypes accordingly. The underlying biochemical networks support their robust

functioning, through coordinated and cooperative dynamical molecular signal processing relying on a vast array of chemical and biological species, protein, metabolites, nucleic acids, which exploit chemical potential energy to function. Of particular interest, the cell membrane is a ubiquitous structure, which constitutes a thermodynamic boundary and supports the digitization of space and biological information/matter to achieve coherent decoupling of biochemical processes. Therefore, this structure requires intense scrutiny to build synthetic protocellular systems. In fact, one could define protocellular systems according to their vesicular nature and their biological information-bearing content. Amongst most described membrane material, authors have mostly used lipids^{386 398 570 571} (i.e. the natural physical boundary of living cells, which remains the most widely used in protocell research), proteins⁵⁷², as well as other inorganic base materials such as block copolymers⁴⁰⁹, polyoxometalates⁵⁷³, or silica⁵⁷⁴ with the intention of building protocellular systems.

Indeed, protocells constitute a research field in itself, at the crossroads of *in vitro* synthetic biology, systems chemistry, systems biology, and bioengineering. The bottom-up synthesis of protocells represents a challenging but described as a reachable goal for synthetic biology, which is increasingly studied. Motivations to solving this complex problem have strongly simulated the advances of *in vitro* synthetic biology⁵⁷⁵, favoring a new technological era for the bottom-up manipulation of biological structures, since it has been proposed that a robust framework to achieve this goal relies on modular design from standard biological parts^{576 577}. In addition, applying a bottom-up approach to build protocells module by module represents a novel strategy to explore the complexity of biological signaling networks.

Having received less attention than *in vivo* synthetic biology, the greatest challenge is probably to explore the vast design space and specific configurations within which biological systems demonstrate complex behavior and emergent properties. However, novel methodologies, supported by computational design and high throughput and quantitative technologies, could be exploited to accelerate the cell-free synthesis of complex protocellular biochemical systems⁵⁶².

Although engineering self-replicative capabilities is still far, the perspective of building *ab initio* biological structures capable of displaying self-organizing behaviors such as signal processing, programmed decision making and bioactuation, is of considerable interest, notably for biomedicine where they could be used as microscale machines to perform diagnostics and therapeutic operations (**Figure 3.1B**).

Considering the application of such systems, *de novo* construction of synthetic cell-mimics offer a vast engineering playground compared to re-engineered cells, which comprise unnecessary biological material imposing energetic burdens. Although the natural cellular chassis may provide certain robustness in specific environment, predicting the behavior of cells has limits that hamper practical application. Instead of using natural cells as the hardware to implement synthetic circuits, protocells can be built with a minimal and well characterized set of biochemical parts without putting a metabolic charge on the system, which makes them particularly amenable to predictive mathematical modeling, and thus ease the bottom-up construction which can be computer assisted. Moreover, they offer a better possibility of control on systems parameters, they minimize unintended cross-talks, and support easy quantitative analysis. Protocells and cell-

free systems may thus constitute the future of applied synthetic biology, as they recapitulate most advantages of the living without the inconvenient.

3.1.2 Implementing biosensing and biocomputing operations using synthetic biochemical circuits

After extensive studies of inorganic chemical reaction networks, *in vitro* synthetic biologists have now started to study how complex spatiotemporal behaviors can emerge from systems of simple biochemical reactions, from abiotic chemistry to life-mimicking functionalities. Since the first description of chemistry based reactive systems exhibiting complex behaviors, such as the Belousov-Zhabotinsky inorganic oscillator, pattern formation systems⁵⁷⁸, and the concept of the first molecular computers⁵⁷⁹, signal processing/computation with molecules at the nanoscale has evolved towards the use of biochemistry, which is probably the most efficient substrate for biocomputation.

Living organisms have managed to adapt chemistry to their interest to achieve autonomy and decision making. As we have previously discussed, biological control mechanisms rely on sensors that integrate molecular information, and processors that treat the information to compute an appropriate response. Cells contain a wealth of inner biochemistry that provides a vast engineering playground to derive a vast range of synthetic circuits. The fundamental question for a synthetic biologist is then: how can one derive cell's complex chemical reaction networks into modular circuits for biosensing and biocomputing?

The first studies were interested by the *in vitro* reconstitution of natural biochemical circuits. For instance, the simple bacterial phosphorylative oscillator constituted by three proteins KaiA, KaiB and KaiC, was one of the first to be studied *in vitro*⁵⁸⁰. It was then showed that naturally occurring cell networks could be decomposed and reconstructed into simple regulatory motifs. Biochemical reaction mechanisms could be reengineered to carry out specific computational functions, such as logic gates and digital/analog circuits^{275 343 474}, switches and memories⁵⁸¹, neural networks²³, universal Turing machines³⁴⁵, or noise filters⁵⁸². The biochemical manipulation of molecular entities have also demonstrated other amazing proofs of concept, such as DNA biocomputers solving hard computational problems^{583 584 585}.

Nucleic acids have been widely used as standard parts for biochemical computing²⁵². As I have previously discussed in Chapter I, they offer the advantage of being deeply modular substrates relying on precisely programmable sequence information. Nevertheless, protein and enzymes are the most ubiquitous and versatile computational elements used in cells⁵⁸⁶, and have the advantage to offer faster kinetics. Moreover, enzymatic circuits can be engineered through the straightforward assembly of synthesized post-translational biochemical components.

Engineering programmable signal processing within biochemical reaction circuits requires a theory and motifs describing information-processing in biological macromolecules. Electrical and biochemical circuits bear some intrinsic resemblance: generators, transistors, resistors, capacitors

and wiring through selective molecular binding interactions. Therefore, synthetic biologists have intended to accommodate the traditional concepts familiar to electric engineering: modularity and standardization to adapt logical operator structures in order to achieve programmed biological information processing. In either equilibrium or non-equilibrium systems, similarly to electrical circuits energy consumption is a key property of biochemical information processing circuits determining specificity, dynamics, variability, signal amplification, and memory parameters⁵⁸⁷. Exploitable biochemical energy mostly relies on redox electron transfer gated by enzymes, which act as elementary molecular components (or biological transistors). Electrochemical potential in biochemical reactions and current flow in electronics indeed display interesting thermodynamic similarities if one compares subthreshold regime of transistor operation to an enzymatic reaction (**Figure 3.1A**). Following the analogy, a signal at some node of a circuit, voltage (or current) in electronic, would take the form of a presence of multiple copies of a particular species at one point of a biochemical circuits. Different biochemical reactions with overlapping substrate and product and operated by different enzymes can then be concatenated to achieve a circuit were a signal can be transmitted from input to output molecules. As a consequence, biological systems show interesting intrinsic properties such as ultra-low power signal processing, compared to equivalent silicon based devices, with 20 kT per biomolecular operation (i.e. corresponding to the hydrolyze of one ATP molecule), or approximately 0.8 pW of power consumption (10 ATP s⁻¹)⁴⁵¹.

The synthetic biology framework consisting in bottom-up construction of biochemical systems from the assembly of standard parts is analogous to designing electronic circuits using standard modules. One way of achieving robust biochemical computation is to use high concentration of species, to stay far from stochastic phenomenon that appear in networks, and sources of noise that could be deleterious if not controlled properly. Digital signal processing in the form of Boolean algebra thus appears has a good choice to implement decision making algorithmic in biochemical circuits. This is supported by the fact that a framework to build such circuits already exists for electronics. A circuit processing biochemical information under the digital domain, requires the definition of *YES* and *NO* thresholds for concentration parameters, for instance for inputs and outputs. *NO* values often identify with absence of input, and *YES* values to a continuous valuation of presence of the output. Concentration of circuits components depend on kinetic rates and one can optimize them to obtain digital processing of information. Nonetheless, other modes of operation have been described, somehow closer to natural signal processing, such as analog computation which is a wider form of the digital subset. In addition, concatenation of biochemical reactions can perform reliable hybrid analog–digital computation. Therefore, a circuit of macromolecules and the computation tasks it performs is encoded in topology and in the kinetic parameters ruling the interactions between components. Programming a biochemical circuit to display a user-defined behavior thus requires a complete specification of (i) the dynamic behavior itself (ii) topology and molecular architecture of the circuit (iii) a thermodynamic and kinetic description of the reactions. For instance, oscillatory networks can be designed *de novo* from knowledge on topological and kinetic parameters only⁵⁶¹.

One can design biochemical circuits to integrate a catalytic activity or allosteric regulation that depends on or necessitate a specific molecular input biomarker. The possibility of designing circuits responsive to molecular triggers has thus prompted the coupling of biosensing with biocomputing. Synthetic biochemical circuits offer an unprecedented versatility in analytical and diagnostic applications since biochemical circuit-based solutions for information processing tasks can be derived in a systematic fashion (i.e. they are fully programmable).

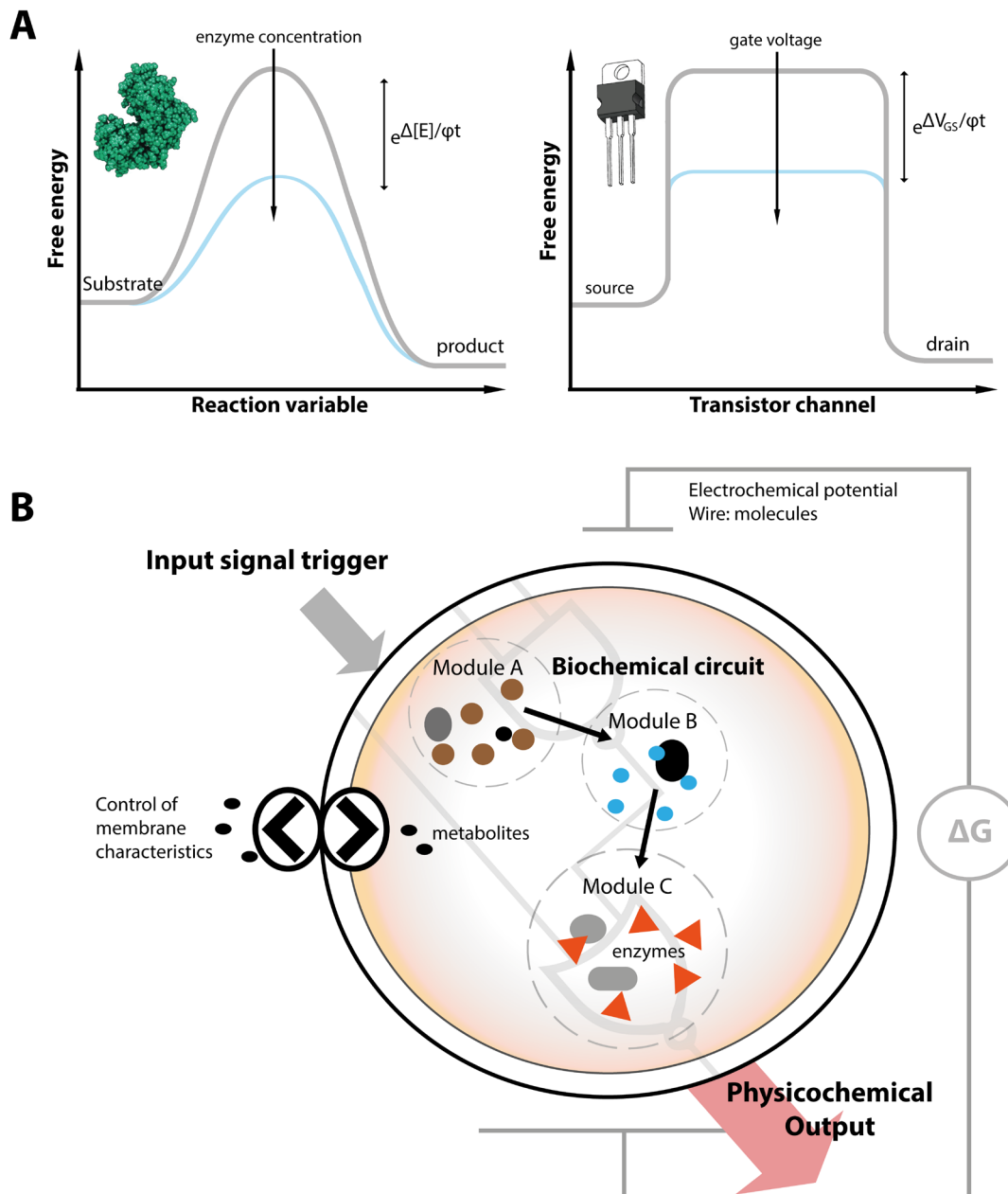


Figure 3.1: (A) Thermodynamic consideration and relation between electron and enzyme based devices to build control circuits from the bottom-up. Substrate and products and electron concentration, current flows and reaction rates, are analogous. Currents in a transistor, and reaction rates in a biochemical reaction, are exponential in voltage differences and Gibbs free energy differences (Adapted from Sarpeshkar⁴⁷⁴). (B) Protocell embedding a synthetic biochemical circuit comprising different enzymatic modules to perform biosensing and biocomputing operation. A protocell can be defined by its vesicular nature associated to biological information-bearing content.

Thus, the final ambition will be to develop a systematic framework to assemble abstract biochemical modules in large-scale circuits, interfacing the molecular environment and recapitulating any arbitrary logic operation with ever increasing precision and capabilities (until potentially achieving life-like behavior³³). To this end, mathematical models and computer simulations are required, and the constant evolution of this field has recently proven increasingly powerful.

3.1.3 Computer assisted design and modeling for bottom-up synthetic biology

Converting design concepts to predicted results is a challenging task when facing the overwhelming complexity of biology. Therefore, the scaling up into increasingly complex biochemical circuits requires automated tools to assist design. It is interesting to point out that electronic design automation for instance, sustained the rapid and geometrical growth in size and capacity of electronic devices (i.e. Moore's law)⁵⁸⁸. Following the analogy, the *de novo* construction of biological circuits and systems according to specifications, could be greatly supported and accelerated by computer assisted modeling, simulation, design, model checking, sensitivity and robustness analysis⁵⁸⁹.

During the last decade, systems biology has given major support to explore design principles linking biochemical circuits' topology to biological processes of interest, since it focuses on the development of computer tools for modeling, simulation to understand complex biological systems⁵⁸⁸. The main reason that prevented synthetic biologists to identify robust configurations to implement within a configuration/parameter space in a biological systems to design, is the number of configurations that grows exponentially with the size of the system that need to be experimentally sampled. *In silico* simulation and model prediction assist in the design process and decouple the pace at which synthetic biochemical circuits can be constructed to solve specific biosensing and biocomputing tasks.

The engineering of programmable circuits requires manipulable abstract entities. Composability in design is crucial to enable the construction of complex systems from the assembly of standardized biochemical parts. To this end, parts, modules, devices, systems are organized in hierarchies⁷, to enable rational concatenation at all levels, as well as the distribution of tasks to carry out in the design process. Computer assisted design tools optimize the efficiency and easiness of iterative steps, especially when supported by graphical user interfaces (**Figure 3.2A**). Crucial to this framework is the identification of standardized basic components that are fully characterized regarding their biophysical, thermodynamic and kinetic parameters. They can further be classified in easy accessible databases to speed up the design process and ease the model calibration. This approach is of major importance in synthetic biology, and proved outstanding capabilities for *in vivo* synthetic gene networks (see the BioBrick registry of genetic parts⁹). Likewise, the laboratory developed the hierarchical biochemical database CompubioticDB³⁶⁷ that stores robust biochemical parts and devices, which we are constantly refining and augmenting.

The design process properly begins with the formalization of systems specifications in mathematical terms, which is followed by an abstract implementation of the circuit using standardized biochemical parts. This design framework uses mathematical description, or models taking into account knowledge on biochemical parameters (either measured or assumed), that can be then used to compute and analyze qualitatively *in silico* the trajectory and behavior at the system-level. The process of making biochemical assumptions that will best account for reality (i.e. model fitting) is crucial since it will impact the reliability and precision of the predictions, while overly complex assumption will be deleterious to computation and optimization speed and efficiency. This step happens before *in vitro* implementation and enables to map the design space, and explore design alternatives.

Design of a synthetic circuit then undergoes quantitative assessment of parameter and compositional space (i.e. initial concentrations, rate constants of enzymes...) to verify performance specifications. Lack of exhaustive characterizations in context for some biological parts often imposes successive phases of refinement and experimental agreement (goodness of fit with data) and theoretical iterations (**Figure 3.2B**). Correct parameters of the model are a crucial to obtain a predictive solution, but iterative optimization can facilitate this process.

A comprehensive mapping of the input-output transfer functions and its sensitivity to the circuits composition is of utmost importance to reduce modes of failure, and for the programming of biochemical circuits ensuring robustness (i.e. defined as the capacity for sustained and precise function even in the presence of structural or environmental disruption⁵⁹⁰). Model analysis often comprises sensitivity and robustness analysis. Sensitivity analysis is used to identify parameters source influencing the performance of the model when perturbed. It enables the estimation of the variation of performance under an admissible parameter variation. This analysis is very informative, particularly in cases where circuit model parameters are uncertain or assumptions are made with large variability. The capacity of different designs to undergo and sustain structural or dynamic perturbations without affecting their transfer function can be assessed through robustness analysis. This analysis can provide precious information for the design process.

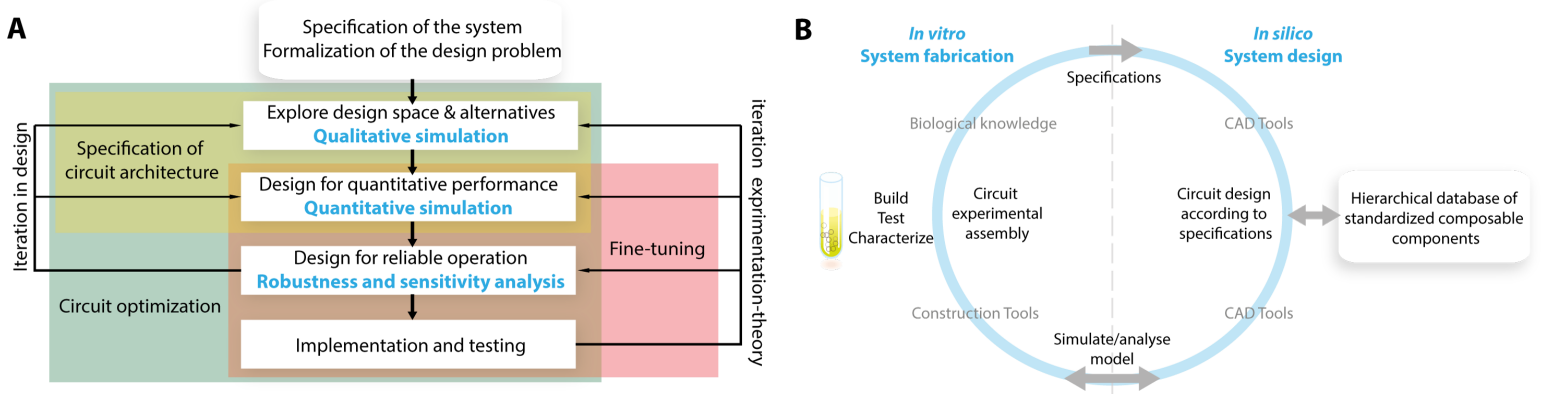


Figure 3.2: Computer-assisted design of synthetic systems. **(A)** Bottom-up engineering of biochemical circuits proceeds from formal specification of the systems behavior to *in silico* implementation and *in vitro* testing, with an iterative approach to design. (Adapted from Marchisio and Stelling⁵⁹¹). **(B)** Iterative circle of system design to fabrication.

The design process requires modeling formalisms, which can be defined as the language in which the model is described. It recapitulates the circuit topology and underlaid mechanistic within a graphical description. The formalism's mechanistic comprises the simulation method, which can be Boolean, deterministic, stochastic, agent-based or hybrid methods. This choice will condition the descriptive and predictive power of a specific formalism. Then, a modeling environment implements *in silico* a chosen formalism to simulate a biochemical system. Numerous software packages have been developed for systems and synthetic biology for this purpose, for instance COPASI⁵⁹², CellDesigner⁵⁹³, or SynbioSS⁵⁹⁴, as well as standardized exchange formats for models, such as SBML⁵⁹⁵ (Systems Biology Markup Language).

The different steps of the process can be supported by various computational methodologies and type of modeling frameworks, such as most common deterministic and stochastic modeling. Deterministic model can describe biochemical systems using analytical equations (usually ordinary differential equations, ODEs or partial differential equations, PDEs) comprising numerical parameters describing molecular interaction, which values are certain. Therefore, the output is by definition exactly reproducible for a defined set of parameter values and initial conditions. Assuming spatial homogeneity, systems can be accurately predicted using ODEs. ODEs describe biochemical reaction, or mass balances (concentration) of species according to:

$$\frac{dX}{dt} = F(N, t; \theta)$$

where dX/dt is the rate of change of concentration of species X , X and N are vectors of species concentrations, θ is a vector of parameters, and $F(N, t; \theta)$ is a vector function that relates rates of change to concentrations⁵⁹⁶. Solving a set of equation describing the systems through dynamic simulation will provide with the time-dependent trajectories of the species concentrations in the model (**Figure 3.3**).

Stochastic models use equations and parameters describing random molecular interactions between species and as such can account for fluctuations inherent to biological systems. This can prove valuable if one is interested by the exploration of noise propagation and influence on systems dynamics. A probabilistic equation is used to describe the probabilistic rate laws, such as the Chemical master equation or Stochastic simulation algorithm (SSA) with the famous Gillespie algorithm⁵⁹⁷ of each reaction between a population of interacting species. The assumption is that the biochemical system within a given time frame obeys rules of randomly interacting molecules under Poisson processes with the rate parameter λ proportional to the reaction rate. Thus, compared with deterministic models, SSA simulations manipulate discrete quantities.

Both approaches may present advantages and inconvenient. However, fundamentally different conclusions about the long-term fate of systems can be reached depending on stochastic or deterministic models, as well as modeling continuous or discrete space⁵⁹⁸. ODEs offer an easiness of implementation, optimization and simulation, and are a good choice to perform extensive mathematical analysis on models, while stochastic methods usually require massive computation power. It should be noted however that stochastic models support closer to reality simulation since they take into account discrete molecular entities, compared to ODEs which process continuous values. Moreover, ODEs can show limits in the ability to capture certain dynamics or

even violate physical laws such as diffusion. This difference can lead to drastically different results at low concentration of species. Specialized for different agendas, these two approaches are complementary to study a biochemical circuit. However, they often require distinct input formats, which require conversion of the original models for compatibility.

Below, I present and discuss the two software that we collaboratively developed, refined and used for the purpose of this study: HSIM^{368 599 600} (Hyperstructure Simulator, developed by P. Amar) along with its recently developed graphical interface NetDraw, and Biocham^{601 602} (Biochemical Abstract Machine, developed by F. Fages *et al.*).

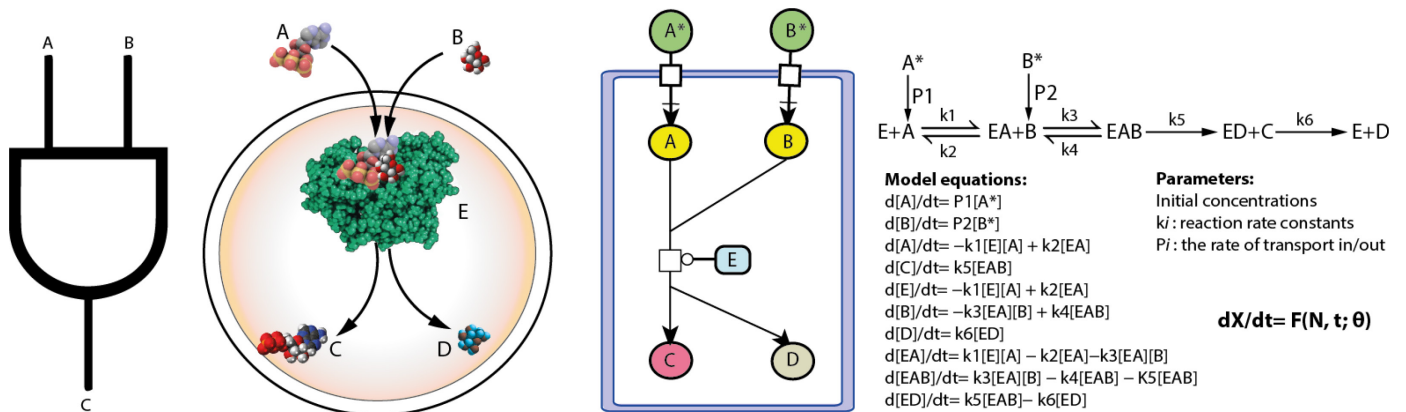


Figure 3.3: Example of a model of an enzymatic circuit mimic an AND Boolean Logic gate inside a protocell, and example of a deterministic ODE model describing its operation.

3.1.3.1 HSIM

HSIM was initially developed to simulate the aggregation and dissociation of large molecular assemblies called Hyperstructures. HSIM is a versatile platform that we used in this study for qualitative and quantitative simulation of biochemical circuits encapsulated in protocells. It exploits two efficient simulation approaches: Stochastic simulation of chemical reactions and entity-centered (i.e. multi-agents) simulation. The stochastic simulation algorithm (SSA) is used with the assumption of spatially homogeneous solutions of reactants, while the entity-centered algorithm enables to take into account spatial heterogeneity in the model when the previous approximation is wrong. Regarding the computational resources, the SSA method shows a limit in the number of reactions it can simulate, whereas entity-centered simulations are limited by the amount of species it can manipulate.

In entity centered simulations, each molecule is considered independently in their environment, where they are allowed to diffuse according to random walk defined by Brownian motion. The spatiotemporal trajectory of the system of interacting entities is then computed according to specific rules (i.e. parameters) and interactions within the volume (i.e. compartments) and other entities (i.e. biochemical reactions: formation or the dissociation of a complex, disappearance, or a change of type of molecules).

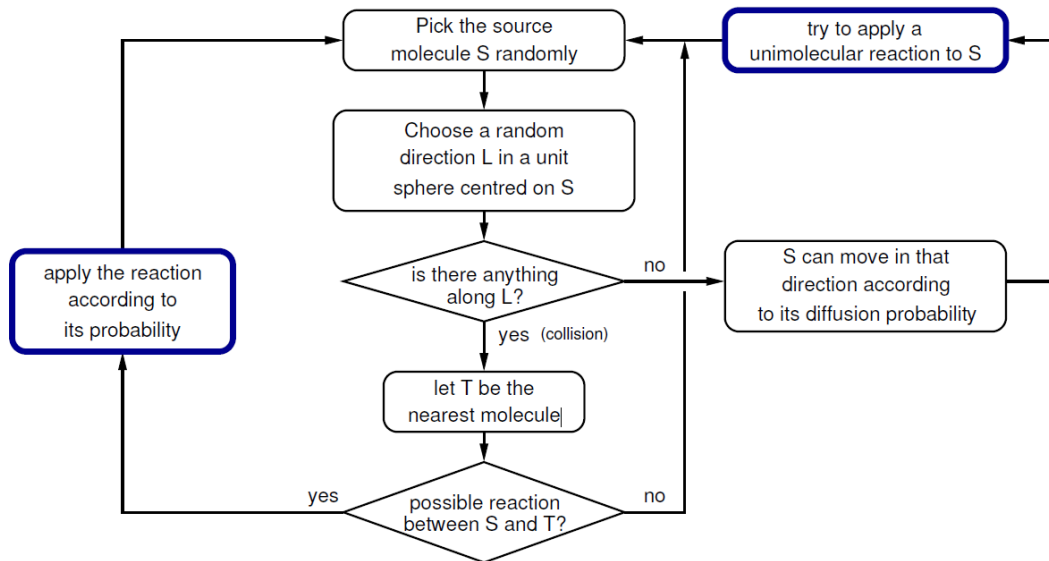


Figure 3.4: Algorithm that HSIM performs at each time step for each molecule defined as entities.

The simulator is a stochastic automaton driven by reaction rules between molecules (**Figure 3.4**). HSIM manipulates biochemical reactions of two categories: unimolecular and bimolecular, while rare natural reactions of more than two reactants are reduced to a combination of bimolecular reactions. Unimolecular reactions describe molecules which can transition between a numbers of finite states with an associated probability. In the case of bimolecular reactions, two reactive molecules collide according to the mass action law, and a probability resending the reaction kinetics is applied to yield products.

HSIM can manage compartments and subcompartments of continuous space and is optimized to manage sizes ranging from small liposomes to eukaryotic cells (100nm-100 μ m). It allows for specification of compartment size, geometry, as well as permeability parameters characterizing the diffusion of species across a membrane of interest.

HSIM keeps a real and discrete computer time record of molecules regarding their types, position, size and interactions. At each time step probability rules are applied to every molecule. The time step used (100 μ s) was optimized for a distance of 10 nm according to the average Brownian displacement of molecules observed for real cytosolic crowding.

Apart from macromolecules, most metabolite molecules can be easily processed in a global way using the SSA algorithm (Gillespie stochastic method), since one can consider that the size will not impact Brownian diffusivity. The high copy number approximation for metabolites enables to treat them as statistically homogeneously distributed in the volume. HSIM can manage both entities and metabolites at the same time, by computing the average number of collisions and subsequent reactions at each time steps.

Below is an example of a HSIM input describing a model with a compartment and subcompartment and two types of molecules (metabolite and entity):

```

title = "Compartment test";
geometry = 100:40;
metabolite s, p; // s et p are small molecules
molecule s, p; // s et p are small molecules
display (s, p);
diffusion (s) = 5e-4; // Permeability coefficient through the compartment membrane of 5.10-4 cm/s
diffusion (p) = 2e-4; // Permeability coefficient through the compartment membrane of 2.10-4 cm/s
compartment {
    geometry = 60:30+10+0+0; // length:diameter+x+y+z
    compartment {
        geometry = 40:20+0+0+0; // length:diameter+x+y+z
    }
}
init (1 μM, s); // Initial concentration of molecules
init (500000, p); // Number of molecules

metabolite s; // SSA
E + s -> ES [0.2]; // Km = 1 mM
ES -> E + s [0.05]; // p2/p1 = 0.25*Km.
ES -> E + p [0.0325]; // k3 = 325 /s

molecule p; // entity
E + s -> ES * s [0.2]; // Km = 1 mM
ES * s -> E + s [0.5]; // p2/p1 = 2.5*Km.
ES * s -> E * p [0.325]; // k3 = 325 /s.

```

3.1.3.2 BIOCHAM

BIOCHAM is a software environment developed for the modeling of synthetic biochemical systems. It comprises a rule-based language and a temporal logic based language and supports the simulation and analysis of boolean, kinetic and stochastic models and the formalization of qualitative and quantitative experimental knowledge of biological properties in temporal logic (i.e. Computation Tree logic (CTL), or Linear Time Logic (LTL)). It manages systems of biochemical reactions with molecular concentrations and kinetic descriptions, in the form of the SBML standard⁶⁰³.

BIOCHAM automates the exploration of a parameter space to optimize and infers unknown model parameters for specific behavior formalized in temporal logic to systematically verify, analyze and optimize models using model-checking methodologies.

It can perform robustness and global sensitivity analysis according to an evaluation function using violation degree of temporal logic formulae³⁶⁹. The violation degree represents the distance between the behavior of the perturbed system and a specified behavior formalized by a temporal logic formula. Biocham automatically computes by numerical simulation of deterministic or stochastic models an estimation of the robustness for dynamical properties and a large number of types of perturbations obtained. It introduced the notion of satisfaction degree of temporal logic formulae to measure the performance of a biochemical system. As a globally relevant measure of system performance, it permitted to broaden the scope of these methods, whereas previous models for synthetic and systems biology were focused on specific behaviors.

3.1.4 Microfluidics for synthetic biology: methods for protocell fabrication

3.1.4.1 Introduction

Microfluidics is an area of science that studies the behavior and precise manipulation of fluids constrained by geometrical micrometric channels (i.e. corresponding to picoliter to nanoliter volumes). Microfluidics exploits the scaling effects of the microdomain to provide with new capabilities in biological engineering. In fact, system miniaturization gives the ability to enhance the control on biological fluids (in terms of speed, accuracy and efficiency) compared to traditional macroscale lab equipments and experimental methods⁶⁰⁴. A flow transforms time to space, which provides with the ability to integrate different dynamic sequential steps, and exert biophysicochemical operations on the flow. Moreover, the main characteristic of a flow in a microchannel is to be laminar instead of turbulent, which favors highly predictable and controllable operations. Transport phenomena can thus be precisely controlled (i.e. mixing, fast thermal and mass transfers, concentration, surface tension, confinement). One of the strongest rationale concerns the integration of lab capacities into multifunctional microfluidic chips (i.e. lab-on-chip), capable of automated sample preparation, high-throughput, fast and parallel processing, increased reliability, protection against contamination and evaporation, low energy and reagents consumption⁶⁰⁵. In a sense, the motivation for using microfluidic systems is analogous to replacing discrete electronic components with integrated circuits.

Importantly, microfluidic devices can be easily manufactured using well established methodologies that had been previously developed for microelectronic industries⁶⁰⁶. Additionally, soft lithography adapted to polydimethylsiloxane (PDMS, an optically clear polymer bearing interesting properties for biology such as gas permeability and non-toxicity) by Whiteside⁶⁰⁷, brought drastic improvement in microfluidic prototyping and accessibility as it opened the technology to non-experts in fluid physics⁶⁰⁸, and even towards *in situ* programmable formats on paper⁶⁰⁹. Moreover, CAD tools for chip manufacturing speed up the prototyping and transition from concept to realization (**Figure 3.5**). Microfluidics' strongest impact has probably been on the analytical and quantitative biology performance⁶¹⁰, especially high speed digital PCR⁶¹¹, structural biology⁶¹², genetic sequencing technologies⁶¹³, or high throughput single cell screening of synthetic organisms⁶¹⁴.

However, as we will see this methodology also concentrates great capabilities for the streamlining of bottom-up synthetic biology design, for fine fabrication of synthetic biological structures such as the assembly of parts and devices as well as the quantitative assessment of systems to support iterative design. It provides a technology to tackle the two major challenges that hinder the progress of synthetic biology, namely the complexity and the biophysical variations of the biological systems. In other words, the highly resolutive environments that microfluidic offers for the control of biological fluids and parts provides with an unprecedented technological solution to achieve the major goals of synthetic biology.

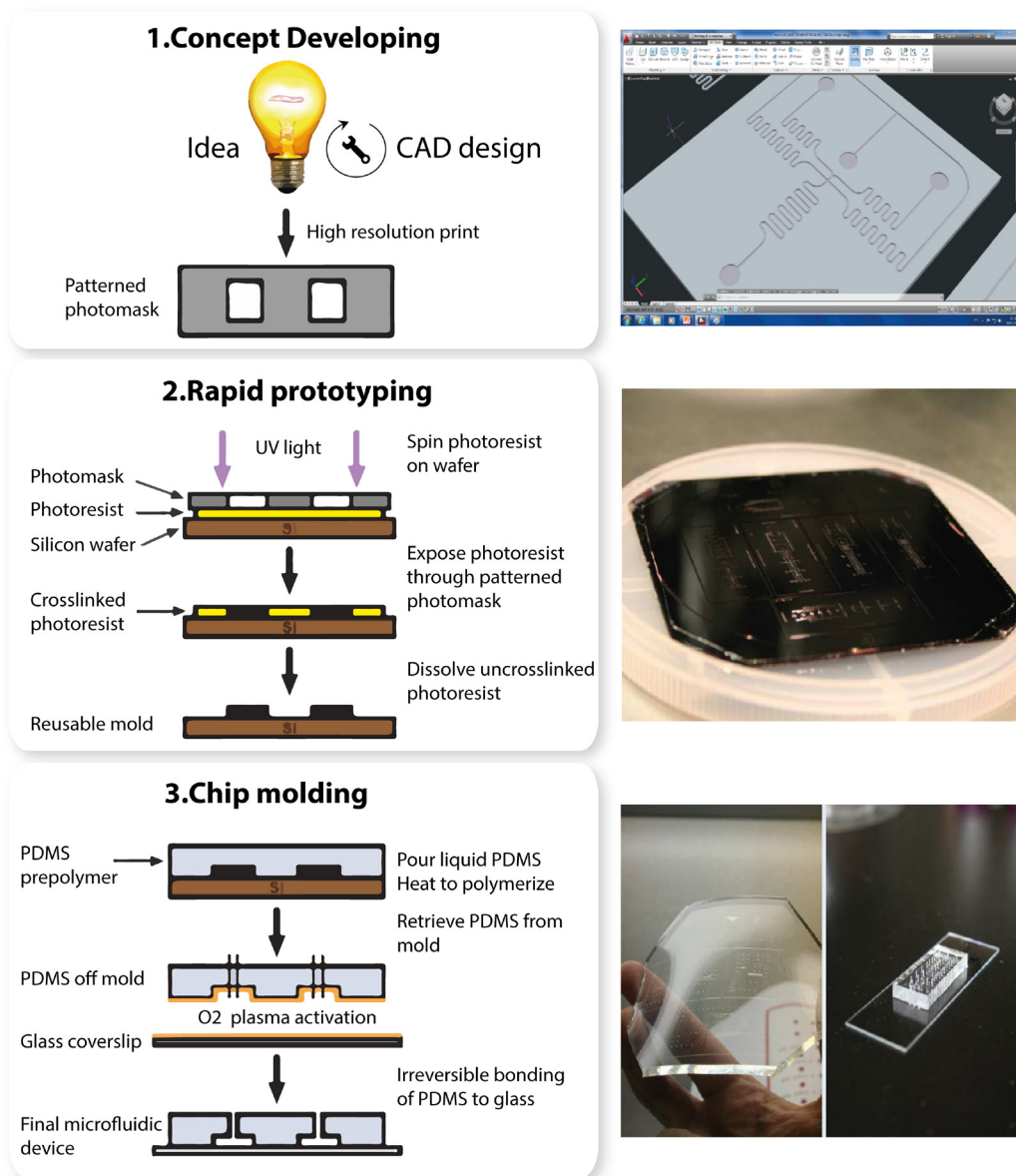


Figure 3.5: Straightforward framework for PDMS chip manufacturing: from concept to realization.

Microfluidics-based *in vitro* compartmentalization and droplet-based microfluidics are highly promising tools for performing fast and high-throughput digitization of space and time, and massive parallelization of biochemical reaction. Particularly, macroscale methodologies to generate highly precise biomolecular assemblies such as synthetic biochemical circuits and protocells are inexistent. Thus, numerous studies have recently focused on the development of microfluidic solutions to intelligent encapsulation. In this thesis, I adapted microfluidic methodologies for accurate *in vitro* protocell programming.

3.1.4.2 Theory: Basic microfluidic concepts

In this section we will briefly describe the physics behind microfluidic devices operation, which we will later exploit to achieve protocell fabrication.

First, one of the most important consequences of flow miniaturization is that fluid properties become controlled by viscous forces rather than inertial forces. This has major importance in microfluidics as we will describe below. This scale dependency can be represented within the dimensionless Reynolds number (Re), according to:

$$Re = \frac{\text{viscous forces}}{\text{inertial forces}} = \frac{\rho v l}{\eta}$$

where ρ is density, η the viscosity, v the characteristic velocity and l the characteristic length (i.e. the hydraulic diameter). Low Re values demonstrate a predominance of viscous forces, implying that the flow will follow strictly laminar behavior (streamlines constant with time) and absence of convection. Consequently, operations at this scale are highly reproductive and controllable. At high Re (i.e. ~ 1000) turbulence appears in the flow, for instance in macroscale tubing (e.g. river rapids). Under the microscale characteristic of microfluidic channels, Reynolds numbers are typically under 1. For instance, a Phosphate Buffer Saline solution at a flow rate of 1 $\mu\text{l}/\text{min}$ in a channel of characteristic length of 10 μm would give a Reynolds number of ~ 0.1 . Therefore, all designs I describe and used in this thesis fall under laminar flow with $Re < 1$. This is the definition of the so-called Stokes-flow.

A second important property comes from the fact that at the scale of microfluidic flows, Brownian diffusion is not sufficient to fully mix fluids. As mixing will only occur due to diffusion, the Peclet number is another dimensionless measure that can be used to represent the ratio between advection and diffusion of a substance, according to:

$$Pe = \frac{vl}{D} \qquad d^2 = 2Dt$$

where d is the distance a molecule travels according to Brownian motion, D is the diffusion coefficient, and t is the elapsed time. It is analogous to how far a molecule is carried versus how far it diffuses within the channel per unit of time. If mixing is required in microfluidic systems calculating the the Peclet number can prove critical. The length to achieve effective mixing is then given by:

$$\Delta y = \frac{vl^2}{D} = Pel$$

However, efficient and fast mixing with low reagent consumption can be achieved using chaotic advection devices. Chaotic advection triggers the mixing in laminar regime flows via the continuous shear forces happening in concentrated solute volumes⁶¹⁵. For instance, one of the most efficient designs for mixing is the staggered herringbone mixer which we incorporate in our design to achieve *in situ* assembly of biochemical circuits.

Further, fluid dynamics is governed by the Navier-Stokes equation, which applies to uniform viscous Newtonian fluids:

$$\rho(\partial t + v \cdot \nabla)v = -\nabla p + \eta \Delta v + F$$

Were the first, second, third and last terms correspond to acceleration (inertia), pressure forces, viscous forces, and volumic forces, respectively. Now, considering a laminar fluid flowing in a microfluidic channel at the steady state, since there is no acceleration/inertia, and viscosity determines flow rate, one can reduce the equation to:

$$0 = -\nabla p + \eta \Delta v$$

which demonstrates the balance between pressure and viscous forces. This consideration implies the mode of displacement of a fluid through microfluidic channels, which is strictly pressure driven flow, also called Poiseuille flow. The fluid is pumped through the device via positive displacement pumps (i.e. syringe pumps). This typically shows a parabolic profile characteristic of the Poiseuille Flow, with zero velocity at the boundary and maximum velocity in the center of the channel (**Figure 3.6**). One can then derives the Hagen-Poiseuille's equation, which relates pressure drop to volumetric flow rate in a channel, for viscous and incompressible fluids in laminar regime (fully satisfied in microfluidics devices):

$$Q = \frac{\pi r^4 \Delta p}{8 \eta L} = \frac{\Delta p}{R_h}$$

where Q is defined as the positive flow rate for the flow from microfluidic inlet to outlet channels, R_h the hydraulic resistance, L is the length of the channel, R is the radius of the channel, and Δp is the pressure drop through a channel. Interestingly, the Navier-Stokes equations reduce to a simple analog of Ohm's law, in the form:

$$\Delta P = Q R_h$$

This relation allows the easy determination of flow rate in a microfluidic chip from external pressure and channel resistance. R_h can be deduced from the dimension of the channel. For cylindrical channels, R_h is given by the Hagen–Poiseuille equation according to:

$$R_h = \frac{8 \mu L}{\pi r^4}$$

However, most chips have rectangular features, which require the use of a somehow more complex equation:

$$R_h = \frac{12\mu L}{wh^3} \left[1 - \frac{h}{w} \left(\frac{192}{\pi^5} \sum_{n=1,3,5}^{\infty} \frac{1}{n^5} \tanh h \left(\frac{n\pi w}{2h} \right) \right) \right]^{-1}$$

Which fortunately can be reduced to a working equation when using chips with high aspect size ratio of channels ($w \gg h$, for example $5 \mu\text{m}$ high with $100\mu\text{m}$ wide channels):

$$R_h = \frac{12\mu L}{wh^3}$$

Therefore, the flow rates in a microfluidic chip can be deduced in a straightforward manner, using methods analogous to electrical circuits. The analogy between fluidic and electric circuit is indeed quite intuitive, where the molecules of fluid are analogous to electrons. This has practical applications for the experimentalist, as it can be straightforwardly used to manipulate flow/pressure relation in the microfluidic chip (**Figure 3.6**).

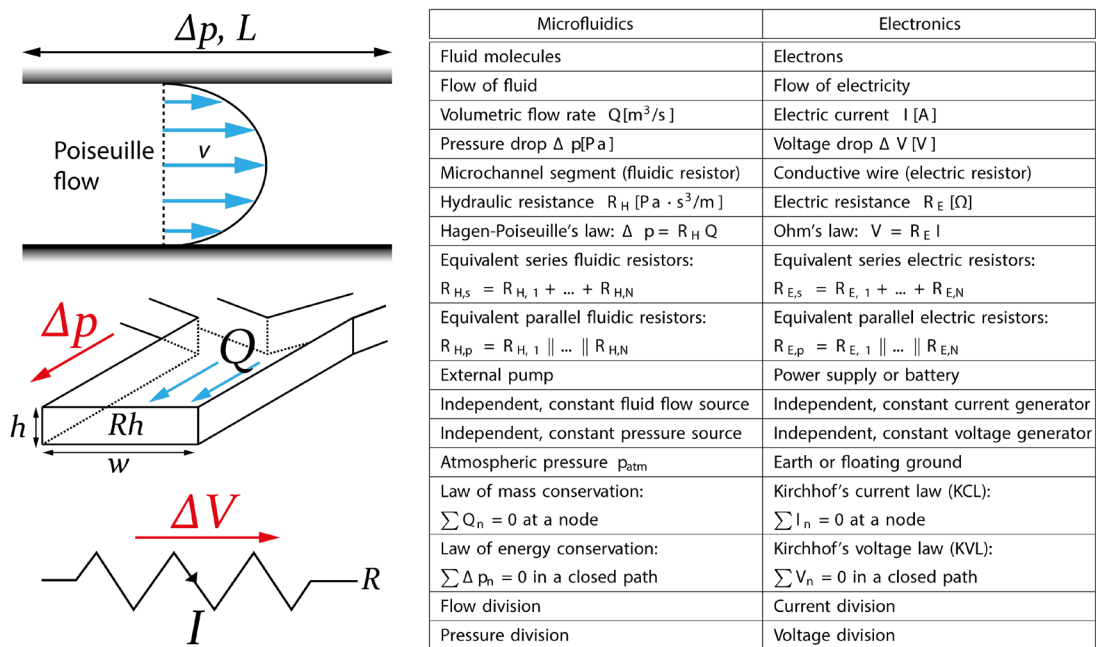


Figure 3.6: Basic microfluidic concept and useful analogy with electronics (Table adapted from Biral⁶¹⁶)

Last but not least, microfluidics is ruled by surfaces, meaning surface tension dominates over other forces. In the case of this thesis, this consideration is of utmost importance since the physical phenomena involved with protocell formation are ruled by surface tension (as all droplet based microfluidics). The so called Capillary number (Ca) is then used to represent the relative magnitude of viscous forces versus surface tension acting at the interface between two immiscible fluids. Like the Reynolds number, it is a dimensionless parameter that can be expressed as:

$$Ca = \frac{\eta v}{\sigma}$$

With η the viscosity of the dispersant fluid in which the droplets are dispersed, v the velocity and σ the interfacial tension coefficient between the two phases. Viscous stress can be regarded as a destructive force, whereas interfacial tension as a cohesive force. The capillary number enables to assess the cohesion of the droplet (i.e. increasing the Capillary number will increase the probability of droplet splitting).

This also implies the Rayleigh-Plateau instability, which is the law explaining how a fluid stream tends to minimize its surface area, and thus will eventually break into monodisperse droplets. One can evaluate its importance from the Young-Laplace equation below:

$$\Delta p = \sigma \left(\frac{1}{R_1} + \frac{1}{R_2} \right)$$

Where Δp is the pressure gradient, R_1 and R_2 are the principal radii of curvature. This describes how as a perturbation in the flow grows, the more important the curvature of the surface, which eventually leads to an increase in pressure gradient and final rupture and droplet formation. The surface tension acts as a gain on this surface shape perturbation, that is, the higher surface tension, the most probable and faster the flow breakage.

3.2 Operational principles, design and architecture of protosensors

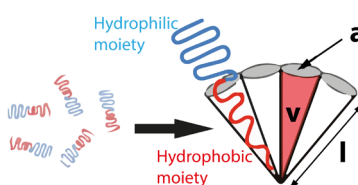
3.2.1 Architecture and general functioning

The protocells we describe in this work consist in vesicular synthetic biological systems which composition is programmed to achieve medical biosensing and biocomputing tasks when driven toward thermodynamic equilibrium. Similarly to cells, in our architectures the biochemical work necessary to support functioning (i.e. signal sensing, processing and output generation) originates from redox reactions. This useful potential biochemical energy is either brought during fabrication and stored as encapsulated electrons donors (such as NADH for example) or is extracted from energy rich molecular inputs (e.g. glucose).

Protosensor architecture thus consists of biochemical synthetic circuits encapsulated within a phospholipid bilayer, which enables to digitize space through the definition of an insulated interior containing the synthetic circuit, and an exterior consisting of the medium to test (e.g. a human clinical sample).

In this study, protocells designate specific types of vesicular compartment of biological size (also known as Giant Unilamellar Vesicles, GUVs $\sim 10 \mu\text{m}$), encapsulating a complex aqueous medium comprising the synthetic circuit. Although many types of protocells have been described regarding the nature of their membranes, they are generally composed of highly ordered amphiphilic molecules. These amphiphiles, for instance phospholipids or synthetic copolymers, comprise hydrophilic headgroups and hydrophobic chains, which can assemble into a bilayer. Orientation of hydrophilic heads in contact with the aqueous medium and hydrophobic chains with the interior in each layer is thermodynamically favored. The physicochemical properties of protocell membranes bilayers strongly depend on the nature of the amphiphiles, which will impact permeability, thickness, stability, or elasticity.

Not all amphiphiles can sustain assembly into vesicles, and chemical structure has an important impact on membrane thermodynamics. The dimensionless packing parameter P represents the molecular shape of amphiphiles in solution, and rules the morphology of the corresponding assembly. It can be expressed as:

$$P = \frac{v}{al}$$


where v is the volume of the hydrophobic moiety, a the hydrophilic-hydrophobic interfacial area, and l the hydrophobic moiety length, as illustrated. In order to form stable bilayer and protocell

vesicular assemblies, P has to be ~ 1 . In our study, we thus investigated the use of common phospholipids with appropriate packing parameter that have been widely used for protocells fabrication, namely Dimyristoylphosphatidylcholine (DMPC), Dioleoylphosphatidylcholine (DOPC), and Dipalmitoylphosphatidylcholine (DPPC).

Importantly, in order to perform biosensing, protosensors require an exchange of matter and information between their interior and with the medium they evolve in. To this end, after encapsulation of synthetic circuits I perform a subsequent membrane modification step. I use an approach in which the phospholipid membrane is rendered selectively permeable to small organic molecules that serve as inputs for our systems, through the self-incorporation of α -Hemolysin transmembrane protein pores, which has a mass cutoff of 3 kDa⁵⁷⁰ (See **Materials and Methods**).

3.2.2 Programming *in vitro* algorithms for the differential diagnosis of acute diabetes complications

Our goal was to engineer medical protosensors capable of implementing *in vitro* diagnostic processes (i.e. diagnosing specific pathologies through the biodetection of patterns of biomarkers) formalized as Boolean functions, using synthetic biochemical circuits as a substrate (**Figure 3.8A**). For this purpose, we chose to evaluate the feasibility of integrating a clinically useful medical algorithm enabling to classify acute metabolic complications of diabetes, namely diabetic ketoacidosis, hyperglycemia hyperosmolar state, hypoglycemia and lactic acidosis (**Figure 3.7**). These disorders constitute medical emergencies, and are known to be associated with high medical and socio-economic burden and with an important mortality and morbidity. In addition, we also intended to implement a screening solution for the early onset of diabetes. Hence, not only this study addresses novel engineering concepts, it also tackles concrete clinical problems that are seeking solutions.

These diagnostic processes are possible via the monitoring the presence of 5 different urinary biomarkers and the assessment of their specific concentrations and combinations in urine, namely: glucose, ketones, lactate, ethanol and nitric oxides (**Figure 3.7**). Therefore, we focused on implementing synthetic circuits capable of processing these specific inputs using simple Boolean operations. With this concrete proof of concept in mind, we also emphasized on developing a universal framework permitting systematic and versatile programming of protosensors for other agendas and pathologies. In fact, a large number of different protosensors could be programmed to answer multiple but specific clinical questions, in order to multiplex disease diagnosis *in situ* (**Figure 3.8B**).

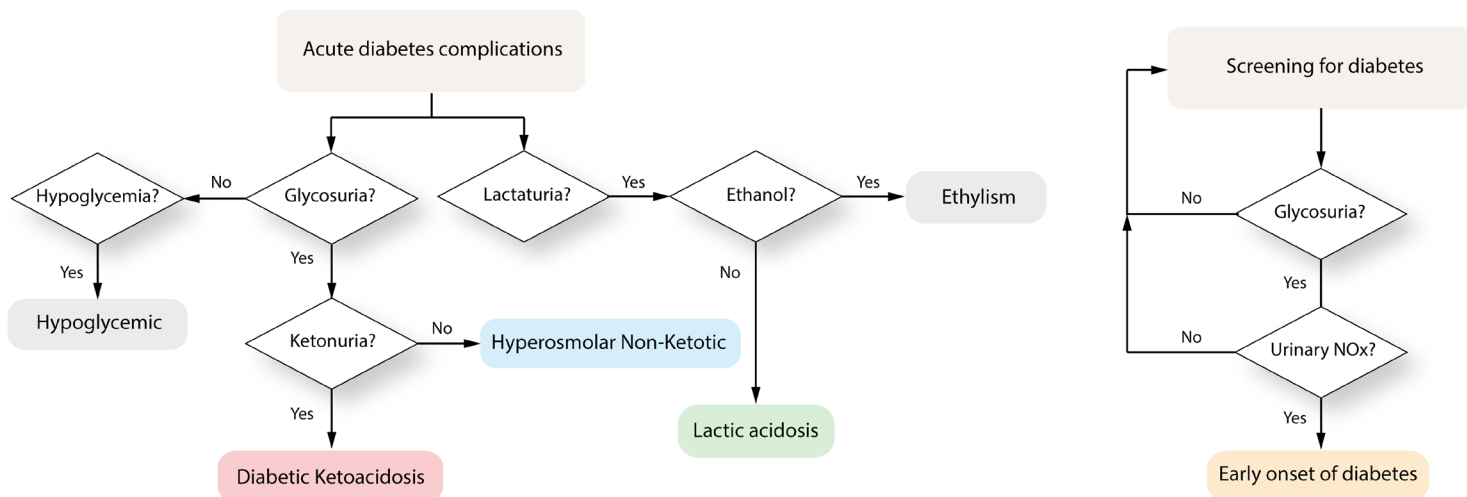


Figure 3.7: Proof-of-concept diagnostic algorithm used in this study, and integrated within protosensors to achieve differential diagnosis of diabetes acute complications, and screening for diabetes. Diabetes associated acute complications, namely diabetic ketoacidosis, hyperglycemia hyperosmolar state, hypoglycemia and lactic acidosis, are clinical emergencies that represent a major health care burden associated with severe mortality, morbidity and frequent complications. Here we propose a diagnostic algorithm enabling diagnosis of these complications, as well as a proof-of-concept screening assay, from markers present in urines.

3.3 Molecular programming of protocells: *In silico* design to experimental validation of synthetic biochemical circuits

3.3.1 *In silico* design, simulation and model checking

The capacity to rationally design biological systems to achieve programmed biosensing and user-defined decision making algorithm and bioactuation requires precise tools within the scope of a systematic approach. Therefore, we first developed an *in silico* framework supporting the design of synthetic biochemical circuits from the bottom-up assembly of biological parts (**Figure 3.8**). This computer assisted framework involves the following steps:

- (i) Design of abstract programs with respect to Boolean logic and molecular input/output specifications according to (medical) algorithm of interest. At this step, one can formalize the temporal logic properties of a synthetic biochemical system regarding expected reference behavior, expressed by a qualitative and/or quantitative temporal logic formula³⁶⁹. This comprises the specification of parameters relative to the initial state (i.e. pathological biomarker concentration thresholds for instance).

- (ii) Implementation of previously defined algorithms using molecular biochemical circuitry: identification of suitable enzymes and metabolites within a network topology obtained from databases and literature to implement molecular Boolean logic operations. Composable and kinetically favorable components are chosen at this step to minimize modes of failure. This process can be automated using recently developed computer assisted extraction of biochemical parts, and Boolean logic gates from metabolic networks of living organisms⁶¹⁷. Experimental characterization of biochemical modules can help this process. Netdraw software is used for user friendly design and mapping of interaction network and to assign reaction rules (with or without kinetic expressions), concentration parameters, as well as spatial parameters such as volume, location of species, and permeability coefficients of compartments. Netdraw is then used for automated generation of HSIM and Biocham code for the model (specification of the initial state, definitions of numerical parameters, compartments volume, invariants, events, declarations of molecular species and locations, specification of the system's behavior).
- (iii) Stochastic simulation (SSA) is first performed within the HSIM software to verify kinetically and functionally favorable circuits, predict the overall behavior, estimate the functioning and manually explore the design space to identify suitable parameter configurations to be refined.
- (iv) Biocham simulations (ODE solver) are then carried out to compute validity domains of specified behavior for the system, perform sensitivity analysis to identify sensitive parameters that can then be iteratively optimized, and measure robustness relative to the variation of specific parameters. Models are thus evaluated with respect to temporal logic specifications. Computing a landscape of satisfaction relative to sensitive parameters enables to visualize and identify suitable parameter space satisfying specified behavior, which can be used to select robust parameters for experimental *in vitro* implementation (i.e. concentration of species for instance, kinetic parameters, initial value or control parameters). One can automate the search, using CMAES methodology (Covariance Matrix Adaptation Evolution Strategy)⁶¹⁸ integrated in Biocham for those parameter values that satisfy a given set of quantitative temporal properties.
- (v) Once the parameters satisfying user-defined system specifications have been found, HSIM stochastic simulator can be used to validate and finely map the complete transfer functions of the protosensors. The systems can then be experimentally implemented and its functioning assessed *in vitro*. Iterations in the design process can occur at each step.

The medical algorithm depicted in **Figure 3.7**, was divided and distributed into three simple biochemical systems (whose truth tables are depicted in **Figure 3.9A**) taking two biomarkers as inputs, which we named GluONe (Glucose and Acetone as inputs), LacOH (Lactate and Ethanol as inputs) and GluNOx (Glucose and Nitric oxides as inputs). The biochemical implementation (i.e. comprising network topology) was assisted *in silico* using the custom computer tools developed for this purpose, which support the systematic mining of natural biochemical network to retrieve enzymatic Boolean logic gates and circuits satisfying user-defined specifications (**Figure S3.5**). We identified three distinct and insulated minimal systems that could operate in parallel, which required 6, 5 and 4 different biochemical entities, comprising 4, 3 and 2 different enzymes respectively. Biochemical knowledge on the parts that enabled formal design was extracted from BRENDA database (See **TableS1** and **S3.3**, and **Figure S3.6** for more detail). Molecular signal processing occurring in these circuits leads to the biochemical synthesis of the following measurable output signals molecules: NADH (Output 1, 340 nm absorbance), Resorufin (Output 2, 571-600 nm fluorescence), ABTS (Output 3, 420 nm absorbance), and DAF (Output 4, 488-515 nm fluorescence). The Boolean formalism and truth tables corresponding to the medical algorithm, as well as the biochemical implementation are depicted in **Figure 3.9A** (Detailed in **Figure S3.6**).

I first performed stochastic simulation in HSIM to evaluate the putative behavior of the three circuits (**Figure 3.9B**). To this end, I used non-optimized models of non-encapsulated synthetic circuits, where initial conditions (i.e. species concentrations) were here determined empirically. I predicted systems state after induction with various concentrations of biomarker inputs, and represented the computed molecular output signals as heat maps. The relation between *in silico* calculated molecular concentrations and experimental measured signal was calibrated beforehand (See methods for detail). This permitted us to validate the Boolean logic operations, with very satisfying theoretical signal fold change, as well as near digital response. In addition, we found that switching thresholds for these models were in good agreement with useful clinical sensitivity for biomarker inputs (pathological thresholds: Ketones > 17 μM ($\sim 10\text{mg/dl}$); Glucose > 1.39 mM ($\sim 25\text{mg/dl}$); Lactate > 10 μM ; EtOH > 17.4 mM ($\sim 80\text{mg/dl}$); NOx > 1000 μM).

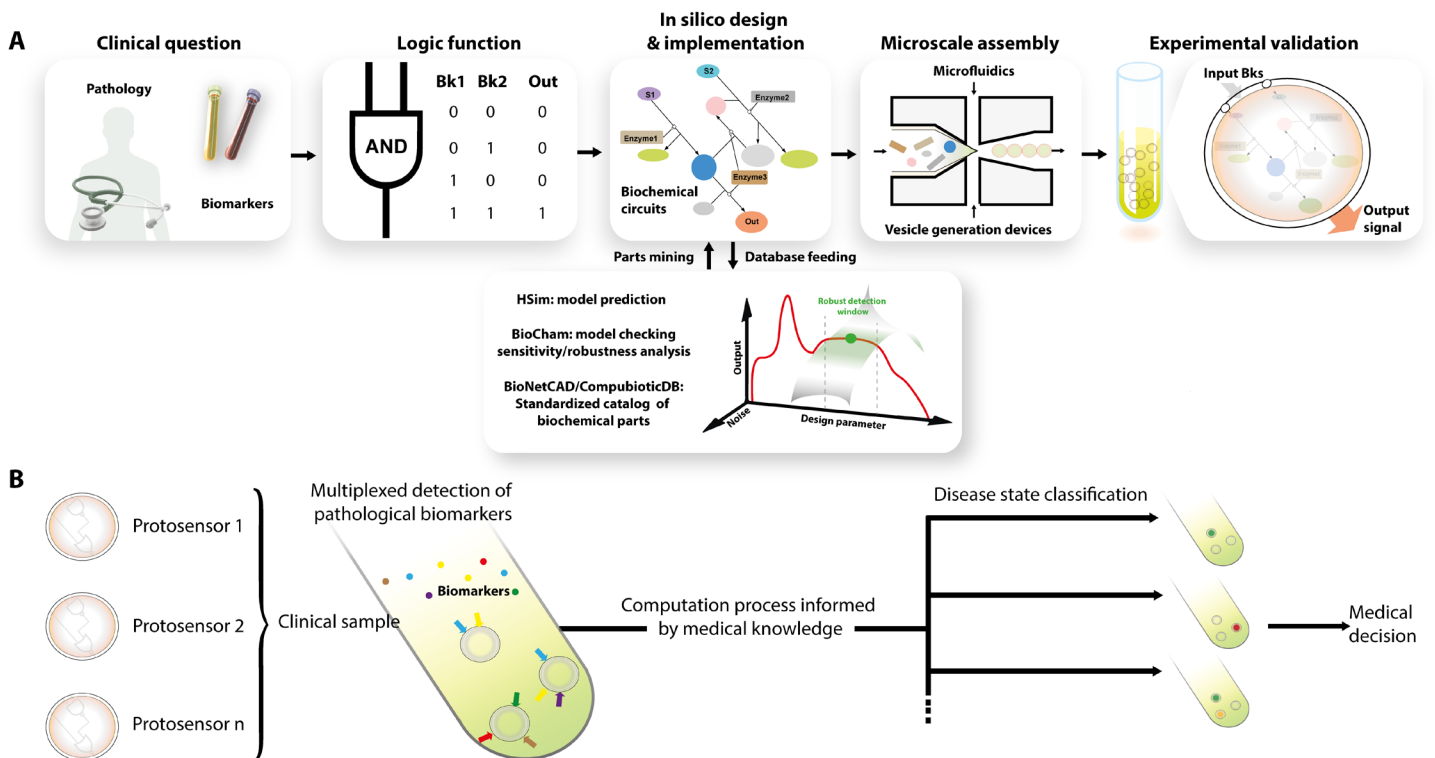


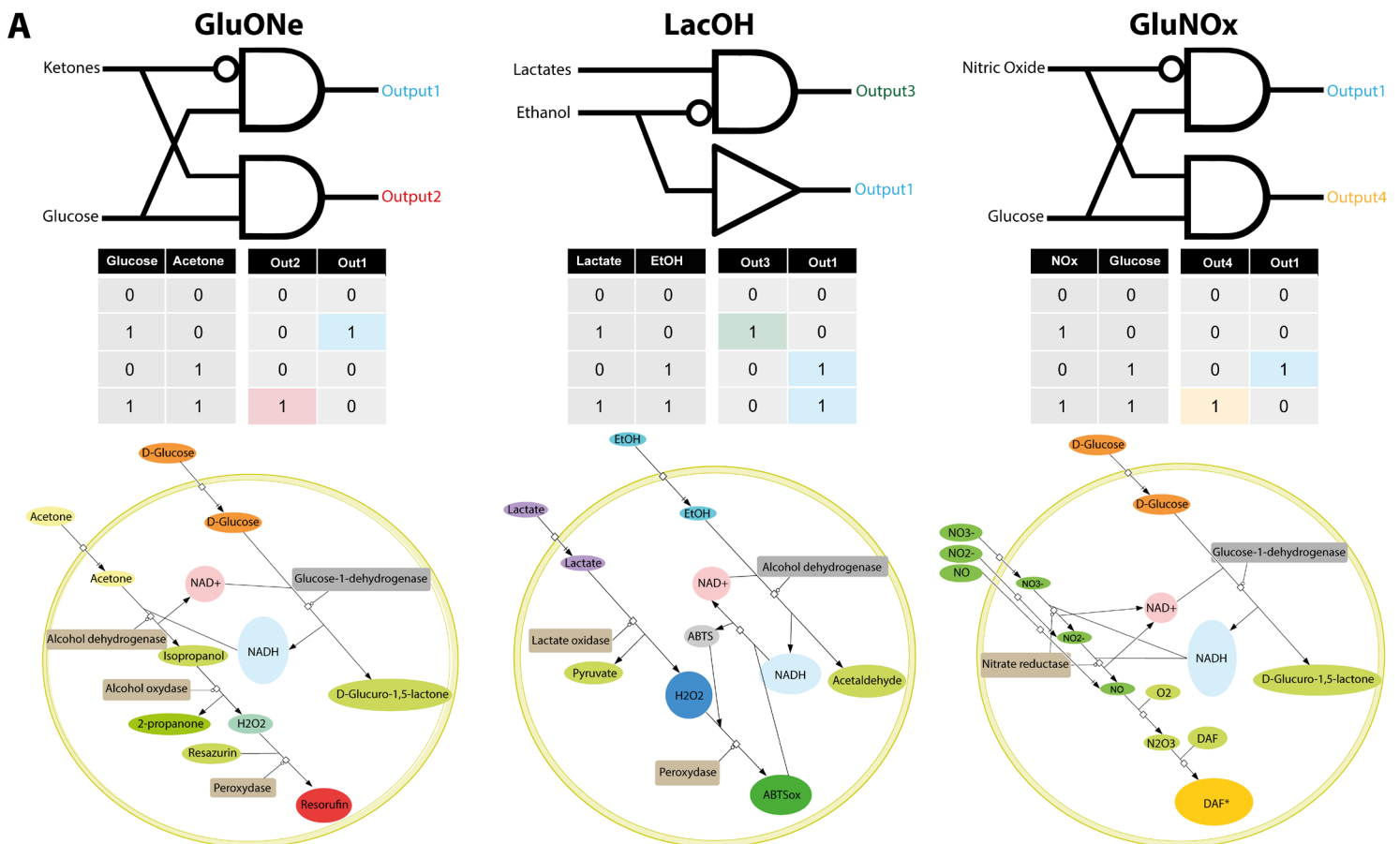
Figure 3.8: General design methodology, architecture and operational principle of protosensors for medical diagnosis. (A) Arising from a clinical need to detect pathologies associated with patterns of specific biomarkers, medical diagnosis can be abstracted to a computational process formalized using Boolean logic *in vitro* and embedded into synthetic biochemical circuits. This *de novo* circuits can be assembled, or programmed *in silico* from naturally occurring building blocks, and encapsulated in synthetic containers to yield diagnostic devices, or protosensors, capable of detecting patterns of specific biomarkers in human clinical samples and integrates these signals in a medical decision algorithm. If a pathological pattern of biomarker is detected, the protosensor generates a colorimetric output. **(B)** Different types of protosensors corresponding to different clinical questions can be used at the same time to enable multiplexed detection of pathological biomarkers and achieve differential diagnosis of pathologies in clinical samples.

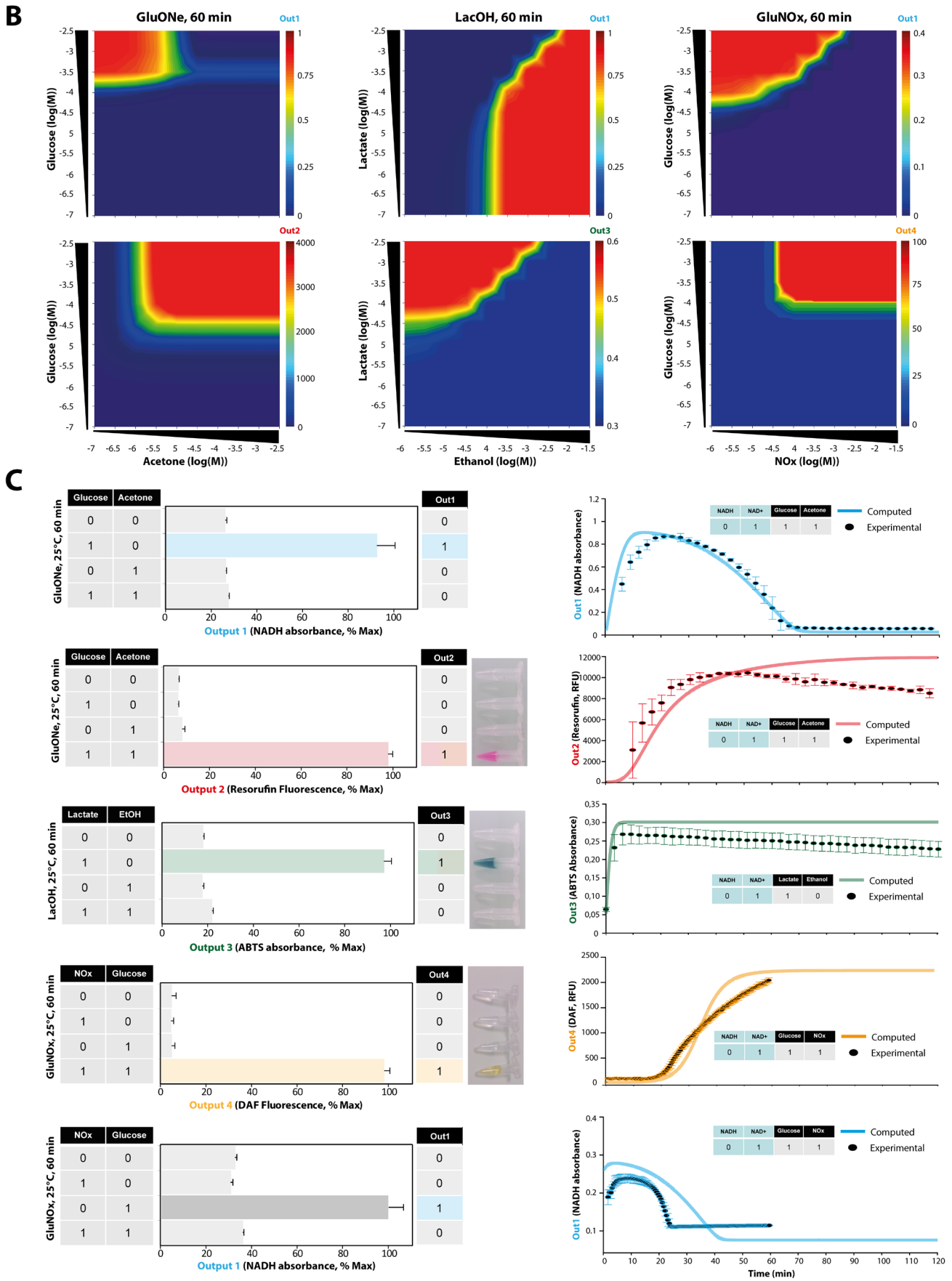
3.3.2 Experimental validation *in vitro*

I then investigated the experimental behavior of the rationally designed synthetic circuits prior to encapsulation inside protocells. I proceeded to *in vitro* implementation in the test tube of previously simulated models with the same initial conditions (i.e. concentration of circuits components) using recombinant enzymes and synthetic metabolites at room temperature. I performed multiple experiments consisting in varying initial conditions (presence/absence of components of the circuit and presence of pathological concentrations of input biomarkers), and measured the generation of output signals (fluorescence or absorbance). This allowed us to get a fine understanding of the functioning and detailed experimental characterization of logic operations (**Figure 3.9C, S3.9** and **S3.10**). Interestingly, we found that kinetic and end point measurement showed very good agreement with HSIM predictions. In addition, the circuits

behaved in exact accordance to Boolean logic specifications with temporal requirements of less than 60 minutes. Outputs 2, 3 and 4 delivered a human readable output signal as expected. Considering signal-to-noise ratio (SNR) as a quantitative measure of biocomputing efficacy⁶¹⁹, we found that these synthetic biochemical circuits showed very good performance in processing molecular signals according to specified Boolean Logic, with calculated SNRs for outputs 1,2,3 and 4 of ~20, 34, 14, 26 dB respectively.

While in this study we were interested in exploring *in silico* the transfer function of the LacOH system for output 1, we did not yet proceed to experimental validation. Indeed output 1 and 3 are in this set up overlapping in absorbance, which could be easily subtracted to obtain a dual readout. However this was not the immediate purpose of this work since it did not present direct clinical utility. Likewise, in the following experimental work we only measured outputs 3 and 4 for LacOH and GluNOx systems. Nevertheless, in a near future these systems will be the subject of further *in vitro* characterization.





3.4 Microfluidics approach for protosensors prototyping and fabrication: encapsulating synthetic circuits in protocells

3.4.1 Theoretical considerations

Developing protocells as robust biosensors for biomedical analysis requires finest control on physicochemical properties. Thus, production methods ensuring fine tuning on parameters was required, to achieve stable protocells with well-defined size, lipid composition, and enzyme content, catalytic performance and stability of encapsulated biochemical species.

Protocell membrane systems exist in states that are kinetically favorable, or *trapped* within local energy minimums, instead of rather real thermodynamic equilibrium⁶²⁰. This implies that the efficiency of fabrication will strongly depend on the choice amphiphile biochemistry as well as on the physical assembly process. This also means that obtaining important yields and reliable encapsulation process are not an easy technological journey.

Since the first published method of cell sized vesicle formation in 1969 through gentle hydration and electrosweeling⁶²¹, most developed methodologies can be separated into two common fabrication techniques: hydration and electroformation. The most described and commonly used techniques involved hydration of a dried phospholipid film on a glass surface. For the purpose of this study, I first explored these basic techniques. This process showed extremely sensitive to perturbation, demonstrated poor yields, no control on polydispersity and size of vesicles, low reproducibility and most importantly was incapable of accommodating precise stoichiometry in encapsulation of various biochemical entities.

A more recent and clever method was then described, now known as the water-in-oil emulsion transfer method. It was first developed by Pautot *et al.*⁶²² and proved of utmost interest in protocell research³⁸⁶. Briefly, this technique relies on the generation in a first step of water-in-oil droplets, which allows phospholipids to organize at the phase interface. Then in a second step, these droplets are transferred through another phase interface, thereby producing a bilayer. Although proving as an interesting, precise and versatile technique, it remained incapable of satisfying monodispersity and high throughput requirements.

After its first apparitions, microfluidics shortly generated avenues to engineer amphiphile vesicles tailored with high degree of precision⁶²³. Traditional approaches for fabricating vesicles rely on the slow and low efficiency of self-assembly of amphiphiles. Microfluidic devices however allow precise control on directed membrane assemblies, and generated a vast scientific literature. Methodologies such as hydrodynamic flow focusing⁶²⁴, pulsed jetting⁶²⁵ or emulsion-templating⁶²⁶⁶²⁷⁶²⁸ emerged and brought new opportunities to engineer vesicular structures and tailor them for specific applications. Hence, in this thesis I explored the encapsulation of synthetic biochemical circuits through two microfluidic approaches: directed and undirected self assembly (**Figure 3.10**).

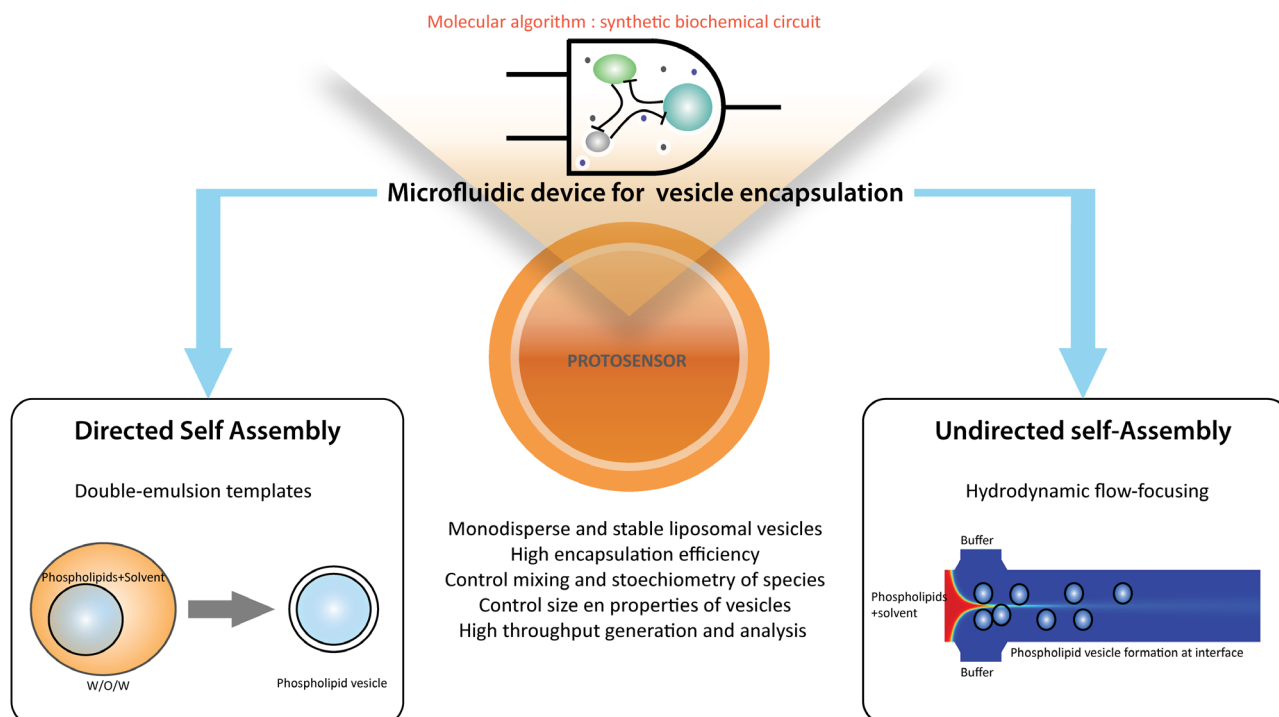


Figure 3.10: Microfluidic approaches explored in this thesis to achieve controlled and high-throughput protocell fabrication.

The approach I used for undirected self-assembly relied on microfluidic continuous-flow mixers, also known as flow focusing devices, where a solvent containing an amphiphile of interest is injected into a center inlet of a microchannel cross junction and pinched into a smaller flow by solvent streams on side channels (**Figure 3.10**, right). Mixing driven by molecular diffusion appears at the interface between the two hydrodynamically focused flows⁶²⁹ in a flow rate dependent way. Flow rates control the width of focused flow, and thus mixing with time and space resolution, enabling complex yet precise mixing processes that could not be achieved in a macroscale context. These performances motivated the use of these devices to control the assembly of phospholipids into structured bilayers, such as the engineering of liposome and polymersome with monodisperse sizes. Amphiphile vesicles are allowed to form at the interface of flow-focused fluids by a complex mechanism relying on sustained bilayer growth by for instance phospholipid membrane self-assembly at the interface. As phospholipid concentration at the interface depends on flow rates, one can tune vesicle size by simply controlling the flow rate^{629 630 624 631}. Although very appealing, when put into application this approach suffered from unreliable encapsulation efficiency, which could show thermodynamically favored encapsulation of small molecules compared to large proteins, a phenomenon which would dramatically affect the programmability of protocell assemblies. Moreover, the largest liposome sizes achievable with this technique were shown to be ~150 nm, which would limit the complexity of the encapsulated material.

In this context, I decided to pursue another microfluidic approach, which this time relied on directed self-assembly and encapsulation. Using microfluidic technologies, it is easy to generate monodisperse single water-in-oil emulsions, which constitutes a whole field commonly known as *Droplet Microfluidics*. Interestingly, this approach can be extended to double water-oil-water emulsions, and further exploited to generate emulsion-templated vesicles, facilitated by the routine capabilities to generate uniform size calibrated emulsions with near total encapsulation efficiency. Double emulsion templates can be generated to mimic a vesicular membrane, which is in this case composed of an oil phase. One can then tune the emulsion generation to obtain double emulsions with thin oil phases. If an amphiphile is dissolved in the oil phase, these architectures can thus be converted into real amphiphile bilayers vesicles via oil removal^{632 633} (**Figure 3.11**).

Therefore, the main microfluidic designs I explored in this thesis exploit droplet generation mechanisms. Similarly to the first designs, they can also be described as flow focusing devices, but this time are used to introduce shearing at the interface of different phases in order to generate droplets. The presence of two phases (i.e. Oil-Water) introduces extra complexity in the system, where inertial, viscous and interfacial forces arising from different fluids will compete⁶³⁴. However, as previously discussed, microfluidic domain deals with Stoke flows, where the laminar characteristics associated with dominating surface tension creates uniform interfaces.

In this thesis, I adapted the one step double emulsion generation which had been demonstrated a decade ago by Utada *et al.*⁶³⁵. Since then, many designs successfully intended to adapt this strategy within microfluidic channel geometries. The underlying mechanism exploits a droplet forming regime of an aqueous droplet into an oil phase, itself dripping in another aqueous phase. This happens when both dispersed phases are in a dripping regime and shear produces break at the same time. The behavior of different phases regarding the wettability to the channel surface is of utmost importance for droplet generation. In fact, the phase to be dispersed into the other one should be non-wetting to the channels, to ensure total efficiency (i.e. reduce phase inversion). Importantly, wetting characteristic in PDMS devices can be increased using a chemical treatment, for example in our case, a polyvinyl alcohol treatment ensures perfect water wettability and forbids oil wettability⁶³⁶. The oil and water phases having different velocities, a viscous shear is induced at their interface. This phenomenon is stabilized by capillary stress coming from interfacial forces, which tends to minimize the area of the interface, thereby producing droplets. As discussed, this mechanism can be represented by the Capillary number. While interfacial forces are ruled by area, viscous forces are volume based. This is of importance since at low capillary numbers (i.e. microchannel assumption), viscous forces and thus flow rates will be the main influences driving droplet breakage. As the capillary number depends on surface tension, one can facilitate droplet breakage and stabilize droplet spherical shape through the addition of a surfactant (i.e. decreasing surface tension and capillary number). In our designs, we used high flow rates in order to stay within a *jet-like* configuration ensuring highest fabrication throughput.

Our strategy thus relied on a microfluidic flow-focusing droplet generation design that generates water-in-oil-in-water (W–O–W: Biochemical circuit in Phosphate buffer - Phospholipid in Oleic acid - Buffer) double emulsions. Double emulsion templates are generated in described flow-focusing channel geometries (**Figure 3.13A**). DPPC phospholipid membranes then self-assemble during a controlled solvent extraction process where Oleic acid is extracted by methanol present in buffer (**Figure 3.11B, 3.11C**). The rationale we use for choosing DPPC concentration in oleic acid consists in calculating the concentration so that there would be a sufficient number of phospholipids to form a lipid bilayer around a 10 μ m diameter vesicle (Detailed calculation can be found in **Materials and Methods**). Oleic acid and methanol present the advantage of being biocompatible and non-toxic to enzymes compared to other inorganic solvents, and was chosen to minimize deleterious chemical interactions.

In addition, we investigated DOPC, DMPC phospholipids for protocell fabrication, and found that DPPC achieved better apparent stability and superior production yields. I found that DPPC vesicles were the most robust phospholipids for protocell construction, since they were capable of withstanding osmotic stress and showed prolonged stability and robust encapsulation at room temperature (**Figure 3.13C**). This can be explained by the fact that an increase in acyl chain length (and therefore lipid transition temperature) as well as the complete saturation of the acyl chain, is directly proportional to stability. Moreover, DPPC offers greater orthogonality and versatility, since they are less permeable, less susceptible to oxidation and disruption by natural proteins (e.g. serum proteins)⁶³⁷.

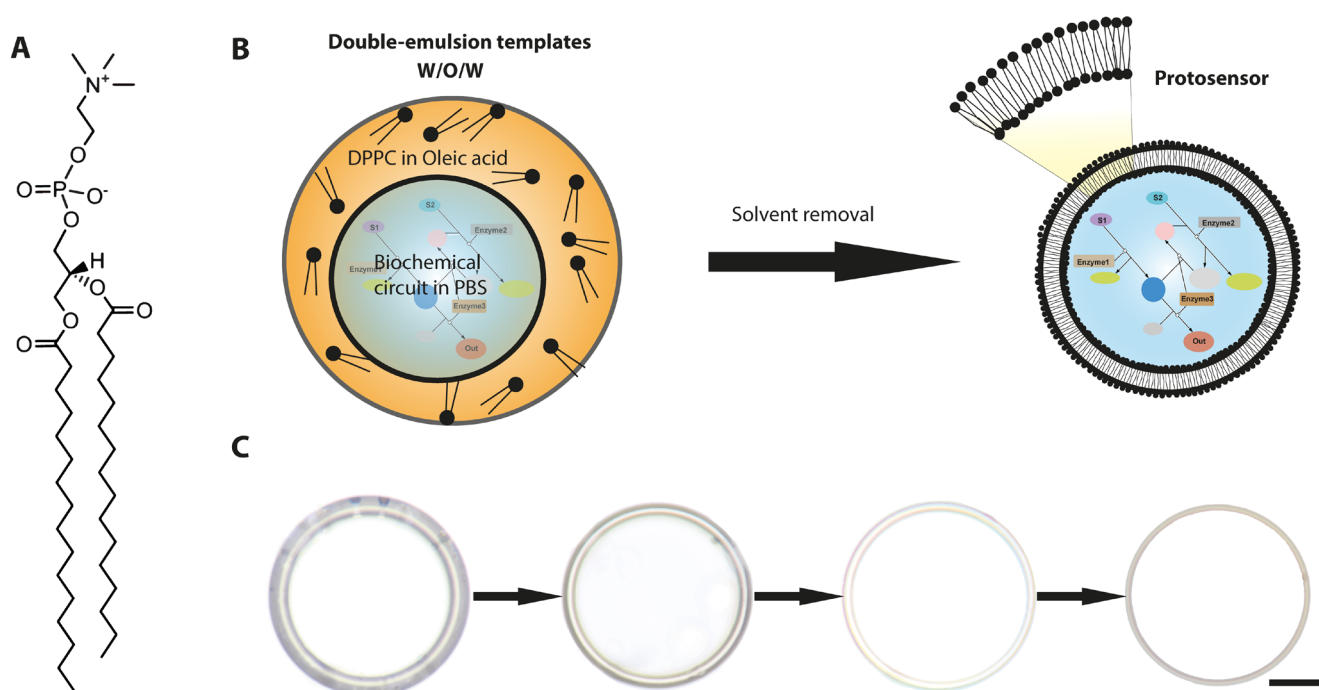


Figure 3.11: Double emulsion template for protosensor fabrication. (A) Chemical structure of Dipalmitoylphosphatidylcholine (DPPC) (B) The microfluidic flow-focusing droplet generation device generates double emulsion templates, and Oleic acid is then extracted to generate protosensors. (C) Visualization of the extraction process at 0, 1, 2 and 3 hours (Scale bar=2.5 μ m).

3.4.2 Experimental set-up and results

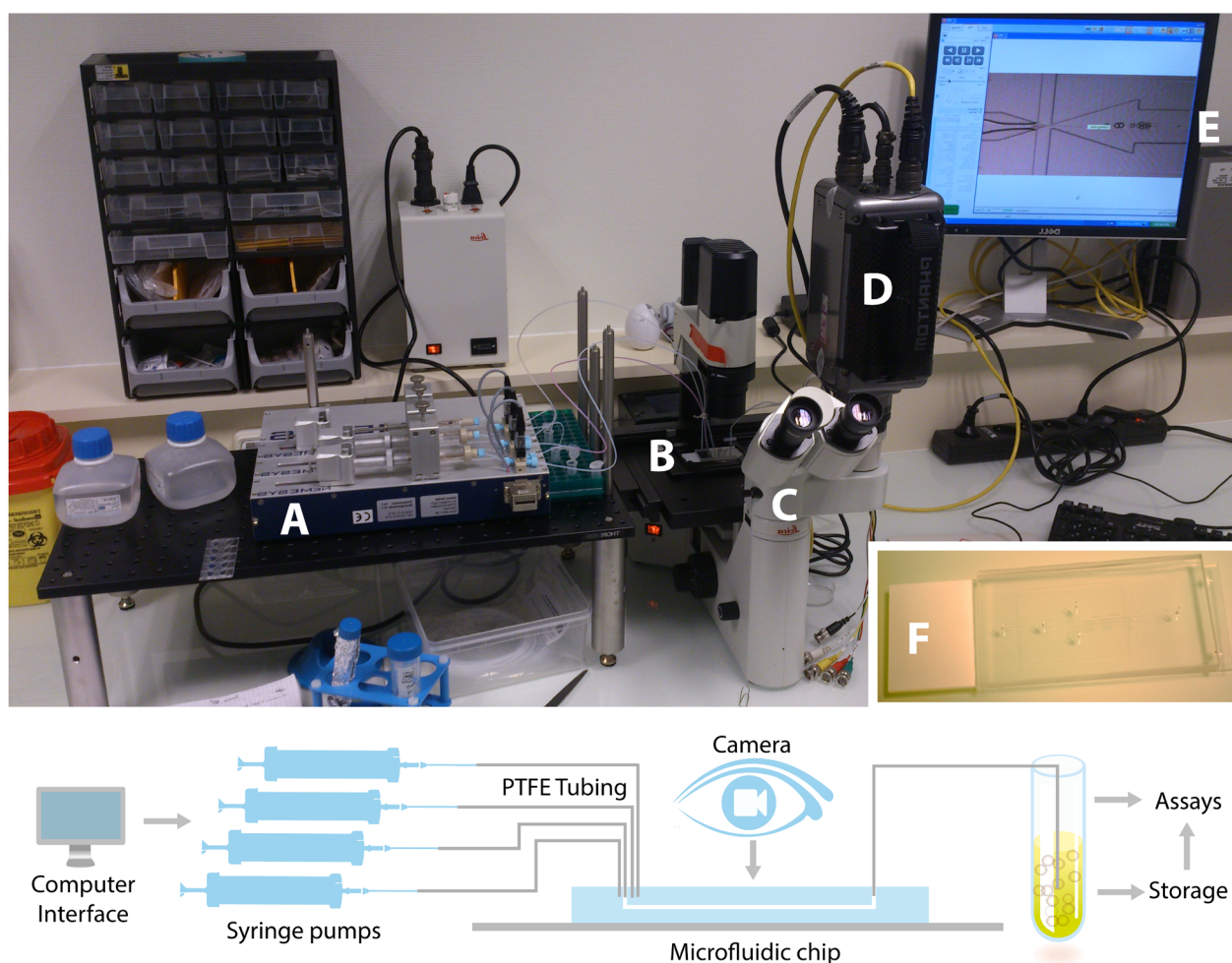


Figure 3.12: Experimental setup for microfluidic production of protocells. (A) Syringe pumps enable displacement driven flow in the microfluidic device (B). Protocell formation is under microscopic control using an inverted microscope (C) mounted with a ultrafast camera (D). A computer interface enables to visualize in real time the fabrication process (E), and gives nanometric control on syringe pumps flow rates. (F) Photograph of the PDMS chip used in this thesis to generate protocells.

The working microfluidic set-up is depicted in **Figure 3.12** (extra experimental details can be found in **Materials and Methods**). Briefly, the flow in microfluidic channels was controlled via displacement driven flow using nanometric syringe pumps equipped with high precision glass syringes, and controlled in real time with a computer interface. I found strong dependence of protosensors yields of production on flow rates, that were kept at $1/0.4/0.4 \mu\text{l}/\text{min}$ (Storage buffer/Oil+Phospholipids/Biochemical circuit, respectively) to achieve best encapsulation efficiency. Real time visual monitoring of the fabrication process enabled precise control on the fabrication process. In addition, measurements from the ultrafast camera at 20 000 FPS allows us to estimate around $\sim 1500 \text{ Hz}$ the frequency of protosensor generation at these flow rates (**Figure S3.2**).

The microfluidic flow-focusing droplet generation design is depicted in **Figure 3.13A** (See **Annexes** for CAD plans). I introduced in our microfluidic design a device previously described, known as the staggered herringbone mixer (SHM)⁶³⁸. It enables efficient passive, chaotic mixing between different solutions under Stokes-flow regime. I integrated this device in our designs to achieve full mixing just before encapsulation of the multiple *upstream* channels carrying biochemical parts, in order to ensure homogeneous internal content, precise stoichiometry, and efficient encapsulation, which I reasoned could have been affected by laminar biochemical gradients and spatial anisotropy of concentrations. Synthetic biochemical circuits can then be spontaneously assembled just before encapsulation, by that mean standardizing the encapsulation mechanism and reducing its dependency on the nature of biochemical materials. Moreover, this design allows for fine tuning on stoichiometry via control on the input flow rates, which proved practical to test different parameters for straightforward prototyping of protosensors.

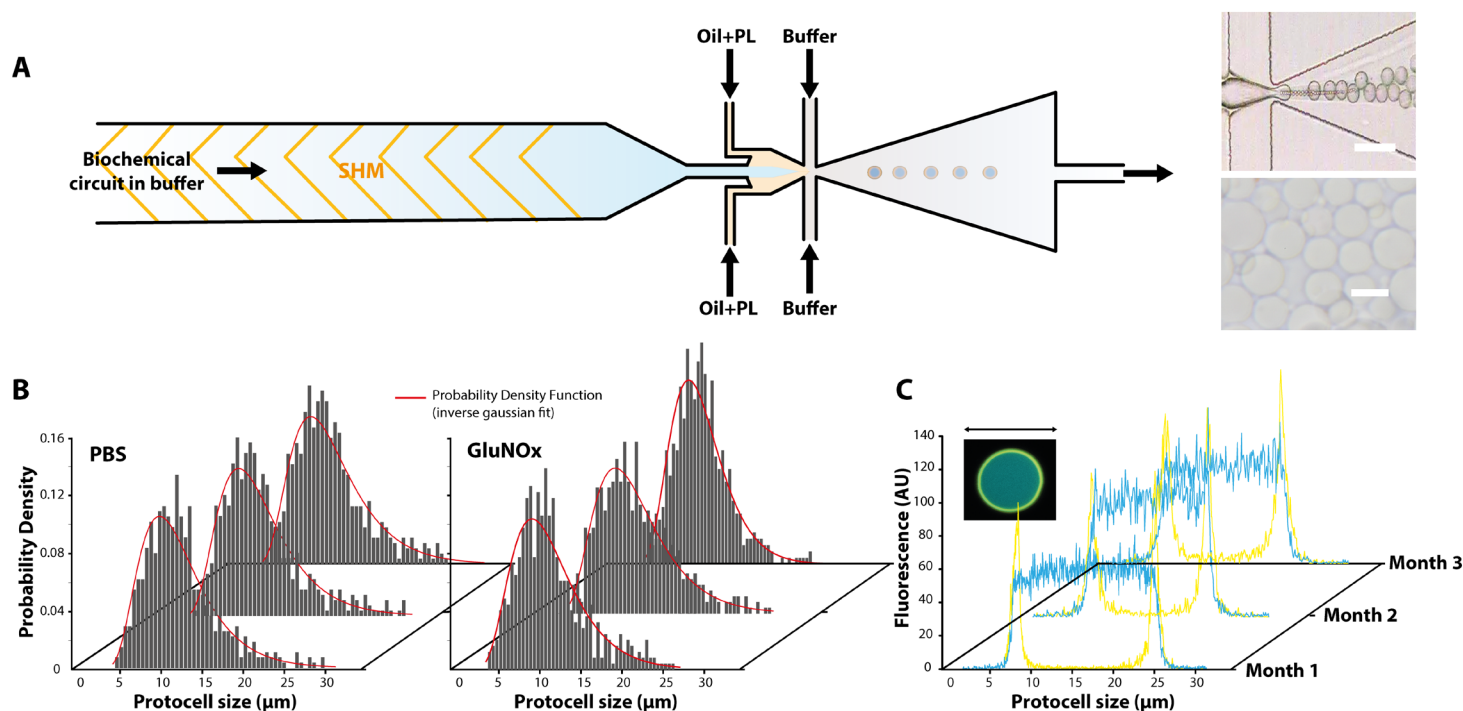


Figure 3.13: Experimental construction of protosensors using microfluidics. (A) Left: Double emulsion templating microfluidic device architecture and operation. This method relies on the generation of W/O/W double emulsions. Buffer: 10% v/v methanol, 15% w/v glycerol, 3% w/v pluronic F68 in PBS (1 $\mu\text{l}/\text{min}$) Oil+PL: DPPC dissolved in oleic acid (0.4 $\mu\text{l}/\text{min}$) Biochemical circuit buffer: enzymes and metabolites in PBS (0.4 $\mu\text{l}/\text{min}$). Right: microscopic validation of protosensor generation *on chip* (Top, scale bar= 20 μm), bright field optical validation of protosensors isolated for subsequent analysis (Bottom, scale bar=10 μm). (B) Size dispersion and stability of protosensors generated with the microfluidic method, after encapsulating PBS (left) and GluNOx network (right). (C) Months long stability of encapsulation. AlexaFluor-488 labeled IgG where encapsulated in protocoells and fluorescence was then monitored by confocal microscopy for three months. Yellow: protocoell membrane fluorescence as labeled with phospholipid dye DiI₁₈. Blue: AlexaFluor-488 labeled IgG fluorescence.

I analyzed size dispersion of protocells in order to characterize the fabrication process using light transmission microscopy (**Figure 3.13B**). At these flow rates we obtained fairly monodispersed protocell with average size of $\sim 10\mu\text{m}$, and an apparent inverse Gaussian distribution of the size parameter. Interestingly, circuit encapsulation showed no influence on size distribution of protocells, which demonstrate that the encapsulation process can be decoupled from the complexity of the biochemical content. Moreover, no evolution of the size parameter was recorded for 3 months, which demonstrated the absence of fusion events between protocells and a very satisfying stability in the conditions of our storage method (i.e. 4°C in storage buffer). I then assayed encapsulation stability using confocal microscopy. To this end, I encapsulated an irrelevant protein bearing a fluorescent label within protocells, and measured the evolution of internal fluorescence over the course of three months. I found that internal fluorescence remained stable, which demonstrates no measurable protein leakage through the protocell membrane after three months in our storage conditions.

Confocal microscopy also gave precious information on protocell's membrane characteristics. To visualize the membrane, I used a phospholipid bilayer specific dye (i.e. DiIC₁₈) which undergoes drastic increase in fluorescence quantum yield when specifically incorporated into bilayers⁶³⁹. I obtained well defined images denoting the complete extraction of oleic acid from the double emulsion and a well-structured arrangement of the bilayer (**Figure 3.13C** and **3.15A**).

Next, in order to validate the encapsulation of biological enzymatic parts inside protocells, I carried out UPLC-Mass spectrometry experiments. I encapsulated two relevant enzymes within protocells: Alcohol Oxidase and Glucose-1-Dehydrogenase. I then performed chromatographic and ESI mass spectrometry analysis on the protocells, and found that we could retrieve the molecular signatures of the enzymes in the interior of protocells, as compared with positive controls (**Figure S3.3**).

In addition to confocal and transmission light microscopy, we performed environmental electron scanning microscopy. This technique validated the structure, size and shape of the protocells and yielded esthetic images, but also gave an interesting way to probe and interact with the phospholipid bilayer (**Figure S3.1**).

Therefore, we successfully validated our microfluidic platform for the fabrication of protocells. This set-up proved capable of generating highly stable protocells with high efficiency and user-defined programmable content.

3.5 Construction and analytical evaluation of medical protosensors

3.5.1 *In silico* optimization of protosensor circuitry

In order to minimize modes of failure and obtain the most robust behavior, we first performed *in silico* optimization of protosensors before *in vitro* implementation. More specifically, the initial state concentration parameters of the species constituting the encapsulated synthetic circuit need to be optimized to take into account membrane selective permeability and give fastest results in accordance with the medical algorithm of interest. We thus incorporated membrane parameters in the models describing molecular inputs passively diffusing in and out through hemolysin pores (see **Materials and Methods** for details). We then defined temporal logic specifications that would best satisfy clinical requirements, that is, obtaining biosensing sensitivities at pathological thresholds, achieve specified signal processing operations, and obtain a measurable output signal in less than 10 minutes for the three systems (**Figure S3.7**, See **Materials and Methods** for details).

Using BIOCHAM, we first performed sensitivity analysis on the models to determine which concentration parameters had the most important influence on the systems' behavior. For each protosensor models, we could identify the two key biochemical species that would constitute the most sensitive components. We then computed 2D sensitivity landscape maps relative to these two dependencies in order to visualize the available biochemical design space for each system (**Figure S3.6**). We found that we could define concentration spaces within boundaries of which to implement desired temporal logic. GluONE protosensors function appeared mostly sensitive to G1DH and ADH enzymes concentration. Interestingly, we found that for LacOH and GluNO_x, their behavior was more sensitive to the initial concentrations of the metabolite NAD⁺ than other enzymes. In fact, for all three systems, the NAD⁺/NADH redox ratio can be seen as a biochemical *current* connecting the two molecular inputs signals, and thus has to be finely tuned to match input thresholds and enzyme levels.

Last but not least, within this computed design space, initial state concentrations were rationally chosen using BIOCHAM automated parameter search (CMAES Method). This approach enables to achieve optimized robustness of operation while satisfying temporal logic specifications according to each model (**Figure S3.7**). The concentrations we obtained were then used to experimentally build the three synthetic circuits embedded in protosensors.

3.5.2 Digital signal processing and multiplexing logic

In order to verify Protocells behavior, we started by mapping *in silico* their complete transfer function using stochastic HSIM simulations (**Figure 3.14A**). As previously, I plotted heat maps of computed outputs signal after induction with various concentrations of biomarker inputs, using calibrated mathematical relations between molecular concentration and measured signal (**Figure S3.8**). I found that Boolean logic was respected with very satisfying theoretical response fold change, as well as near-digital, sharp response profiles. I also found that theoretical switching thresholds for these models matched useful clinical sensitivity for biomarker inputs (pathological thresholds: Ketones > 17 μM ($\sim 10\text{mg/dl}$); Glucose > 1.39 mM ($\sim 25\text{mg/dl}$); Lactate > 10 μM ; EtOH > 17.4 mM ($\sim 80\text{mg/dl}$); NOx > 1000 μM).

The next goal was then to investigate the behavior of protosensors *in vitro*. Therefore, I proceeded to microfluidic fabrication of GluONE, LacOH and GluNOx protocells using optimized concentration parameters as previously defined. In a first experiment, I reasoned that a preliminary exploration of models validity would be to achieve the same transfer functions *in vitro* as previously predicted by simulations. I thus exposed and incubated the three protosensors systems to increasing concentrations of respective input biomarkers, and measured their individual output signal response using flow cytometry. I hypothesized that this technique would give most precise measurements by cancelling sample noise effects in order to get finer verification of protosensors behavior at the single (*proto*)cell level (**Figure 3.14B**). Interestingly, when comparing these data to HSIM model simulations, I found that the experimentally measured switching thresholds using this technique showed very good agreement with predictions, along with near-digital responses. Although this approach alone is not sufficient to map the complete Logic behavior in response to the combinations of different inputs, which would require extensive experimental sampling, it is a preliminary unidimensional validation of useful analytical properties of these systems. In addition, we have yet to measure the *in vitro* responses for output 1, which is the focus of ongoing experimental work.

I then sought to further visualize the spatial and analytical digitization of output signals. I used confocal microscopy to quantitatively measure and precisely visualize output signals generation in induced protocells (**Figure 3.15A**). I obtained bright images with high SNRs and important response fold changes. Molecular output signals appeared well localized to the interior of protosensors, although I did not quantify possible leakage. These experimental data strongly corroborate previously flow cytometry acquired data.

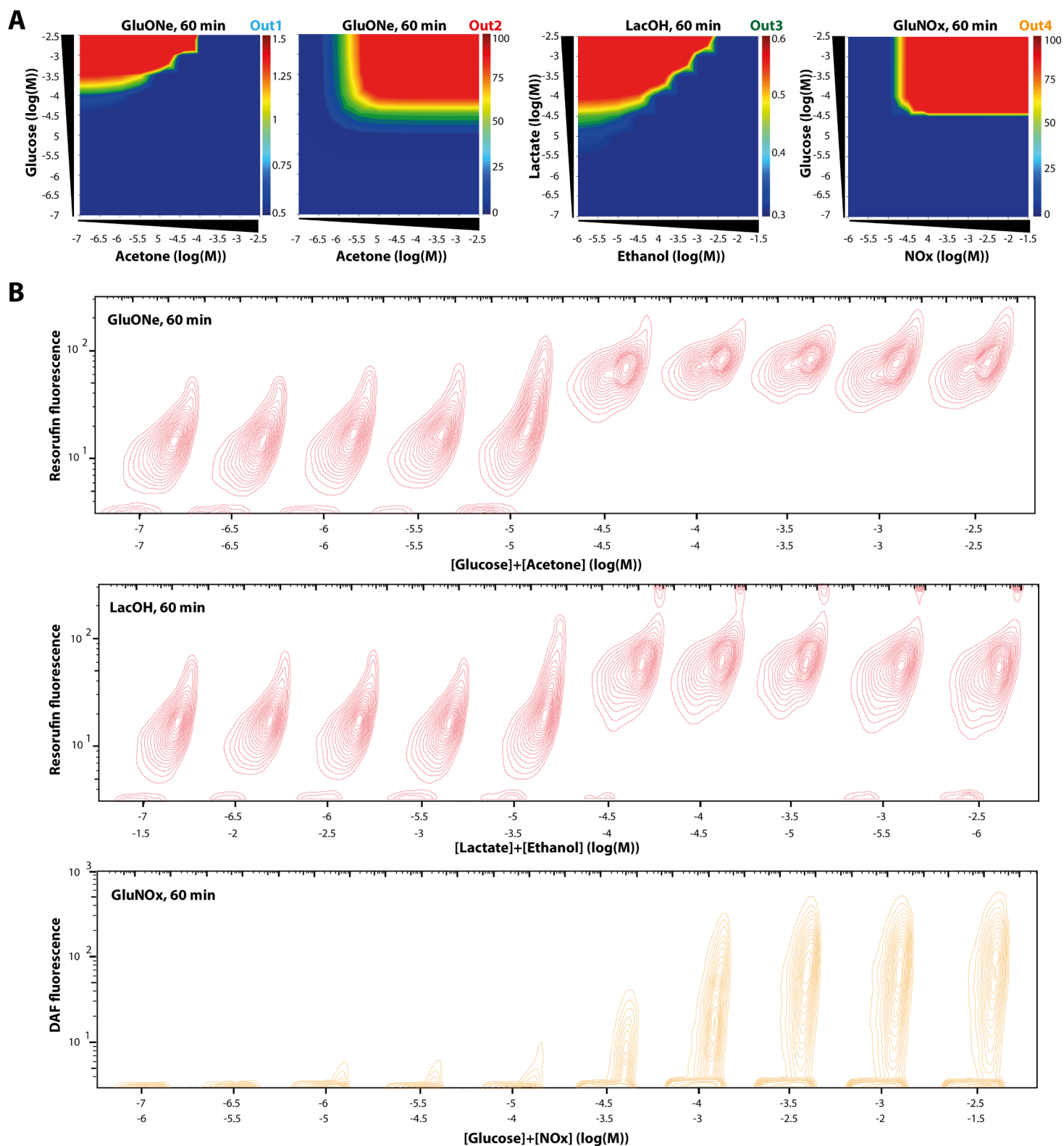


Figure 3.14: *In silico* and experimental validation of analytical properties of protosensors. (A) Heat maps depicting computed output signals at 60 minutes using HSIM models simulations, for the 3 different systems after induction with increasing biomarker concentration (B) Experimental validation of computer prediction at the single (proto)cell level using flow cytometry. For LacOH, in order to get a fluorescent output signal measurable with a flow cytometer, we exchanged ABTS with Resorufin, which are analogous.

I then performed multiple experiments consisting in mapping the experimental truth tables of the three protosensor systems. I measured output responses at the population level, while varying input conditions (i.e. presence/absence of pathological concentrations of input biomarkers). This allowed us to get a fine understanding of the functioning and detailed experimental characterization of logic operations (**Figure 3.15A, 3.15B and S3.11**). I obtained clear digital-like behaviors with important fold changes and exact accordance to Boolean logic specifications with temporal requirements of less than 60 minutes (**Figure S3.11**). I calculated signal to noise ratio, which showed very good performance SNRs for outputs 1, 2, 3 and 4 of ~8, 35, 5, and 11 dB respectively.

Even though very satisfying for analytical applications, we found greater background noise compared to non-encapsulated synthetic biochemical circuits. I hypothesized that it was due to the introduction of auto-fluorescent species such as surfactant and phospholipids, as well as probable scattering and absorbance phenomenon emerging for spherical protocellular structures in solution.

In fact, the rationale behind encapsulation was to achieve greater analytical robustness and obtain insulation from context, while achieving true composable and multiplexed Boolean logic. In other words, multiple types of protosensors in solution should be able to operate independently in a standardized way without interacting with each other, this way achieving multiplexed analysis of the molecular environment. To verify this assertion experimentally, I first addressed multiplexing logic capabilities. For this purpose, I set up different experiments where I measured output signals in media spiked with multiple biomarkers, using all three synthetic biochemical circuits in either batch mode analysis (i.e. non-encapsulated synthetic circuits) or protosensors analysis (i.e. protocell encapsulated synthetic circuits). I compared the measured output signals obtained to the expected theoretical *true* outputs. I found that simply mixed synthetic biochemical circuits were incapable of achieving correct signal processing tasks, probably due to molecular interactions between circuits' components. For this experiment, I rationally chose example combinations of biomarkers that would be most likely *wrong*, although different combination could have given *true* behavior. On the other hand, mixing the three types of protosensors did not affect biosensing and signal processing capabilities, which were capable of coordinating *true* Boolean logic and output signal generation (**Figure 3.15C**). Although I did not yet test all the 32 possible input combination that would require extensive experimental work, the data shown here suffices to demonstrate the interest of insulating signal processing biochemical circuits within protocells, thereby increasing the scale in parallelization of biocomputation tasks operating simultaneously.

I then sought to address the potential effects of clinical urine media on protocell structure and operability using flow cytometry (**Figure 3.15D**). I reasoned that measuring protocell fluorescence signal along with forward scattered light (FSC) would give insights on protosensor stability, as FSC is correlated with vesicular size, and internal fluorescence with membrane integrity. I found that induction and prolonged incubation in urine does not impact stability, structure or inducibility of the system. In addition, I found no difference between operation in standard PBS buffer and urine media (**Figure 3.15B and Figure S3.12**).

Taken together, these data demonstrate that protosensors enable the implementation of programmed biosensing and biomolecular logic gated operations displaying robust and predictable behavior *in situ*.

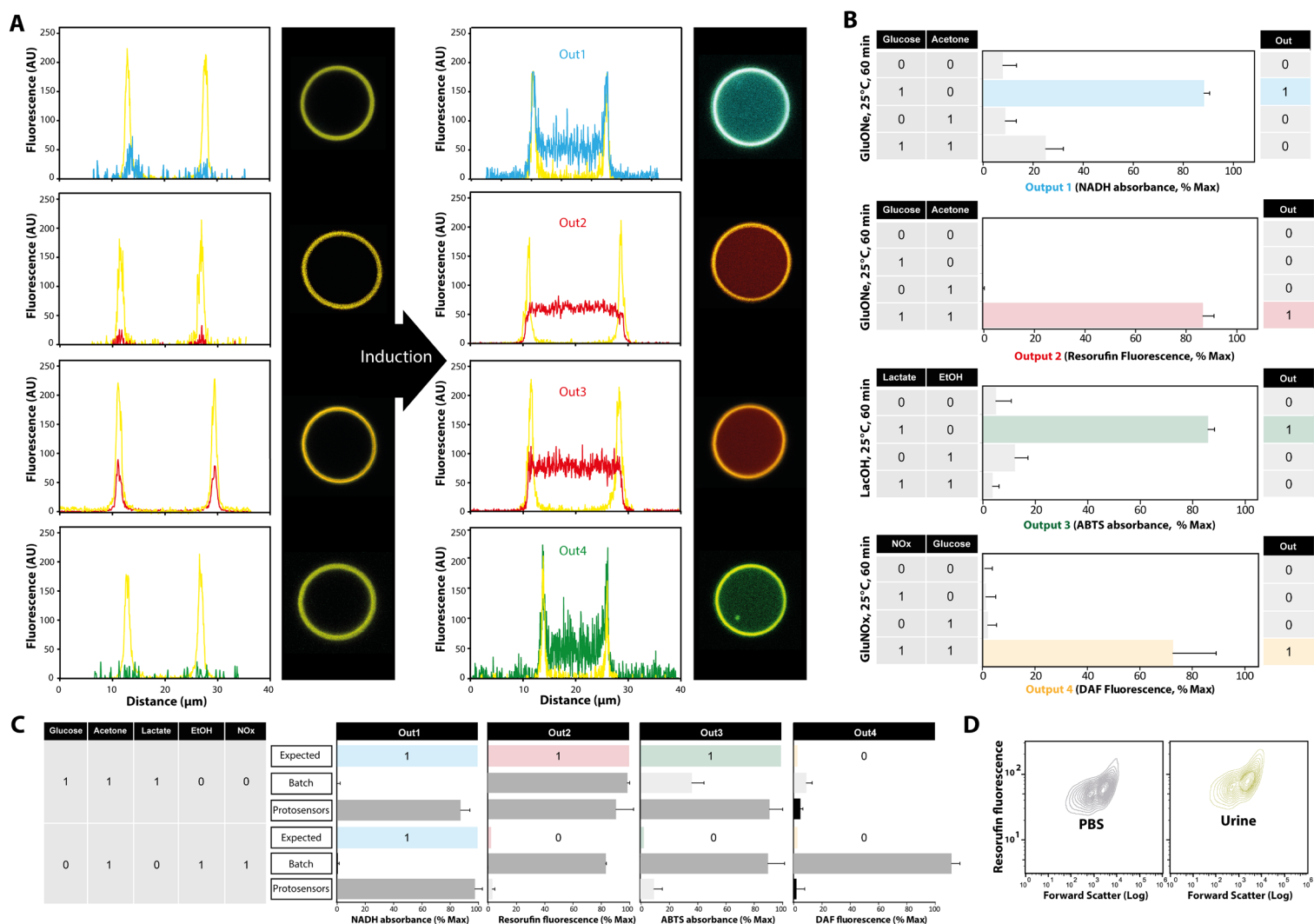


Figure 3.15: Experimental validation of medical protosensors robust multiplexed biosensing and logic. (A) Confocal microscopy validation of “ON” output signals responses at 60 minutes after induction with respective biomarkers. From top to bottom: GluONe (Out1, Out2), LacOH (Out3), GluNOx (Out4). The phospholipid bilayer is stained in yellow using the dye DiC₁₈. For LacOH, in order to get a fluorescent output signal measurable in confocal microscopy we exchanged ABTS with Resorufin, which are analogous. (B) Experimental truth tables of protosensors operating in human urines. (C) Multiplexing Logic: example of comparison between expected (valid analytical response according to molecular inputs present in the sample), batch mode analysis (non-encapsulated synthetic circuits) and protosensors analysis (encapsulated synthetic circuits). PBS media was spiked with multiple biomarkers (in this case, acetone, ethanol and nitric oxides or glucose, acetone and lactate), and output signals were measured at 60 minutes (D) Flow cytometry evaluation of protosensors structural robustness in urines. GluONe protosensors were induced and incubated for 2 hours at 25°C in human urine, and analyzed using flow cytometry while recording Resorufin fluorescence and forward scattered light.

3.5.3 Assaying pathological clinical samples: protosensors mediated diagnosis of Diabetes

After successful characterization of protosensors analytical capabilities in spiked samples, we then proposed to perform real world diagnostic evaluations. To assess the relevance of protosensors for disease detection in a clinical assay, we sought to evaluate a proof-of-concept that could detect endogenous levels of a pathological biomarker in clinical samples from patients. Although we did not benefit from a large sample library of diabetes related metabolic complication to test the complete implemented diagnostic algorithm in detail, we disposed from previously collected urine samples from naive diabetic patients. These patients presented simple glycosuria with negative ketonuria as confirmed with a urinary dipstick. Therefore, I reasoned that assaying pathological glycosuria in these urine samples would be a good simple testbed evaluation for GluONE protosensors and would provide with a interesting diagnostic evaluation (i.e this satisfies the GluONE algorithm Output 1=Glucose AND NOT Acetone).

I proceeded to incubation of GluONE protosensors with either diabetic urine samples or non-diabetic control urines, and as previously described measured output signal responses (**Figure 3.16**). I also concomitantly performed glycosuria analysis using the clinical gold standard urinary dipsticks. I found very good correlation between output signals from protosensors and visual examination of dipstick. Moreover, we performed Receiver Operating Characteristic (ROC) analysis on these data and found that the assay reliably detected glycosuria in samples from diabetic patients, with a near ~100% sensitivity and specificity, and an Area Under Curve of ~0.9981, which defined GluONE protosensors as an excellent diagnostic test. Together, our data demonstrated that protosensors can discriminate between normal and diabetic patients with excellent diagnostic accuracy. Therefore, I concluded that rational biomolecular programming of protosensors can be used to generate clinical grade assays to detect endogenous biomarkers of disease in patient samples.

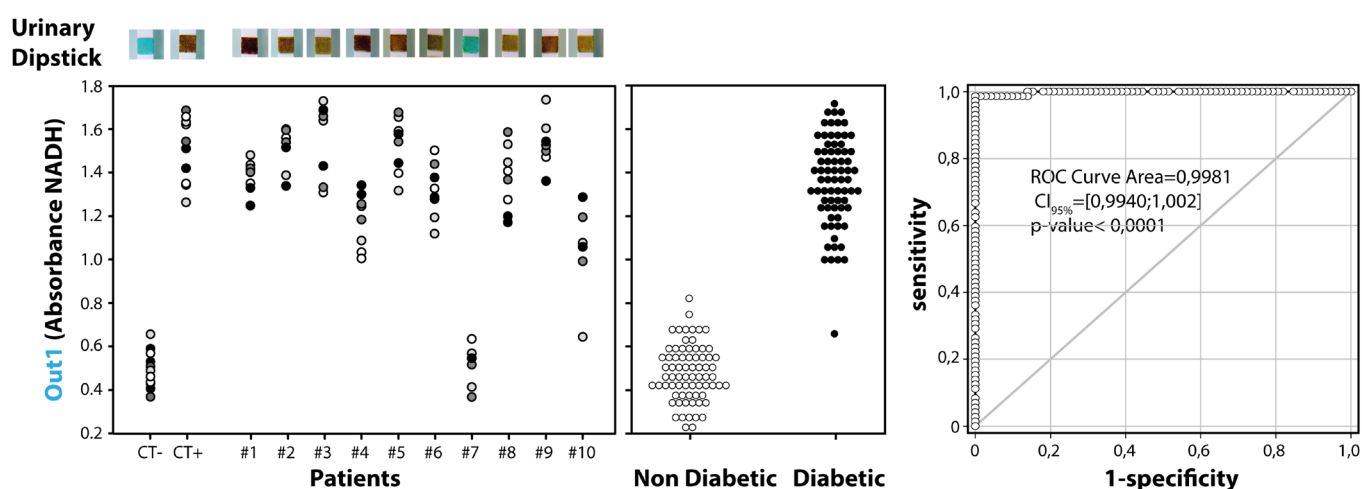


Figure 3.16: Protosensor mediated detection of pathological glycosuria in patient clinical samples. Left: Dot histogram of data used for statistical analysis. Right: Receiver Operating Characteristic (ROC) analysis curve depicting statistic sensitivity versus (1-specificity). A set of 72 measurement performed in non-pathological urine were compared to 72 measurements performed in urine from diabetic patients. I used GluONE protosensors for this assay, and measured Out1 signal (NADH absorbance) after 60 minutes.

3.6 Conclusion and discussion

This study demonstrated that protosensors are highly promising tools to perform multiplexed *in vitro* diagnostics integrating medically relevant algorithmic processes. As a prototype for the clinics, we showed that this technology could be successfully applied to solve real clinical problems and demonstrated that protocells could overcome several hurdles faced by classical *in vitro* diagnostics. Although portable, it also offers multiplexed detection and complex analytic capabilities with sharp near-digital response profiles at tunable thresholds, coupled to expert decision making.

We brought rational programming of synthetic biological circuit closer to real world application by addressing some of the previous technological limitations, namely design and scalability. We proposed a systematic computational approach that combined a directed exploration of the design space according to time dependent quantitative and qualitative specifications, to automated robustness optimization of initial parameters for experimental implementation. This ensures that robustness in operation and functionality are maximized and led to successful automated design and construction of synthetic biochemical circuits from vast and comprehensive enzyme databases. Last, we provided a quantitative *in silico* framework to evaluate the function and analytical properties of protosensors.

We showed that by confining biological complexity within membrane boundaries, one can achieve more complex circuitry by preventing deleterious molecular *short-circuits*. Moreover, protocell membrane confinement provided a robust architecture to insulate and decouple the biochemical *software* from crosscoupling with the complex medium it operates in. Complex and susceptible enzymatic activities can thus be confined within protosensors as standardized and insulated sensing and computing units.

Along with computer simulation and assisted design, we demonstrated that these devices could be synthesized using a straightforward, versatile and scalable methodology relying on microfluidics. We envision that the capabilities brought by microfluidics to bottom-up design of protosensors will help bridge the gap towards effective translation to the clinic.

Although the experimental implementation of protosensors seemed straightforward in this study, this may unfortunately not be the case for more complex circuits. Indeed, complex and unpredictable quantitative molecular interactions often govern the dynamics of biological matter, in a way that man-made predictive mathematical models can fail due to over-simplification. Most biological parts available for the bottom up engineering of synthetic biochemical circuits remain poorly characterized in context, and combination of species can lead to biological discrepancies. Moreover, a lack of distinct suitable parts may limit the scaling up of biochemical biosensing/biocomputing devices. Some input biomarkers may in certain cases lack robust biochemical sensors interfacing with computation modules. Additionally, incomplete knowledge of molecular reactivity and physicochemical properties of species, as well as unpredictable stochastic noise arising at the network level may also hinder computer assisted construction frameworks. However, there is no reason to assert that these hurdles are insurmountable. For instance, instead of systematically capitalizing on the huge number of catalytic activities that can

be exploited, it could be interesting to investigate in the future how fully orthogonal biochemical parts could be constructed to operate in highly precise ways. For instance, computational protein design methods have been capable of creating enzymes with novel non-natural catalytic functions⁶⁴⁰.

In order to overcome these limits, it will also be necessary to improve the reliable classification, standardization, and robustness assessment of biological parts. This initiative is underway through CompubioticDB database, which stores and is fed with useful and abstract biological parts implemented with proteins and metabolites, as well as modules such as sensors, switches, or logic gates. Constant refinement and quantitative improvement of the database should lead in the future to geometrical enhancement in bottom-up design capabilities. Future work will focus on developing automated tools to feed CompubioticDB with robust and well characterized biochemical parts to reduce the errors of *in silico* and *in vitro* implementation processes. Although design still requires some human expertise and verifications at this point, future work will focus on accelerating the transition from human assisted error-prone process to full automation.

In addition, efforts will be directed towards augmenting the models' depth of description from basic Michaelian kinetics to molecular mechanistic, using for instance prior work conducted in the laboratory on Basic Elementary Actions (BEA) within the biological function description scheme (Bio Ψ)^{641 642}.

The bottom-up approach I applied here to synthesizing protocell models using phospholipid bilayer membranes may impose some intrinsic limits. Indeed, I have yet to explore the consequences of osmotic pressure on the systems, or how to finely tune selective exchanges through the membrane. In addition, I have yet to determine how transmembrane signal transduction or trafficking may be achieved. This will necessitate a wide variety of discipline, from soft matter physics, and biochemistry, to chemistry to construct stable multifunctional compartments. We are thus also currently working on novel encapsulation strategies, wishing to evolve compartments towards greater orthogonality. For instance, I envision that protein nanocages, nanogels, polymersomes, could provide good alternatives for future designs. Moreover, valuable alternatives to enzymes as substrate to biosensing and computing could be found in nucleic acids and their orthogonal derivatives, since they provide with extended capabilities, programmability and versatility. In addition, while output multiplexing as shown in this study will reach intrinsic limits related to the number molecular signals that can be used simultaneously, the problem could be solved through spatial patterning for example. In fact, protosensors could be amenable to ultrahigh throughput applications through high density *on chip* spatial patterning.

This study paved the way for the development of integrated biochemical circuits capable of sensing their molecular environment, achieving biomolecular signal processing and decision making at the molecular and cell scale. I suspect that the approach to biosensing described in this study, which relies on autonomous and programmable entities, to be of great interest for novel kinds of local measurements and bioactuation *in situ* since protosensors can be addressed to specific biological structures or cells through external receptors. In addition, these systems could be engineered into *sense-act* micromachines where the systems would conditionally generate cell

actuation signals or therapeutic responses *in situ*, as well as integrated or interfacing living systems. They could also be used for high-throughput screening of complex phenotypes or biological functionalities.

A vast landscape of open problems in biology and medicine has remained unsolved due to our inability to grasp biological networks dynamics and process information at the molecular scale. Gaining insights on this phenomenon through a bottom-up and systematic approach to control and design molecular programs acting within biological substrate and interfacing organisms, could be of outstanding interest and bring progress for both basic and applied sciences.

The last decade witnessed a growing interest for the study of the modalities with which biological systems process information and solve computational problems, largely inspired by applications pervasiveness of portable, autonomous, and programmable sensor devices. A top-down approach has traditionally been favored due to an increasing ability to experimentally probe the molecular mechanisms of living organisms using a variety of new technologies. However, the development of novel frameworks to build from the bottom-up orthogonal structures could provide unique opportunities to model, analyze, and unravel biological systems, bring new insights on the fundamental principles governing biological information processing, and open up the way for tremendous biomedical applications. In a long-term vision, I envision establishing protocell and bottom-up biomolecular programming as a universal framework to produce a wide array of tools for research, from probing and interfacing biological structures to cellular reprogramming and microscale biomolecular computing.

3.7 Materials and methods

Study Design:

To evaluate the robustness of our system and its functionality in clinical samples, we used urine pools from healthy individuals as well as urine samples from healthy individuals and diabetic patients. Regarding collection of clinical samples, non-pathological (control) and glycosuric (diabetic) urine samples were obtained from the Department of Endocrinology of the Lapeyronie Hospital, Montpellier, France, under the supervision of E. Renard. Individual informed consents were obtained from the patients and control individuals. Glycosuric urine samples were collected from 10 newly discovered, non-stabilized diabetic patients. Urine samples were stored at -80°C before use.

Protocell microfluidic construction

PDMS Microfluidic chips were designed and prototyped using AutoCAD software and fabrication was carried out by the Stanford University microfluidic foundry. The microfluidic chip was connected with PTFE tubing to neMESYS V2 syringe pumps (Cetoni GmbH, Germany). Microfluidic processes were imaged using a Leica DMIL microscope mounted with a Canon 750D or a Phantom V7.3 ultrafast camera. Detailed information can be found in SI.

Spectrometric assays, flow cytometry and microscopy

To test the operability of the different systems, synthetic biochemical networks and protosensors were inoculated in 100 μl total volume of PBS with or without inducer, or urine from patients diluted at a ratio of 1:2 in PBS for a total volume of 100 μl in a 96-well plate. If not differently specified, After 1 hour of incubation, fluorescence and absorbance was read using a synergy H1 plate reader. We concomitantly tested urine clinical samples from non-stabilized diabetic patients using the Siemens Multistix 8 SG reagent strip according to the manufacturer's protocol. Flow cytometry experiments were performed on a Guava EasyCyte bench top flow cytometer (Merck) equipped with a 488nm laser, and analyzed using FlowJo vX software. Confocal microscopic assays were performed on a Leica SP8-UV equipped with 63x oil lens and 355nm, 488, and 561nm lasers.

Data analysis and statistics

Experimental values are reported as means \pm SD. All experiments were performed in triplicates. Data, statistics, graphs, and tables were processed and generated using MATLAB (MathWorks) and SigmaPlot (Systat Software Inc.). For receiver operating characteristic analysis, a set of 72 measurements performed in non-pathological urine were compared to 72 measurements performed in urine containing 1% w/v glucose.

3.8. Supplementary materials

3.8.1 Microfluidics and protosensor preparation and characterization

PDMS microfluidic chips were designed and prototyped using AutoCAD software and fabrication was carried out by the Stanford University microfluidic foundry. 20 Gauge holes were punched in the PDMS chip, allowing the use of customized made stainless steel adapters (New England Small Tube) for PTFE tubing connections (1/16 OD, 0.8 mm ID). The flow in microfluidic channel was controlled via displacement driven flow using Cetoni neMESYS syringe pumps equipped with high precision glass syringes. I found strong dependence of protosensors yields of production on flow rates, that were kept at 1/0.4/0.4 $\mu\text{l}/\text{min}$ (A/B/C) to achieve best encapsulation efficiency. Movies taken with an ultrafast camera (Phantom v7.3) mounted on a LEICA DMIL inverted microscope at $\sim 20\,000$ FPS allows us to estimate around ~ 1500 Hz the frequency of protosensor generation at these flow rates.

I introduced in the microfluidic design a device previously described, known as the staggered herringbone mixer (SHM)⁶³⁸. It enables efficient passive, chaotic mixing between different solutions under Stokes-flow regime. We introduced two times more cycles (i.e. 10), as we calculated that ~ 5 cycles were sufficient to achieve efficient mixing, according to the equations provided in Williams *et al.*⁶³⁸. I integrated this device in the designs to achieve full mixing of the two C channels carrying biochemical species just before encapsulation, in order to ensure homogeneous internal content, precise stoichiometry, and efficient encapsulation, which we reasoned could have been affected by laminar biochemical gradients and spatial anisotropy of concentrations. Synthetic biochemical circuits can then be spontaneously assembled just before encapsulation, by that mean standardizing the encapsulation mechanism and reducing its dependency on the nature of biochemical materials. Moreover, this design allows for fine tuning on stoichiometry via control on the input flow rates, which proved practical to test different parameters for straightforward prototyping of protosensors.

Our strategy relies on a microfluidic flow-focusing droplet generation design that generates water-in-oil-in-water (W–O–W: Biochemical circuit in PBS - Phospholipid in Oleic acid – Storage Buffer A) double emulsions. Double emulsion templates are generated in described flow-focusing channel geometries. DPPC phospholipid membranes then self-assemble during a controlled solvent extraction process (Oleic acid is extracted by methanol present in Storage buffer A). The rationale we use for choosing DPPC concentration in oleic acid is adapted from Teh *et al.*⁶²⁸. Briefly, oleic acid solution was composed of 1.1mM DPPC. This concentration was chosen so that there would be a sufficient number of phospholipids to form a lipid bilayer around a 10 μm diameter vesicle. The average area per molecule of DPPC in a bilayer membrane is estimated around a value of 0.64 nm²⁶⁴³. Considering that a 10 μm liposome would have a lipid area of 3.14 $\cdot 10^{-10}$ m², at least 8.15 $\cdot 10^{-16}$ moles of DPPC would be required to compose the full lipid bilayer of a liposome. Assuming the maximum thickness of oleic acid contained by the primary

double emulsion to be 5 μ m, we can calculate the volume of oleic acid to be 3.665 pL. A 1.33 mM DPPC concentration would be sufficient, as is it 5 times the amount of phospholipid needed to form a lipid bilayer around the vesicle. In addition, we briefly investigated DOPC, DMPC and DPPC phospholipids for protocell fabrication, and found that DPPC achieved better apparent stability and superior production yields.

In order to achieve selective biomarker input entry and matter/information exchange between protocell content and exterior medium, we capitalized on passive pore forming bacterial protein *α* -hemolysin. *α* -hemolysin pores have several properties that identifies it as a robust transmembrane channel suitable for biosensing applications. *α* -hemolysin pores are self-assembled in the membrane and do not require specific assembly conditions, they are stable over a wide range of pH and temperature and are open in normal conditions. The transmembrane pore of *α* -hemolysin operates the delivery of ions and small organic compounds such as sugars, metabolites or nucleotides in a selective way through the walls of synthetic lipid vesicles⁶⁴⁴ with a passive diffusion rate of $5.5 \pm 1.5 \times 10^{-4} \text{ s}^{-1}$ as previously measured⁶⁴⁵.

The output channel of the microfluidic chip containing newly formed protosensors was connected with PTFE tubing to a collection vial containing buffer A kept on ice. After 5 hours of fabrication which also allows methanol to evaporate, we obtained ~ 1 ml of protosensor (by encapsulated volume). The final solution containing protosensors corresponding to a $\sim 1:5$ final dilution in buffer A could then be stored at 4°C for a maximum time of 1 week before further use.

For measurements, Hemolysin treatment of protosensor was performed 15 minutes prior to induction. Hemolysin was added for a final concentration of 1 μ M. Protocells solutions were then back diluted at a 1:1 proportion into the sample to test (e.g. Urine). Induction was carried out under slow agitation at 25°C. All fluorescence and absorbance measurements were performed on a synergy H1 plate reader, in 100 μ l p96 microwells.

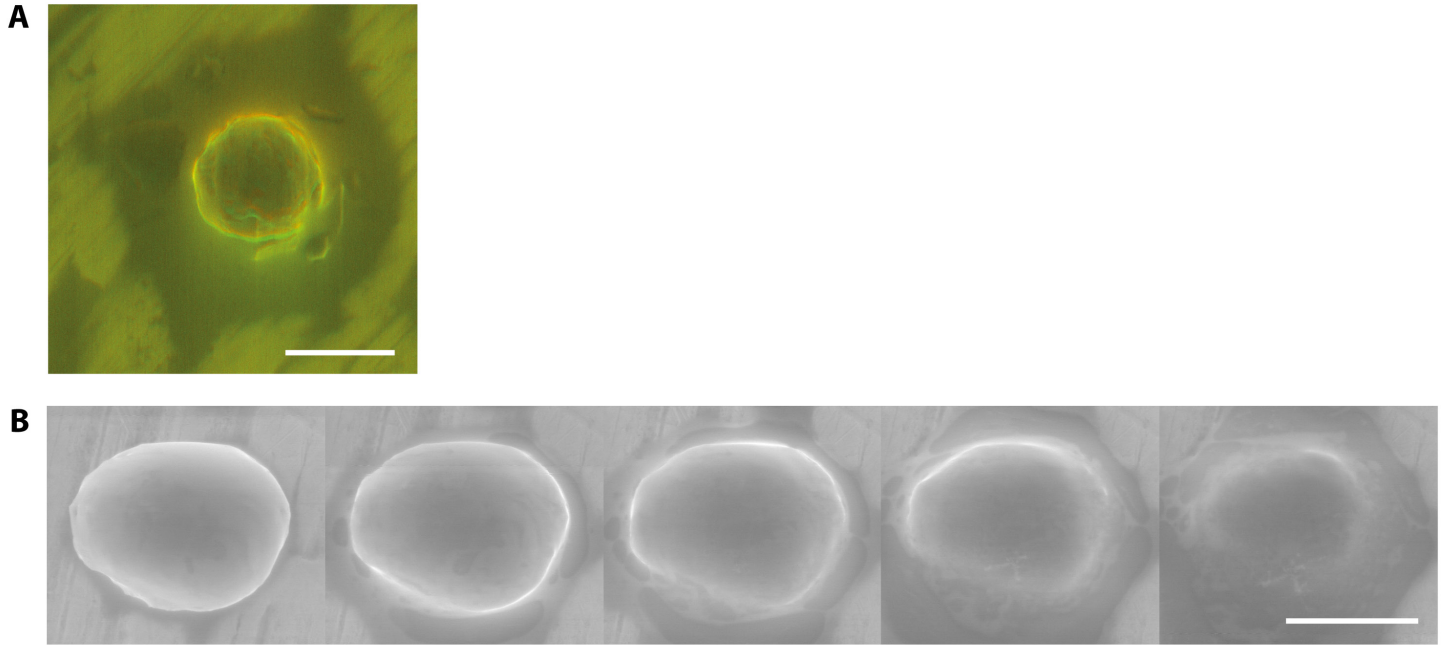


Figure S3.1: Environmental Scanning Electron Microscopy (ESEM) photomicrograph of protosensors. DPPC vesicles were fixed overnight at 4°C in 2.5% glutaraldehyde solution in PBS, and then washed with water prior to direct observation. **(A)** Stereoscopic micrograph of individual protosensor **(B)** Kinetic visualization of electron beam interacting with a protosensor. From left to right: 10 seconds were sufficient to melt the DPPC bilayer and release intra-vesicular content. (Scale bar=10µm)

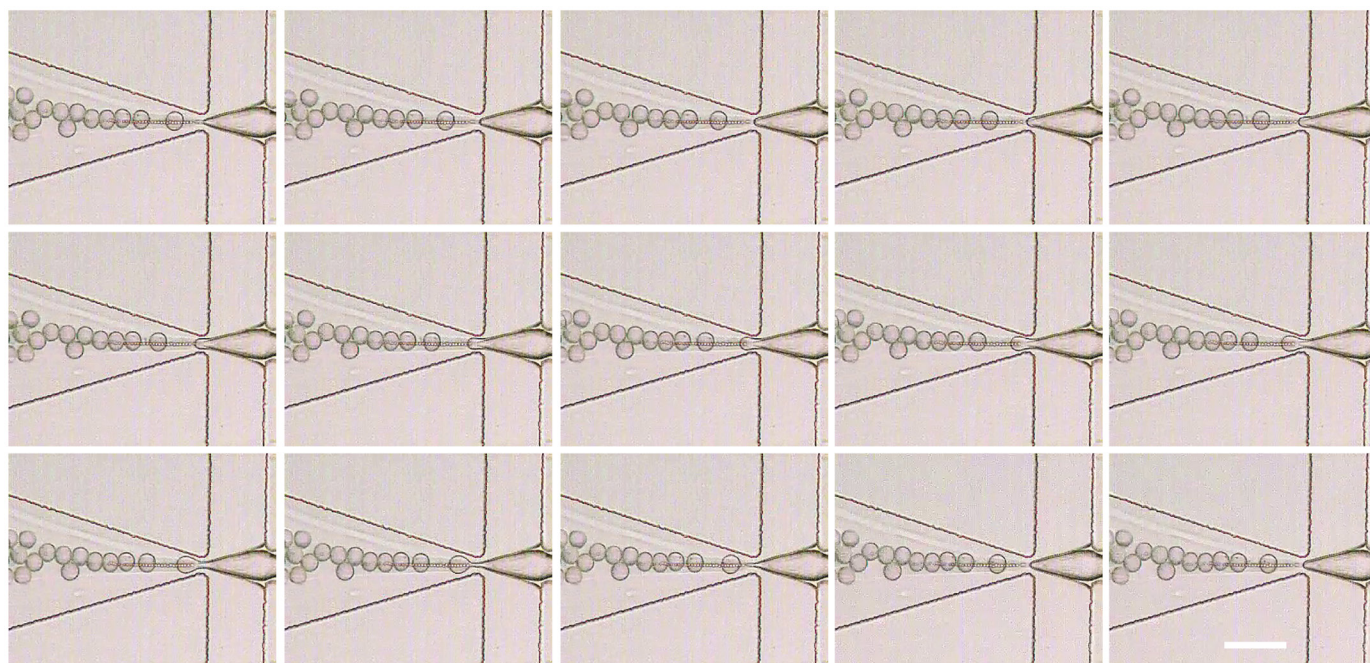


Figure S3.2: Timelapse photomicrograph of protosensors fabrication within microchannels. (Read from left to right, top to bottom). This movie was recorded at 20 000 FPS and corresponds to $\sim 0.5 \mu\text{s}$. (Scale bar= $40 \mu\text{m}$)

Molecule designation	Stock solution	Concentration of use (batch/protocells)		
		GluONE	LacOH	GluNOx
NAD+	50 mM PBS	250 μM / 4 mM	250 μM / 5 mM	100 μM / 2 mM
NADH	50 mM PBS	-	-	-
Acetone	100 mM PBS	1 mM (unless specified)	-	-
Ethanol	100 mM PBS	-	20 mM (unless specified)	-
Glucose	50 mM PBS	1 mM (unless specified)	-	5 mM (unless specified)
NO ₃	50 mM PBS	-	-	5 mM (unless specified)
Lactate	100 mM PBS	-	0.5 mM (unless specified)	-
Isopropyl alcohol	100 mM PBS	-	-	-
G1DH	3.4 U/ μl PBS	8.5 U/ml= $0.354 \mu\text{M}$ / 7.4 μM	-	138 U/ml= $5.7546 \mu\text{M}$ / 127 μM
ADH	55.7 U/ml PBS	0.2785 U/ml= $14.06 \mu\text{M}$ / 221 μM	0.2785 U/ml= $14.06 \mu\text{M}$ / 317 μM	-
AO	0.1 U/ μl PBS	0.75 U/ml= $0.02725 \mu\text{M}$ / 0.59 μM	-	-
NR	1U/ml PBS	-	-	4.2 μM = 0.5 U/ml / 92 μM
HRP	10 U/ml PBS	0.015 U/ml= $0.00104 \mu\text{M}$ / 0.0208 μM	0.05 U/ml= $0.00347 \mu\text{M}$ / 0.0754 μM	-
LO	6.425 U/ μl PBS	-	0.1 U/ml= $1.12 \mu\text{M}$ / 23.4 μM	-
Hemolysin	250 μM PBS	1 μM	1 μM	1 μM
Resazurin	10 mM water	50 μM / 1 mM	-	-
ABTS	10 mM PBS	-	100 μM / 2 mM	-
DAF-2	5 mM DMSO	-	-	10 μM / 200 μM

Table S1: Stock solutions and concentrations used in this study. All chemicals and enzymes were purchased from Sigma Aldrich. G1DH: Glucose-1-dehydrogenase, ADH: Alcohol dehydrogenase, AO: Alcohol oxidase, NR: Nitrate/Nitrite reductase, HRP: Horseradish peroxidase, LO: Lactate oxidase, ABTS: 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid), DAF-2: 4,5-Diaminofluorescein. See Table S2 below for more information about enzymes used in this study. Stock solutions were kept at -30°C until use. Resorufin and ABTS solution were prepared the same day of the assays or kept no longer than a week.

Enzyme/reaction	Organisms	Substrate 1	Km (mM)	Substrate 2	Km (mM)	Kcat1 (/s)	Kcat2 (/s)	Source	Mechanism
Nitrate Reductase (nitrate-> nitrite) EC 1.7.1.1	<i>Arabidopsis thaliana</i>	NADH	0.004	NO ₃ ⁻	0.015	-	210	BRENDA	Random bi-bi
Nitrate Reductase (nitrite-> NO) EC 1.7.1.1	<i>Arabidopsis thaliana</i>	NADH	0.004	NO ₂ ⁻	0.0074	-	2	^{646, 647}	
Glucose 1-Dehydrogenase EC 1.1.1.47	<i>Pseudomonas</i> sp.	NAD ⁺	80	Glucose	0.86	200	400	BRENDA	Ordered bi-bi
Alcohol Dehydrogenase EC 1.1.1.1	<i>Equus Caballus</i>	NADH	0.0025	Acetone / acetaldehyde	135 / 6	0.717	0.33 / 31.8	BRENDA	Ordered bi-bi
Alcohol Dehydrogenase EC 1.1.1.1 reverse	<i>Equus Caballus</i>	NAD ⁺	0.34	Isopropanol	268	0.41	0.75		
Alcohol Oxidase EC 1.1.3.13	<i>Candida</i> sp.	isopropanol	10	-	-	150	-	BRENDA	Ping-pong
Horseradish peroxidase 1.11.1.7	<i>Armoracia rusticana</i>	Amplex Red	0.081	H ₂ O ₂	0.005	240	-	BRENDA	Ping-pong
		ABTS	0.18	H ₂ O ₂	0.005	760	-		
		NADH	0.012	H ₂ O ₂	0.005	0.009	-	^{648, 649, 650}	
Lactate Oxidase 1.13.12.4	<i>Pediococcus</i> sp.	(S)-Lactate	0.2	-	-	-	283.3	⁶⁵¹ , BRENDA	Ping-pong
Alcohol Dehydrogenase EC 1.1.1.1	<i>Equus Caballus</i>	NAD	0.0074-0.01	Ethanol	2.46	308	-	BRENDA	Ordered bi-bi
NO decay	N.A.	NO	-	-	-	1.9.10 ⁻³	-	⁶⁵²	N.A.
NO reaction	N.A.	O ₂	0.001	-	-	2	-	^{653, 647, 654, 655, 646, 656, 657}	N.A.
		NO ₂	0.001	-	-	2000	-		
		DAF-2	0.001	-	-	6.28	-		
N ₂ O ₃ reaction	N.A.	DAF-2	0.001	-	-	2000	-		

Table S2: Biochemical parts and associated kinetic parameters used in this study to build biochemical circuits.

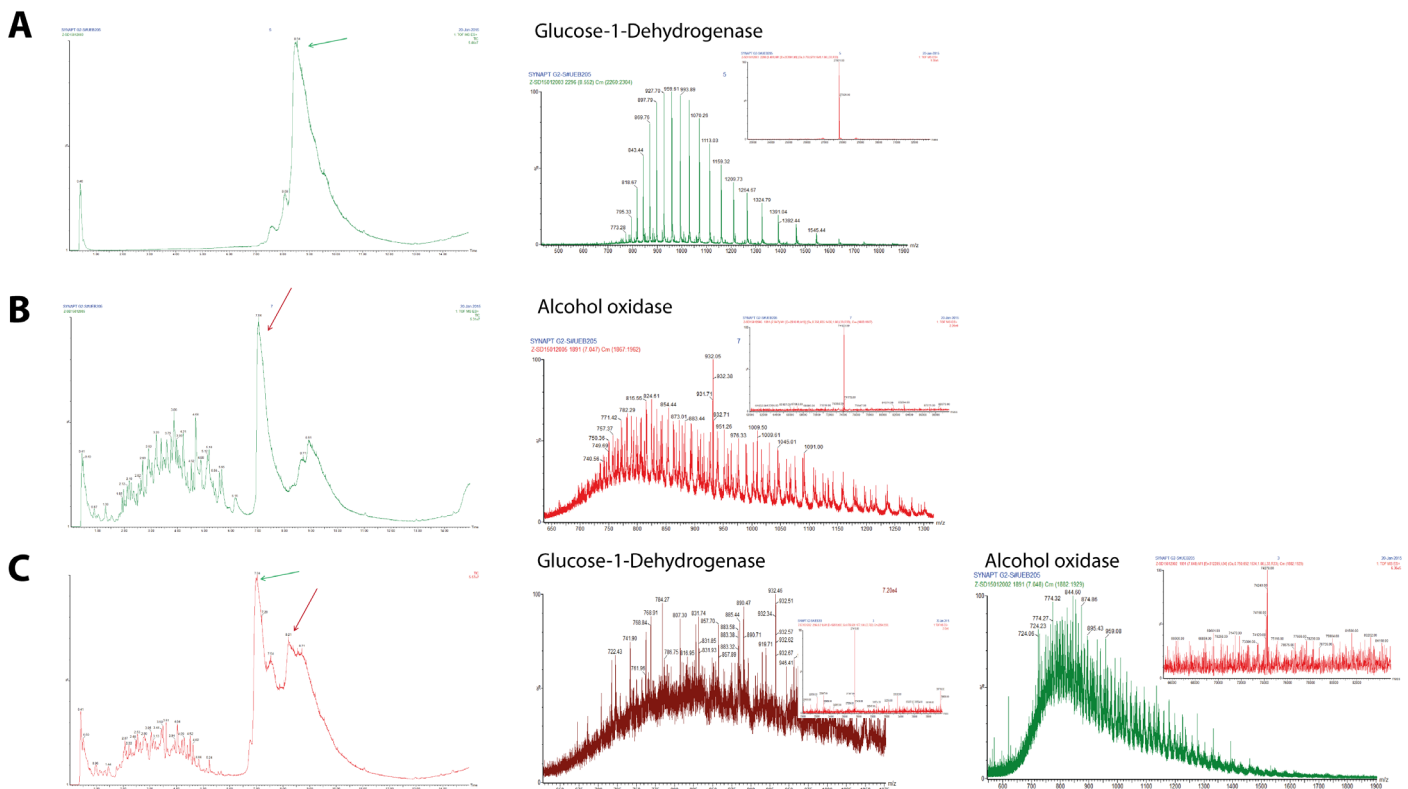


Figure S3.3: UPLC-Mass spectrometry experiments to assay enzyme encapsulation in protocells. Experiments were performed on an Acquity UPLC (Waters) coupled with a TSQ Quantum (Thermofischer). Briefly, we used a Kinetex C18 column (100 x 2.1mm 2.6µm) with H₂O+0.01% formic acid and acetonitrile+0.01% formic acid as eluents and a flow rate of 0.5ml/ml, with ESI+ detection. **(A)** Chromatogram of G1DH enzyme in PBS buffer and MS spectra of main peak, which enables to identify the enzyme with a mass corresponding to literature (~30kDa) **(B)** Chromatogram of AO enzyme in PBS buffer and MS spectra of main peak, which enables to identify the enzyme with a mass corresponding to literature (~74kDa) **(C)** Chromatogram of protocensor encapsulating G1DH and AO enzymes in PBS buffer, and MS spectra of main peaks, which enables to identify the two enzymes.

3.8.2 *In silico* modeling, simulation, and computed output signals

HSIM modelisation

For all simulations, we used a protocell diameter of 10 μ m. To generate heat maps, we ran 5 simulations for each input concentration point in order to average for stochasticity. **Figure S3.4** depicts the enzymatic multisubstrate mechanisms for biochemical reactions, with the associated HSIM model equations we used for simulation.

Modeling protocell permeability

For the purpose of this study, we implemented in HSIM the diffusion rate dn/dt (in mol/s) of input biomarker metabolites from the medium to the inside of protocells through the membrane. This is driven by passive diffusion, given by a modification of *Fick's law*, which states that the diffusion rate across the membrane is directly proportional to the *permeability coefficient* P , to the difference in solution concentrations $C_1^{aq} - C_2^{aq}$, and to the area A of the protocell, or

$$\frac{dn}{dt} = PA \left(\frac{C_{\text{exterior}} - C_{\text{protocell}}}{dx} \right)$$

With for any molecule, the value of P , and thus its rate of passive diffusion, is proportional to its partition coefficient K :

$$P = \frac{kD}{x}$$

Where D is the diffusion coefficient of the substance within the membrane and x is the membrane thickness. These experimental parameters can be easily found in the literature for a wide panel of molecules diffusing across phospholipid bilayers. HSIM supports the introduction of permeability coefficient P ($\text{m}\cdot\text{s}^{-1}$). For this study, we used a P value for ethanol and acetone diffusing passively across DPPC bilayer of around $0.01 \text{ m}\cdot\text{s}^{-1}$. Phospholipid bilayers being naturally impermeable to other organic solutes, we introduced staphylococcus α -hemolysin pore forming protein in the membranes of protocells, in order to allow passive diffusion of input biomarkers metabolites. The diffusion coefficient has been widely measured, and according to a recent measurement by Wanatabe *et al.*⁶⁴⁵ was estimated around $5.1\cdot 10^{-11} \text{ m}^2\cdot\text{s}^{-1}$, which is interestingly only ~ 10 time smaller than in free aqueous solution. Considering the DPPC bilayer to be 3.2 nm wide⁶⁵⁸, one can calculate the permeability coefficient, which gives us $1.6\cdot 10^{-2} \text{ m}\cdot\text{s}^{-1}$. I further hypothesize that only one third of protocells surface would be accessible to hemolysin treatment, which would then correspond to a value of $0.5\cdot 10^{-3} \text{ m}\cdot\text{s}^{-1}$.

BIOCHAM modelisation

Validity domains were computed to extract concentration thresholds (N and R) at steady state (T in seconds) satisfying temporal logic specifications. We first performed sensitivity analysis on concentration parameters with a specified logic formula and numerical temporal properties corresponding to requested systems behavior, with 0.5 variations on parameters. This permitted us to identify the two most sensitive initial state concentration parameters of the systems which we then used to visualize the design space through comprehensive map of configurations satisfying specifications. We then conducted an automated concentration parameter search according to the stochastic optimization method CMAES implemented in Biocham (covariance matrix adaptive evolution strategy⁶¹⁸). For instance, for the model GluONE, this process would require the following commands in Biocham (complete code can be found in **Annexes**):

```
% Trace analysis: extraction of output concentration thresholds (N and R) and switch time (T) at steady state (FG):  
validity_domain(F(G((Time>T) & (N > [NADH]) & ([resorufin] > R))))).
```

```
% SENSITIVITY analysis of concentration parameters c, a, b and f with 0.5 variations:  
%Temporal specification with and time horizon of 10 minutes and concentrations N>10000000 & R>1000000  
sensitivity([c,a,b,f],[0.5,0.5,0.5,0.5],F(G((N > [NADH]) & ([resorufin] > R))), [N,R], [10000000,1000000], 600).
```

```
% VISUALIZATION of satisfaction landscape:  
landscape([c,a],[[0,10000000],[0,100000000]],F(G((N > [NADH]) & ([resorufin] > R))), [N,R], [10000000,1000000], 10, 600,  
landG1DHADH).
```

```
% PARAMETER SEARCH:  
search_parameters_cmaes([c,a,b,f],[[0,10000000],[0,100000000],[0,10000000],[0,2000000000]],F(G((N > [NADH]) &  
([resorufin] > R))), [N,R], [10000000,1000000],600).
```

All analysis and *in silico* modeling were performed according to input concentration parameters corresponding to pathological threshold values. The following pathological threshold for input biomarkers were rationally specified according to clinical requirement, and used for calculations and parameter optimization:

```
Ketones>17 µM (10mg/dl, pathological if >0)  
Glucose>1.39 mM (25mg/dl, pathological threshold)  
Lactate>10 µM (pathological if >0)  
EtOH>17.4 mM (80mg/dl, equivalent to DIU)  
NOx>1000 µM
```

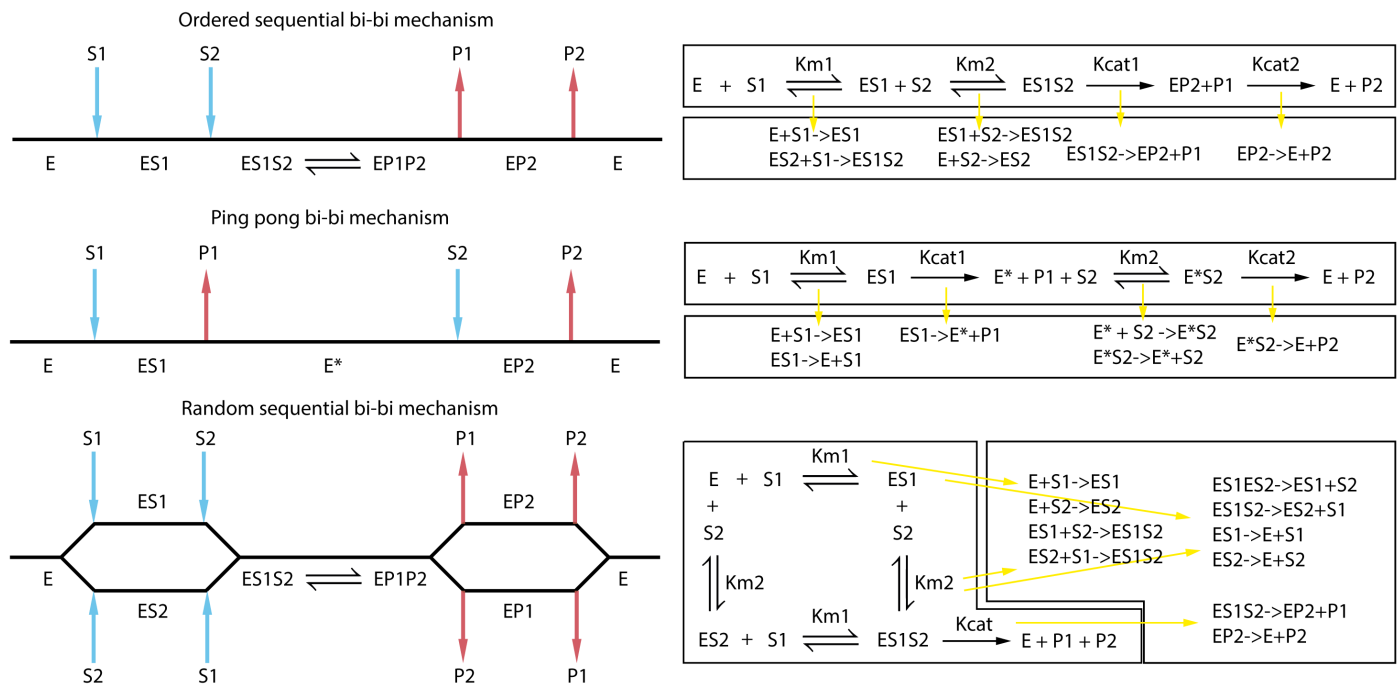


Figure S3.4: Multisubstrate mechanisms for enzymatic reactions and corresponding HSIM equations

Translation of probability from HSIM stochastic simulator to Mass action rates for Biocham ODE solver

HSIM manages molecules in terms of copy number, and not in concentration terms, as it takes into account compartments volumes. Biocham uses concentrations, where Mass action rate factors intrinsically integrate the volume parameter. For a monomolecular reaction, in Biocham as K is in volume/time and $[A]$ in copy number/volume, $K*[A]$ is the number of reactions that will happen per units of time. Likewise, HSIM integrate a constant time step in its probabilities. Similarly, for a bimolecular reaction, in Biocham K is in volume²/time and $[A]$ and $[B]$ in copy number/volume, which gives us a number of reaction events per units of time. HSIM probabilities have thus to be translated into BIOCHAM mass action rates. Mass action rate (MA) used in the Biocham ODE solver can be related to HSIM stochastic simulator probabilities (P) according to:

$$(MA) = \frac{1}{\tau} \frac{\alpha}{V} (P) \quad \text{For bimolecular reaction of the form } A + B \rightarrow C [P] \text{ (order 1 reaction)}$$

$$(MA) = \frac{1}{\tau} (P) \quad \text{For monomolecular reaction of the form } A \rightarrow B [P] \text{ (order 0 reaction)}$$

With τ , V , and α corresponding to HSIM iteration step ($100 \mu s = 10^{-4}$ sec.), experimental proportionality factor, and protocell volume, respectively.

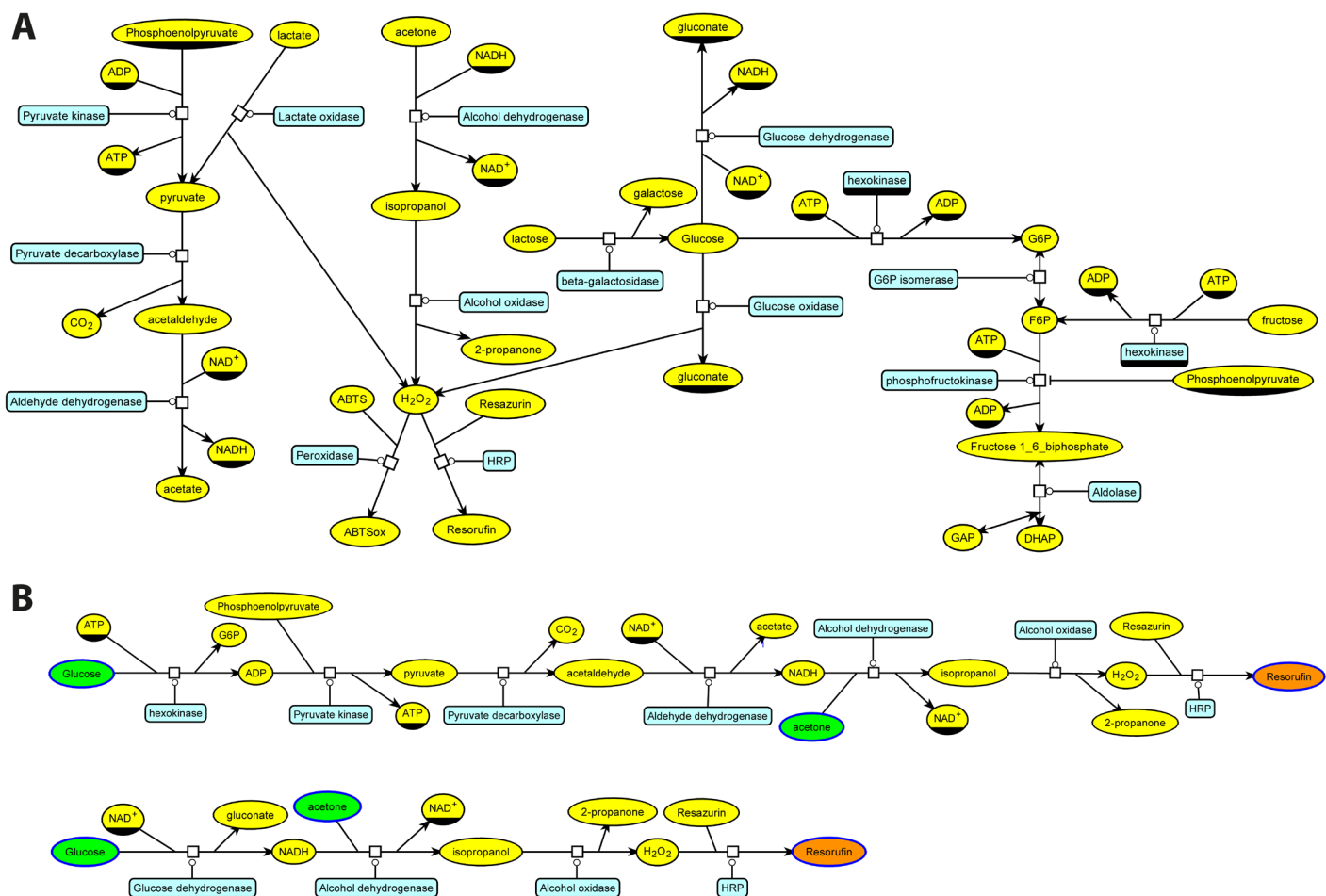


Figure S3.5: Example of automated biochemical implementation of user defined enzymatic Boolean logic gate from natural metabolic networks. (A) Input arbitrary metabolic network comprising biochemical reactions from glycolytic pathways, which will be mined to find enzymatic logic gates. An implementation of a logic gate is a subnetwork where appropriate biomolecular inputs and output are identified. (B) A given Boolean function with its truth table is given to the *in silico* tools (NetGate and Netbuild) developed by Patrick Amar and Marc Bouffard⁶¹⁷, in this case the GluONE system, and the metabolic network will be searched for corresponding concatenated logic gates. Briefly, in a first step, all the possible implementations of the logic gates present in the input network are enumerated. In a second step, these implementations are checked against the given truth tables and the gates found are sorted. In this example, the software found 2 implementations satisfying GluONE biomolecular logic from the input network. The simplest one is the synthetic circuit chosen in this study.

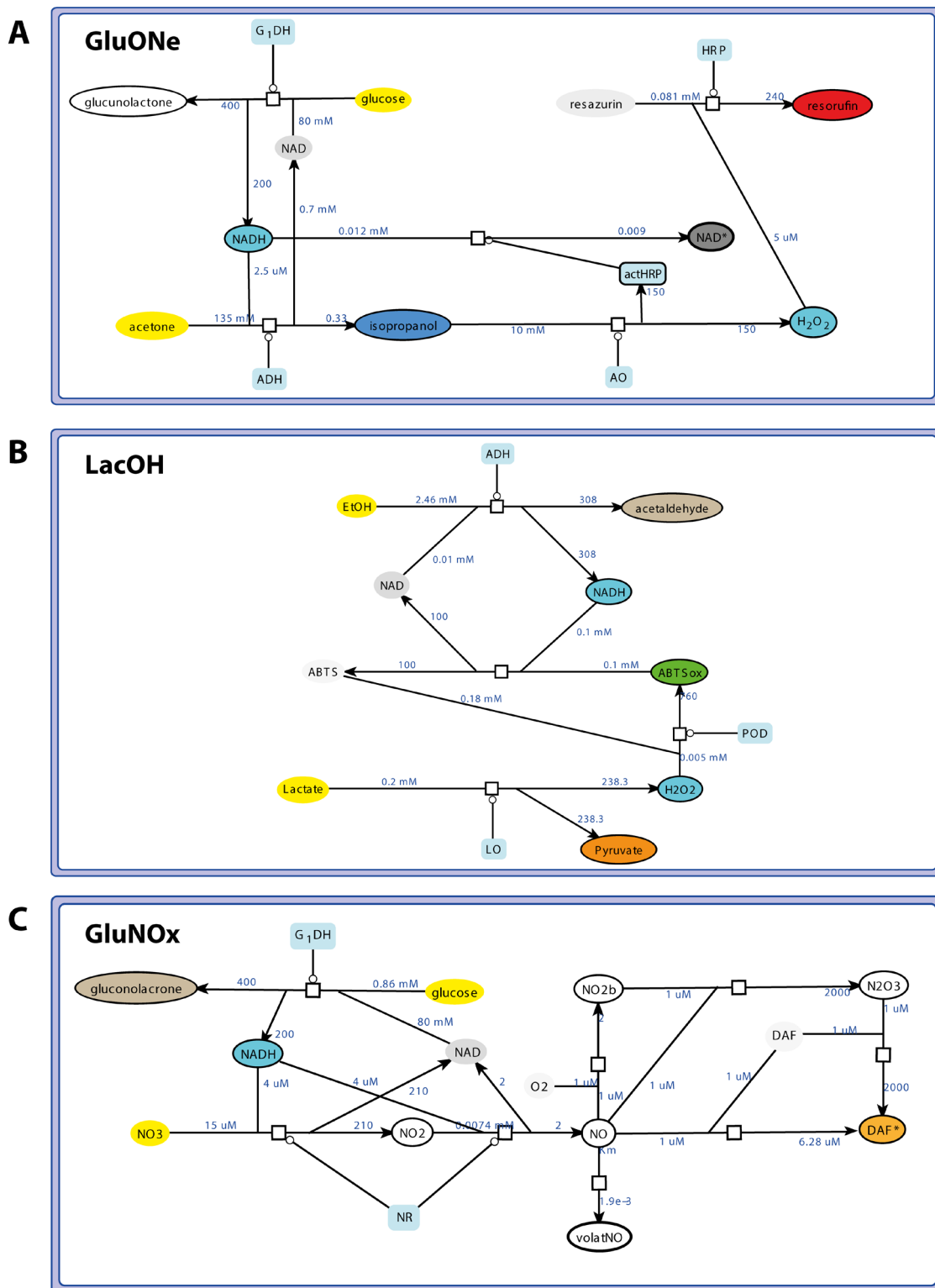


Figure S3.6: Topology of biochemical circuits with kinetic parameters that were designed and used in this study. Metabolites corresponding to systems inputs are depicted in yellow. **(A)** The GluONE system takes Glucose and Acetone as inputs in the medium and applies AND and N-ImPLY Boolean logic to inputs, to generate a absorbance and fluorescent output signal in the molecular form of NADH and Resorufin, respectively. It comprises 4 different enzymes and 2 different metabolites. **(B)** The LacOH system takes Lactate and Ethanol as inputs in the medium and applies N-ImPLY Boolean logic to inputs, to generate an absorbance and colorimetric output signal in the molecular form of oxidized ABTS. It comprises 3 different enzymes and 2 different metabolites. **(C)** The GluNOx system takes Glucose and NOx as inputs in the medium and applies AND Boolean logic to inputs, to generate a fluorescent output signal in the molecular form of nitrosylated DAF-2. It comprises 3 different enzymes and 2 different metabolites.

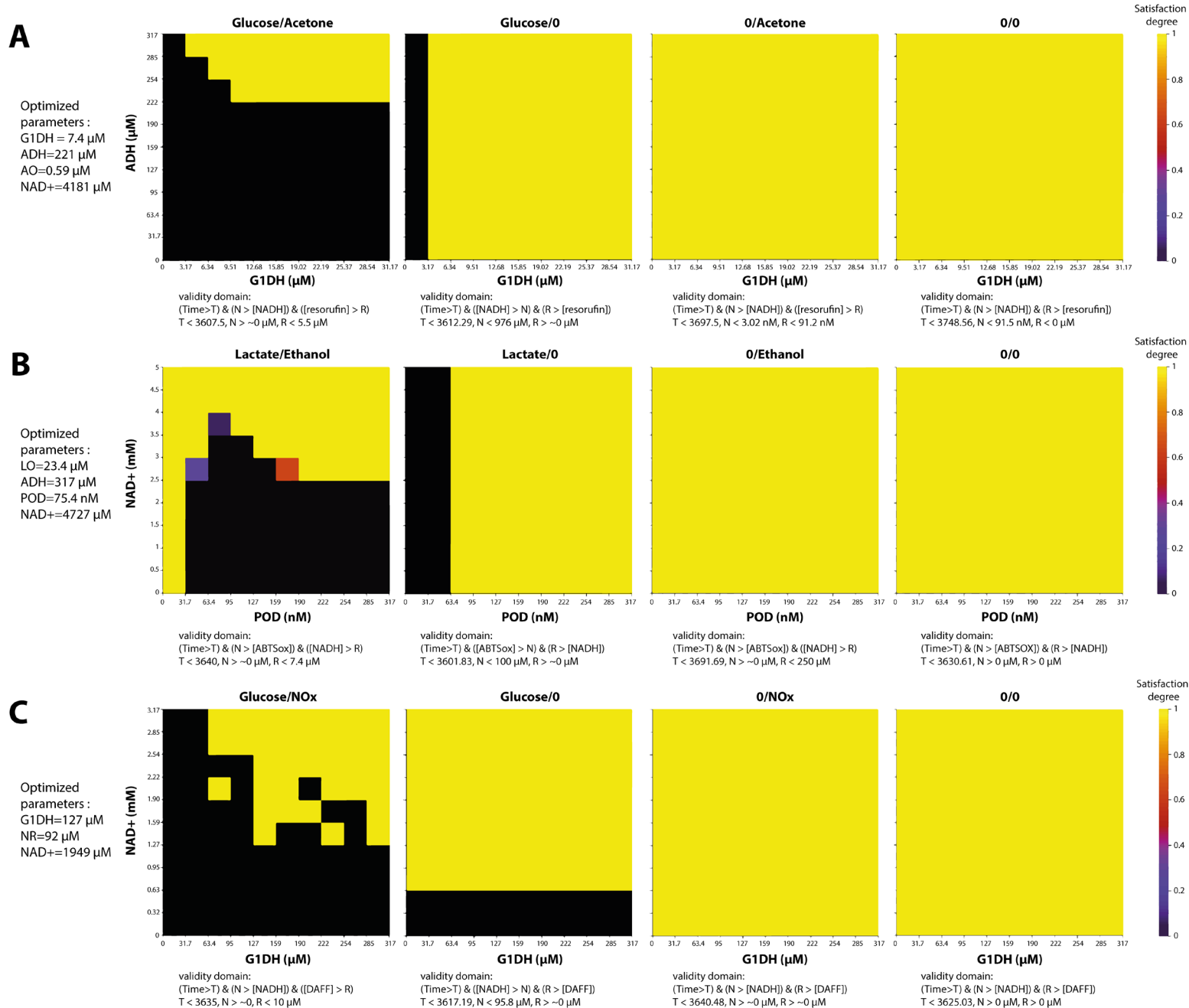


Figure S3.7: Mapping satisfaction degree landscape for (A) GluOne, (B) LacOH and (C) GluNOx biochemical circuits in protocells and different input biomarkers according to clinical specifications. Satisfaction degrees of temporal logic formulas at 10 minutes were computed while varying the two most sensitive parameters of respective models (e.g. ADH and G1DH), for each combination of inputs. Validity domain for temporal specification of output concentration thresholds at steady state are depicted below each maps. Optimized concentration parameters were then computed using C-MAES method implemented in Biocham.

Computed absorbance and fluorescence maps

In order to predict system's fluorescence and absorbance outputs *in silico* depicted in **Figure 3B** and **6A**, we needed to calibrate computed output concentration to experimental values. This could be achieved by generating experimental calibration curves, which could be then analyzed to yield a mathematical relation between concentration and signal (**Figure S3.7**).

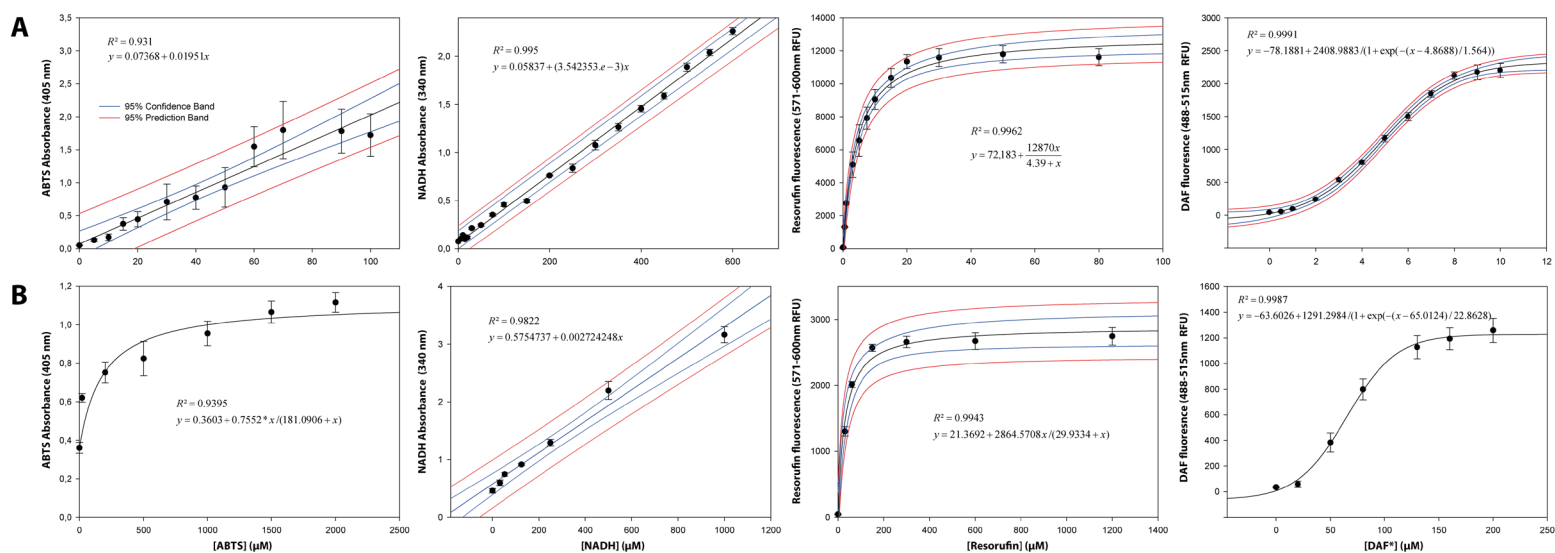


Figure S3.8: Experimental calibration curves used to compute output signals. In order to obtain a mathematical relation between concentration of outputs and experimental fluorescence and absorbance measurements, we measured signals from samples spiked with known concentration of output molecules. **(A)** Output molecular signal in PBS buffer **(B)** Output molecular signals in protosensors. Depicted are the mathematical formulas used to compute output signals in *in silico* simulation.

3.8.3 Complete characterization and kinetics measurements of synthetic biochemical circuits

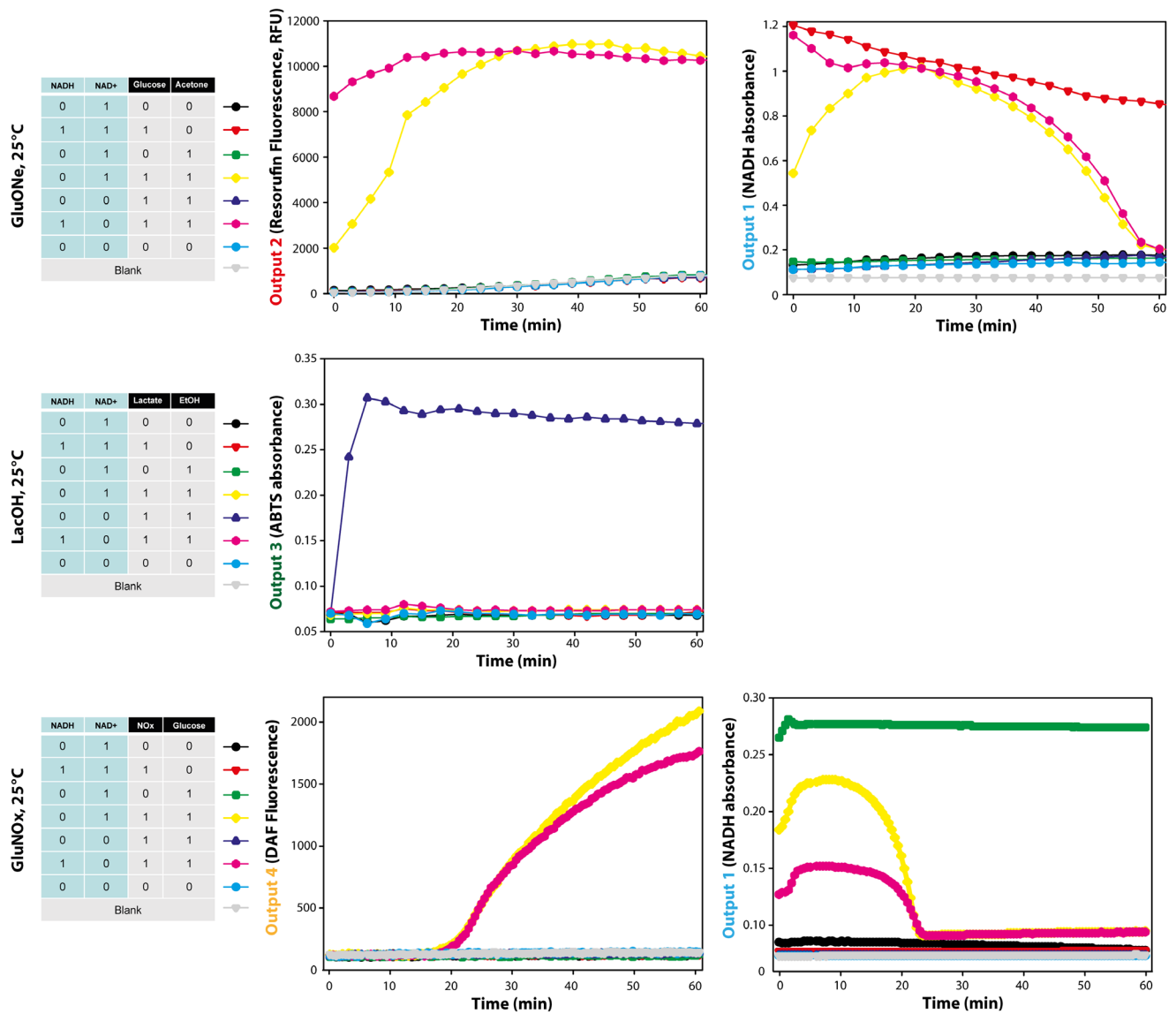


Figure S3.9: Detailed experimental kinetic characterization of synthetic biochemical circuits *in vitro*. Enzymes and metabolites were mixed in p96 100µl wells in PBS, homogenized via smooth agitation, and inputs were added last. Kinetic measurements were performed on a Synergy H1 plate reader, under slow agitation at 25°C.

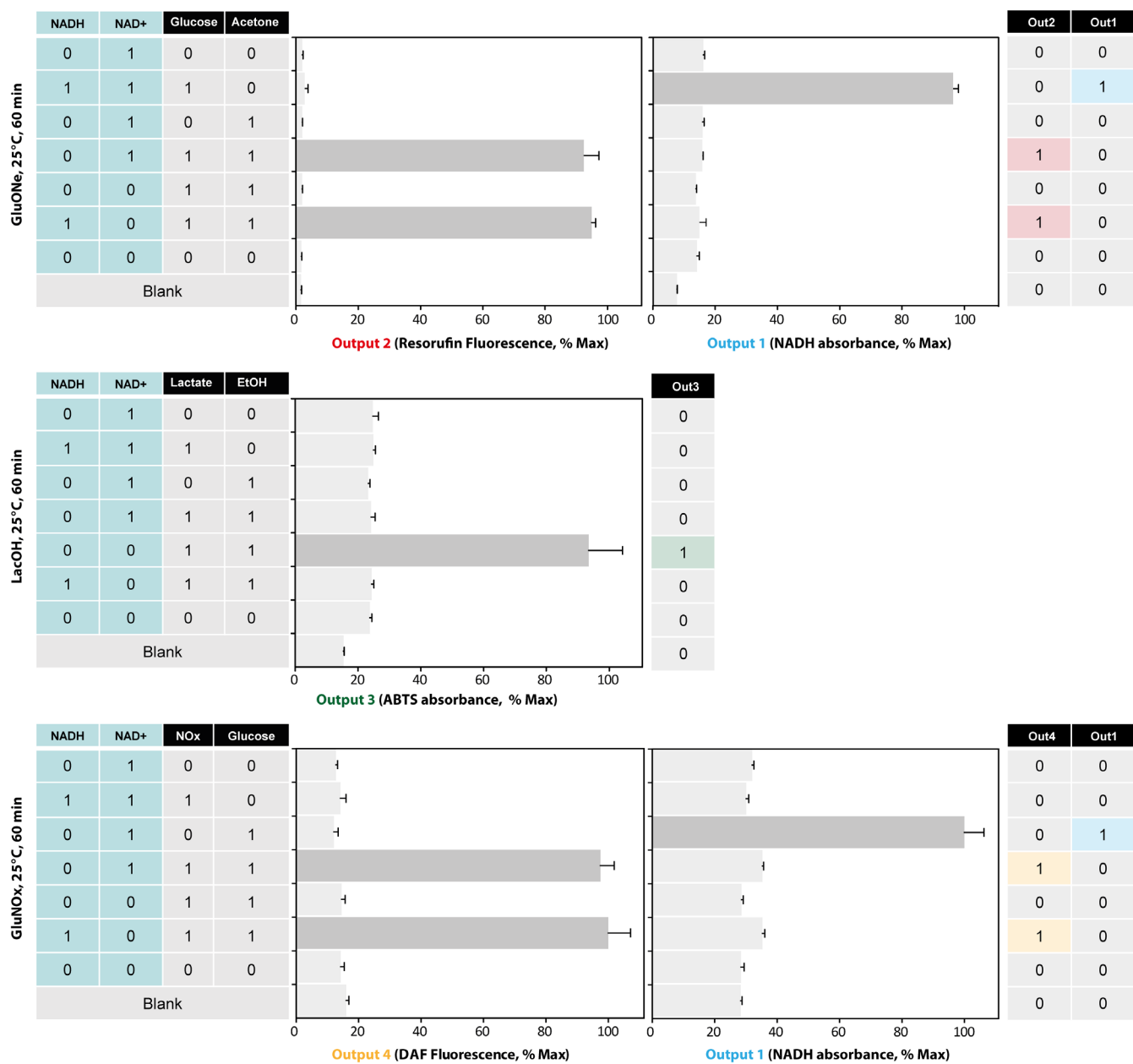


Figure S3.10: Detailed experimental logic characterization of synthetic biochemical circuits *in vitro*. In this figure are depicted measurements obtained as previously detailed in **Figure S3.9**.

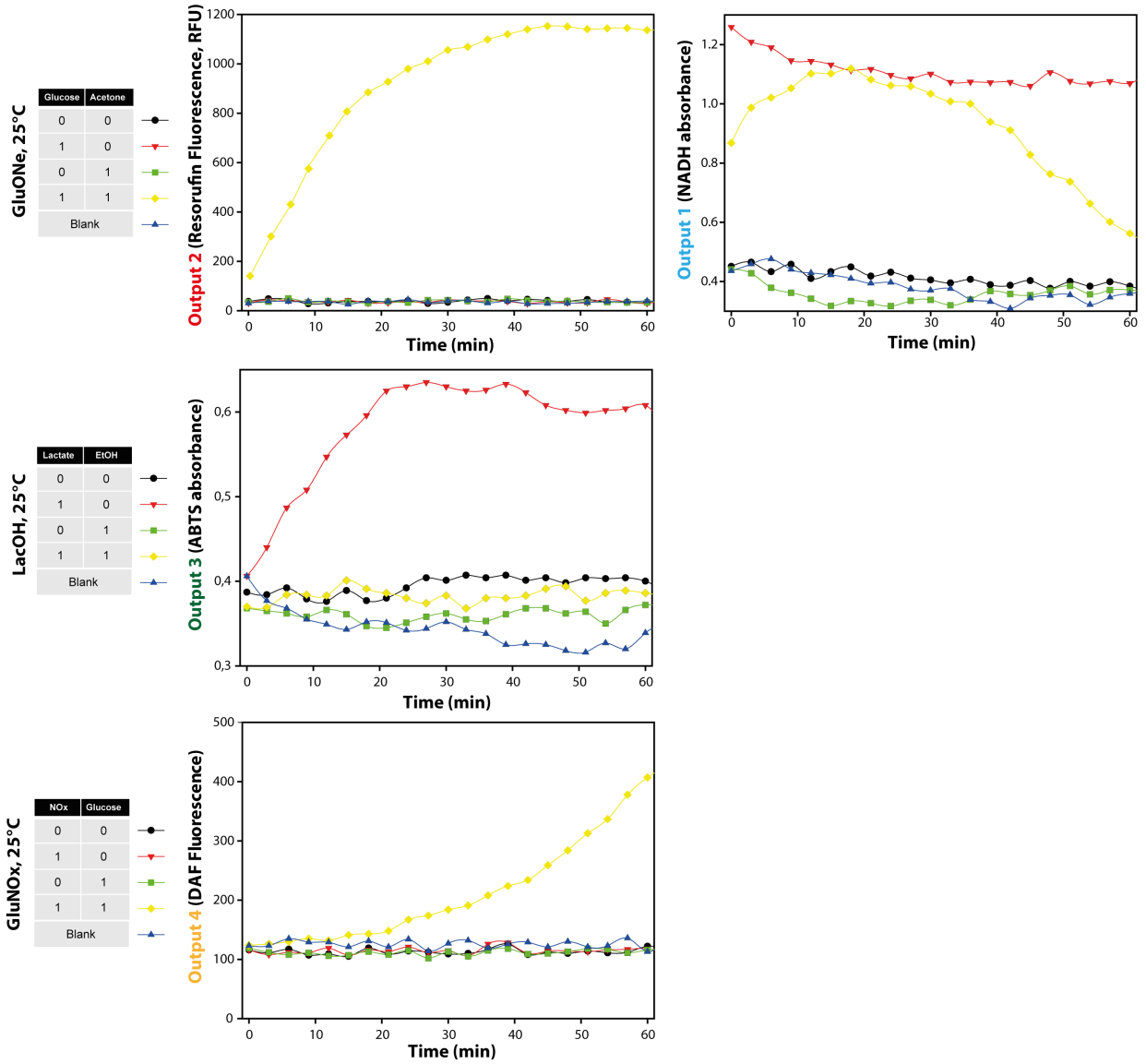


Figure S3.11: Detailed experimental kinetic characterization of synthetic biochemical circuits in protocells.

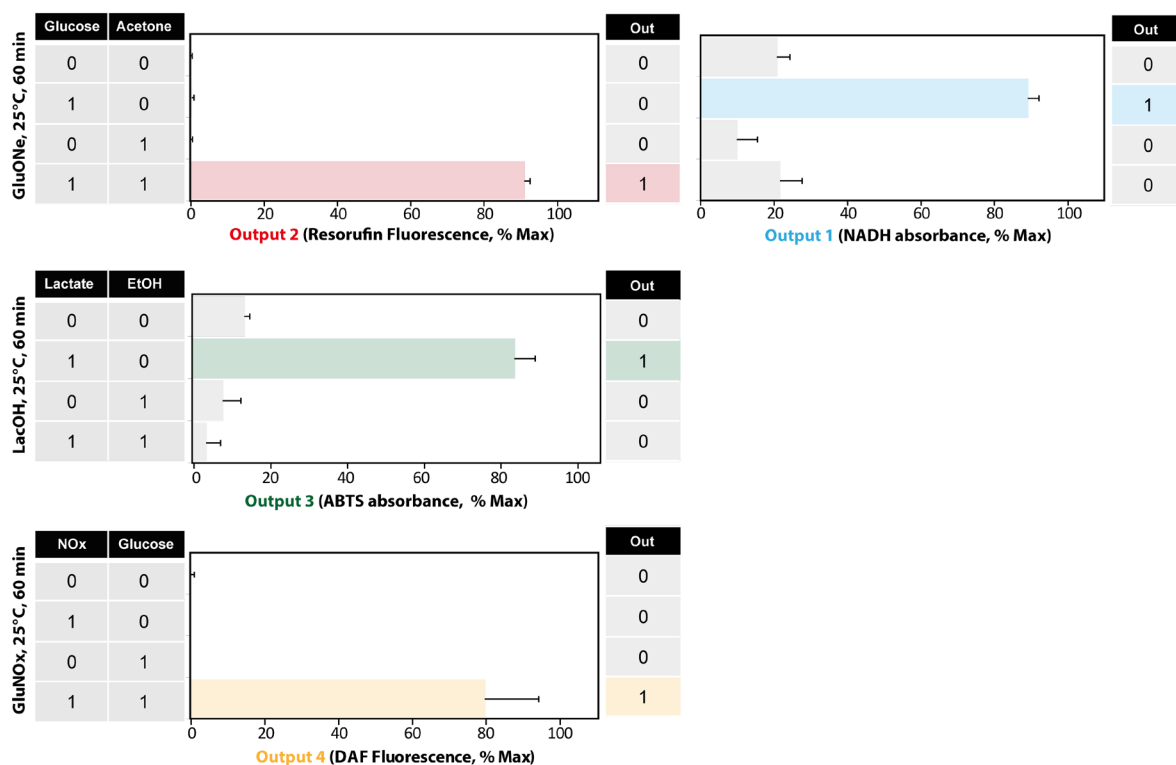


Figure S3.12: Experimental truth tables of protosensors operating in PBS buffer.

3.8.4 Microscopic size dispersion measurements

To assess the size dispersion of protocells, we took random size calibrated microphotograph of protocell preparations using a LEICA DMIL inversed microscope equipped with a 40x lens and a Canon 550D camera mounted on a phototube. Microphotographs were then processed using ImageJ software and a custom script, which allowed for automated size analysis of protocells. Size dispersion figure were plotted and fitted with using the *allfitdist.m* script in Matlab software. (See **Annexes** for script code)

3.8.5 HSIM and Biocham code for models used in this study

The reader is kindly referred to **Annexes**.

I do not fear computers. I fear the lack of them.

Isaac Asimov, *The Age of Miracle Chips*, 1978

Chapter 4

Engineering universal protocell biocomputers

Portions of the text below were adapted from *Computing with synthetic protocells* by Courbet, Molina and Amar, *Acta Biotheo.* (2015)³⁹⁰

Biomolecular computation has been widely seen has never to compete with conventional computers made of silicon integrated circuits. As we have precedently discussed throughout this thesis, traditionally the ultimate goal of synthetic biology has been to apply biocomputing to create *embedded controllers in situ*, such as programmed synthetic circuits in eukaryotic cells, bacteria, viruses, or fully synthetic devices engineered to perform useful and simple computation, for instance to connect bioanalytical operation through biochemical sensing with bioactuation such as drug delivery or targeted cell death.

However, if biocomputing were to compete with *in silico* sequential computing on specific tasks, we wondered which, if possible, could be the practical modalities. In this perspective, protocells appeared as interesting substrate for biocomputing since we showed they could be programmed with important design space, for instance recapitulating simple Boolean operations. Moreover, it appeared that one could take advantage of the microscale of protocells, which enables to process information a huge number of times in a restricted volume, a capability that could prove relevant as novel versions of massively parallel computing.

To my knowledge, the theoretical work we present here constitutes a totally novel approach, which has, to my knowledge, received very little if no attention. Apart from practical biosensing application permitted by programmable protocells, we wondered if the systematic framework we developed could be extended to encompass any sort of operation at the microscale. More specifically, we investigated how protocells could be engineered into standardized components, programmed to implement arbitrary Boolean function, and how they are further concatenated into complex machines, and used as universal computers to solve complex problems.

Abstract

In this article we present a new kind of computing device that uses spatially constrained biochemical reactions circuits as building blocks to implement logic gates. The architecture of a computing machine relies on these standardized and composable building blocks, computation units, which can be used in multiple instances to perform complex Boolean functions. Standard logical operations are implemented by synthetic biochemical circuits, encapsulated and insulated within synthetic vesicles called protocells. These protocells are capable of exchanging energy and information with each other through transmembrane electron transfer. In the novel paradigm of computation we propose, *protocomputing*, a machine can solve only one problem and therefore has to be built specifically. Thus, the programming phase in the standard computing paradigm is represented in our approach by the set of assembly instructions (specific attachments) that directs the wiring of the protocells that constitute the machine itself. To demonstrate the computing power of protocellular machines in a practical example, we apply it to solve a NP-complete problem, known to be very demanding in computing power, the 3-SAT problem. We show how to program the assembly of a machine that can verify the satisfiability of a given Boolean formula. Then we show how to use the massive parallelism of these machines to verify in less than 20 min all the valuations of the input variables and output a fluorescent signal when the formula is satisfiable or no signal at all otherwise.

4.1 Introduction

What is Computation? One definition could be *the goal-oriented process that transforms a representation of input information into a representation of output information*. The process itself can be iterative (or in another form recursive), in this case it is called an algorithm, but other forms of processing can be used, such as neural networks or first order logic.

A computation process, whatever it is, has to be run by a computer, which can be a human using pen and paper, or a machine specifically built for that purpose. The most popular form of computer is an electronic device that uses a digital representation of data, and manipulates this representation according to a set of instructions that implements the algorithm transforming them into results. The set of instructions is then called a computer programme. Electronic computers use numbers, integer and floating points, to represent data. These numbers are commonly coded in base 2, which can also be directly used to encode Boolean values and therefore easily implement conditional calculations. Electronic computers are mainly built from basic blocks, logic gates, which are interconnected to make the arithmetic and logic units, memory registers and microcontrollers that form the Central Processing Unit which in turn, along with the Main Storage Unit, and the I/O Controllers constitute the computer itself.

Therefore, one can build a digital computer using any technology that can mimic logic gates and their interconnections. We intend to demonstrate in this article how to implement single Boolean logic gates using synthetic minimal biological systems embedded in a vesicle, or protocell, and how to connect them together to fabricate a device, or protocellular machine, that computes a complex logical function. The computing model that underlies our biochemical implementation

of a computer is similar to the one of an electronic computer, giving their computing capabilities are the same.

The fundamental characteristic of electronic computers is their ability to run a potentially infinite number of algorithms doing a wide variety of computations on data, because they are programmable: the same computer can run sequentially (or pseudo-concurrently) as many different programmes as those that can reside in its main memory storage, along with the associated data.

Here, we propose a methodology where programming the computer is analogous to its physical assembly. The program thus resides in the set of instructions given for the assembly process. Furthermore, we show how to build (i.e. program) a kind of computer that can solve one problem belonging to a class known to be hard to solve: a NP complete problem.

The computational complexity theory explores the feasibility of computational problems, in terms of computing time (or memory space) needed to solve a problem of a given size. In the Von Neumann based architectures (standard electronic computers) the number of computing elementary steps (instructions) is often used to approximate the computing time, since each instruction takes approximately the same amount of time to be performed.

There are two main classes of computational problems, those that can be solved by a deterministic machine in a number of steps which can be expressed as a polynomial of the problem size (class P), and those that can be solved in polynomial time, but on a non-deterministic machine (class NP). Typically decision problems where (i) a solution can be verified in polynomial time and (ii) there is no other known algorithm except generate and verify all the potential solutions, are NP problems. Solving these problems on a Von Neumann computer requires an exponential number of steps with respect to the problem size. A NP problem is said to be NP-complete if any other NP problem can be transformed into this problem in polynomial time⁶⁵⁹. In consequence NP complete problems are more difficult to solve than any other NP problems because if one NP-complete problem is quickly solved (in polynomial time) then all the NP problems will be quickly solved. Of course all these complexity classes collapse if $P = NP$ (which is one of the great open conjectures in computer science).

We have chosen the 3-SAT problem, a variant of the boolean satisfiability problem (SAT), as an example of NP-complete problem⁶⁶⁰ a protocellular computer can solve elegantly. This is mainly because the very small size of protocells and their 3D packing allow us to build a machine made of billions of logic gates specifically connected to solve a given 3-SAT problem. Another characteristic of our protocellular machines is that they are disposable in the sense that once the computation is done for a given set of input values, the machine is no more usable. But the counterpart is that the energy needed for the computation is very low⁴⁵¹. Finally, the biochemical nature of the protocellular machines makes them very easy to interface with living organisms. For example, they can be used for medical diagnosis to implement biosensing coupled with medical decision algorithm

4.2 Methods

4.2.1 Protocells as computation units: definitions

The bottom-up design of biological systems is made possible by the synthetic biology approach that applies engineering principles to biology in order to design standardised biological parts, devices, systems in a systematic and rational manner. Hierarchical abstraction of biological functions enables the assembly at the system level of new biological systems with user-defined functionalities^{6 7 8}. The behaviour of synthetic systems is predictable and design can be automatised *in silico* before attempting to implement them with biological components⁵⁹¹. In addition, the remarkable capacity of biological building blocks to compute in highly sophisticated ways has led scientists to design and engineer biomolecular computers³⁹. Thus far, most biocomputing has been investigated from the top down perspective, that is, by modifying existing organisms⁴⁶⁰. The strategy we propose here, prototyping, is interested in implementing protocells from the bottom-up perspective to perform computation, where very little attention has been given^{568 567 391}.

Starting from an abstract operation that is to be computed, one can rationally and systematically choose biochemical species for the implementation (metabolites, enzymes, nucleic acids...) (**Figure 4.1A**). Standardised and robust biomolecular components and reactions can be engineered, tested and optimised to implement different types of biological functions or computations⁵⁸⁹: simple Boolean operations, memory devices, amplifiers, analog to digital converter, oscillators etc. (**Figure 4.1B**). In addition, we previously demonstrated how this process can be automatised using CAD tools recently developed for this purpose^{589 367 588}. For example, an AND biochemical logic gate taking reduced metabolites as inputs (NADPH and FADH₂) can be implemented using a network of 3 different enzymes and 4 different metabolites connected by 3 biocatalytic reactions, and transferring electrons to NADH as an output. In the same way, we can implement a set of standardised computation units that recapitulate all Boolean logic gates (see **Figure 4.3** for examples of implementations of AND, NOT and NOR gates). Electron transfer can also be coupled to various output biological functions to produce human readable signals (**Figure 4.2**) or enable the selection of machines with specific behavior for further analysis. We propose to exploit specific reduction of species to trigger readable outputs, either luminescence or fluorescence (i.e. reduction of rezasurin into fluorescent resorufin) or the transport of a ligand (or its receptor).

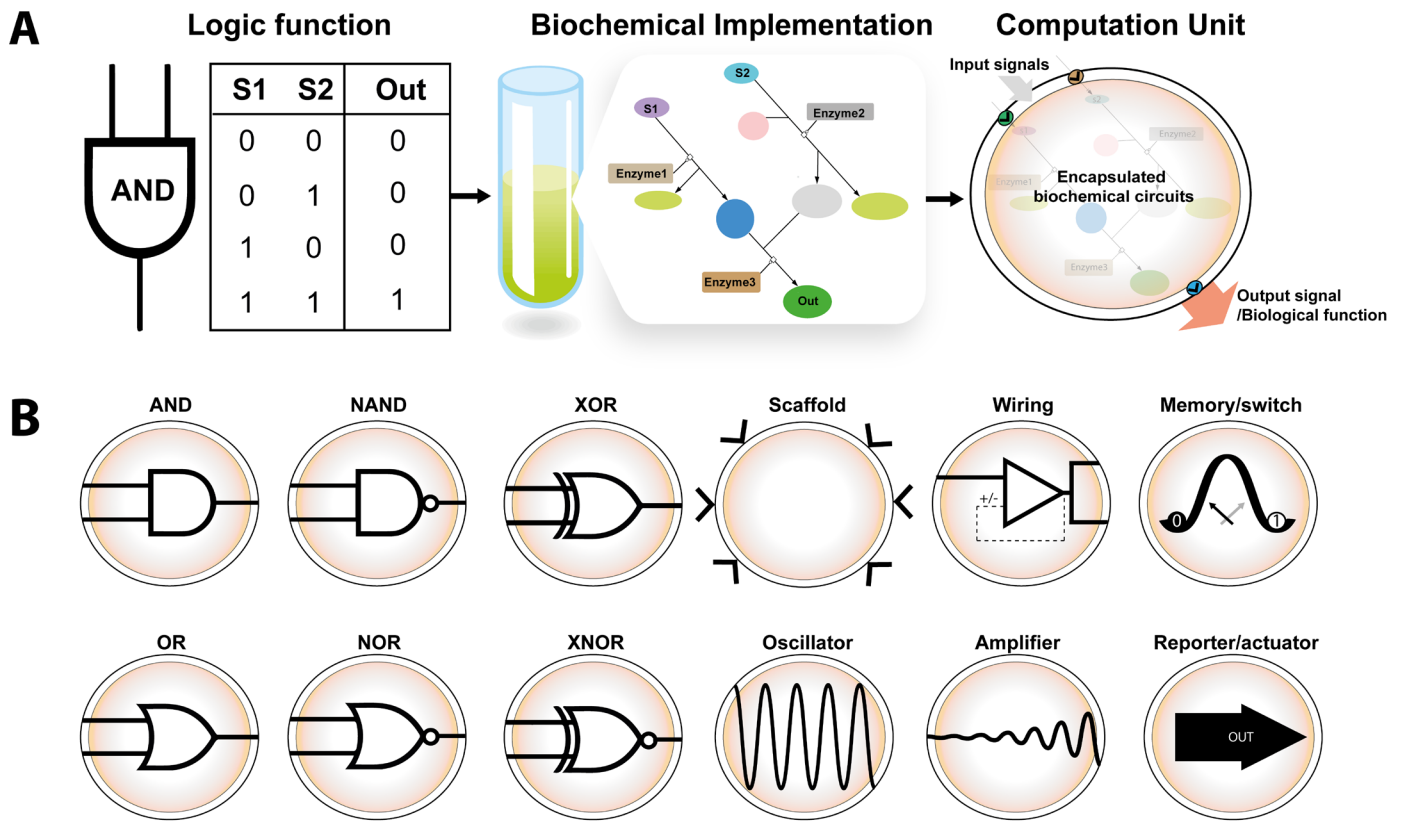


Figure 4.1: (A) Rational design of a computation unit implementing a given logical function. (B) Different types of computation units. An AND gate outputs true only if the two inputs are true; An OR gate outputs true if at least one of the inputs is true; A XOR gate outputs true only when one of the inputs is true; The NAND, NOR and XNOR gates outputs the opposite value of the AND, OR and XOR gates respectively.

Our approach improves the modularity of biomolecular computing systems by the fact that biochemical networks implementing Boolean logic are encapsulated within synthetic vesicles, or protocells, distinguished by their high degree of organization and control over biological processes provided by the membrane boundary³⁹⁸. Such architecture of insulated computing units allows us to use many instances of the same type of protocell anywhere in the circuit when the same logic gate is needed. Moreover, this enables the concatenation of multiple layers of protocells to achieve complex information processing capabilities. In such architectures, input information arrives from upstream connections with previous protocells, to output connections to following computation units. As each logic gate is encapsulated within an impermeable vesicle, the reactions that compute the output value will go from the non-equilibrium initial state to an equilibrium state. Therefore, once a logic gate has finished to compute the output, it is no more able to do another computation. So this first model of protocellular machine is in essence a kind of disposable computer.

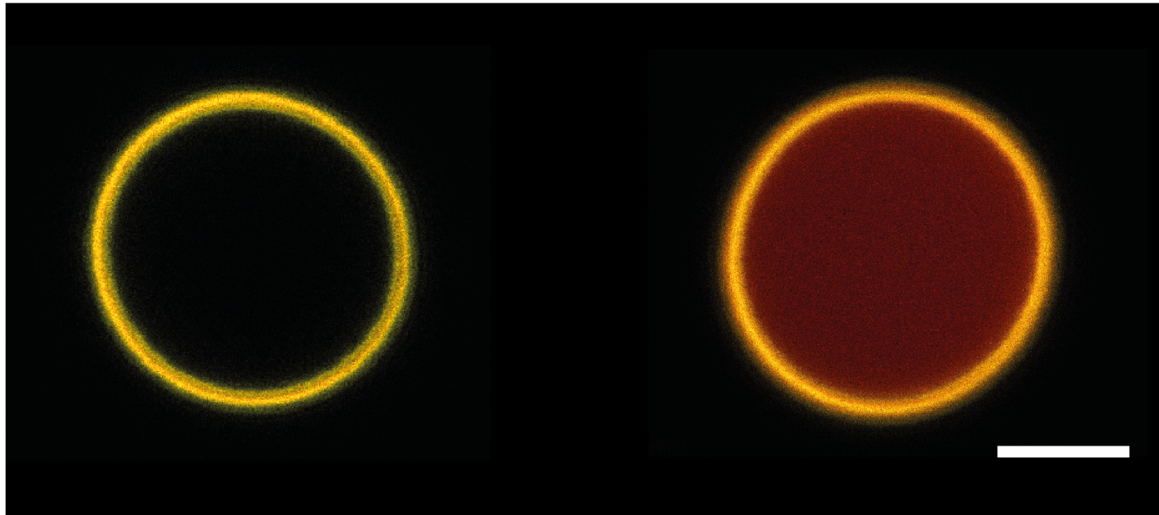


Figure 4.2: Example of experimental fluorescence signal triggered in micrometric protocells. Enzymatic electron transfer from carbohydrate to the redox sensor probe (in that case resazurin is reduced into the red fluorescent product resorufin). Phospholipidic protocells encapsulating biochemical species were generated using microfluidic devices, and imaged using a confocal microscope. Left no induction; right induced with glucose (scale bar=5 μ m).

Encapsulation of biochemical networks can be achieved using natural bilayer membranes (e.g. phospholipid bilayers, liposomes)³⁸⁶, or engineered membranes (e.g. copolymers, polymersomes)⁴⁰⁸, with respect to stoichiometry of internal species and incorporation of membrane proteins for connections^{661 572 409}. This process is also known to stabilise enzymes, prevent cross-talk, denaturation or proteolysis and improve enzymatic properties^{662 663}. In addition, streamlined workflows, for example relying on microfluidics as we discussed previously, are already available for the high-throughput generation of protocells that encapsulate various substrates^{625 664 665}. This strategy, extensively used in our lab, allowed us to test the implementation of various protocellular logic gates. Such vesicle have proven to be sufficiently stable (i.e. not prone to fuse together or physical disruption) to enable the construction of such multi vesicular assemblies^{666 628}. Tunable sizes ranging from 50 nm to 50 μ m can be obtained, although in our approach, size should be kept as small as possible to obtain the highest density of computing operators.

4.2.2 Circuit wiring

To obtain a full circuit implementing a given Boolean function, we then need to concatenate and wire basic logic gates. The design of a function-specific protocellular machine exploits the composability of *computation units*. Amongst a specific set of protocells, multiple instances of the same logic gates can be wired together to implement a user-defined function. One way to achieve successive reactions in each layer of a protocellular machine, from input to output protocells, is to drive them using electrochemical potential (e.g. oxido-reduction reactions). By analogy with electronic computers, electrons are energy carriers and the redox potential is the current of the system, which could be measured with an electronic device. The major difference is that inside a

protocell, wires are replaced by free molecules (e.g. NADH, NADPH, FADH₂), and effective wiring is achieved using chemical selectivity of enzymes. Molecules are either electron donors or acceptors, obeying biological enzymatic rules resulting in current and energy for computation. In such systems, the in -> out direction is driven by the thermodynamics of the redox reaction. In our example, a protocell giving the *true* value would have a reductive state with high concentration of NADH, which can then transfer its electron to reduce the input of the next protocell. Conversely, a protocell giving the *false* value does not output any electron. In addition, electron transfer occurs only between physically connected protocells, through tight junctions putting into close contact electron transfer complexes, which carry out the connections between protocells and therefore between logic gates (**Figure 4.3**).

Here, we propose build a protocellular machine from a set of protocell logic gates assembled in a tree-like layout (see the following section). When set to *true*, the inputs of the machine initiate electron transfer through the chain of protocells that constitutes each branch of the tree, down to the *root* protocell.

In these input protocells, electron production is started by the specific oxidation of molecular species by oxidase enzymes. Electrons are then transferred down the protocell chain via transmembrane electron transport complexes that enable electron coupling (reduction) of specific molecular species. In that sense, input protocells can be seen as the generators that power the machine. Moreover, fuel protocells with a switch like behaviour, could be used to amplify and reshape the signal and therefore counteract its decay.

In order to implement specific electron transfer modules, we propose to exploit the modularity and thermodynamic reversibility of natural oxidative phosphorylation and photosynthesis complexes, which catalyse the electron transfer across natural membranes with specificity to NADH (Complex I), FADH₂ (complex II), and NADPH (NADPH quinone oxidoreductase)⁶⁶⁷. This includes quinone (or chemically related) and cytochrome c shuttle, which are delocalised mobile electron carriers that could be used as inter-protocell transfer molecules. In our design, we propose that a first quinone carrier (or related), could transfer electrons from a specific output signal (substrate specificity given by the first complex: I, II...) to a close complex III, which would then via a mobile cytochrome c transfer these electrons forward to the complex III belonging to the next protocell. This mechanism constitutes efficient reversible energy coupling, which has been shown to work via electron-tunneling across the proteins⁶⁶⁷. Furthermore, recent studies have highlighted the possibility to engineer natural prokaryotic complexes for efficient and substrate specific synthetic electron transporters^{668 669 670}.

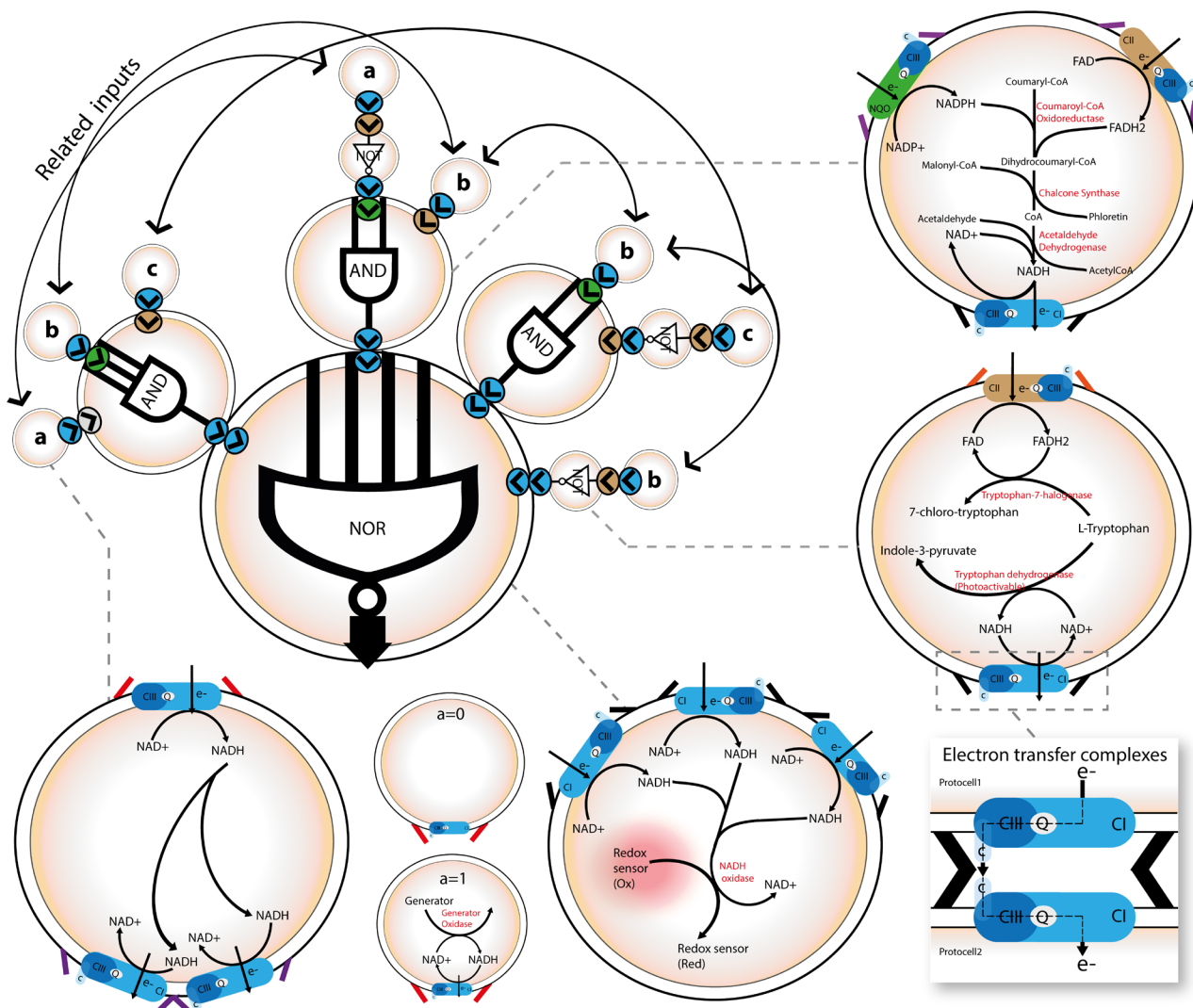


Figure 4.3: Detail of a possible implementation of each type of protocell gate. Each type of logic gates has been simulated in silico with HSIM, and some of them are under test in the lab. The detail of the electron transfer mechanism is shown in the bottom right cartoon. For example, the fluorescent NOR gate uses a cascade of two enzymatic reactions (NADH oxidase, Horseradish Peroxidase) to consume the fluorescent oxidised scopoletin when NADH is present in the protocell, that is when at least one input is set to true, so is transferring electrons to make NADH from the initial pool of NAD⁺.

The architecture of a machine is controlled by the functional wiring of input and output of specific protocells. This can be achieved by using programmable junction modules, which can be selected to implement any protocellular machine in a plug-and-play way (**Figure 4.3**). Biological function for these programmable attachments could be supported by couples of ligand/receptors with high binding affinity, such as aptameric nucleic acids⁶⁷¹⁻⁶⁷² or peptidic binders⁶⁷³, that could be straightforwardly produced in large combinatorial synthetic libraries using SELEX⁶⁷⁴, or ribosome display respectively⁶⁷⁵⁻⁶⁷⁶. Starting from a pre-built stock of computation units, the user can define a set of attachment instructions that corresponds to the Boolean function to implement. Irreversible constructs can be achieved using cross-linking chemicals, so that no unbinding would occur⁶⁷⁷⁻⁶⁷⁸. We assume that the kinetics associated with such an assembly process would be of the order of minutes. Some attachments can also be set as random, to enable

stochastic wiring of different types of protocells to specific positions. This could be used for example to solve problems involving the navigation through a large parameter space where protocellular machines could be used to compute a fitness function. Additionally selection methods could be implemented to isolate protocellular machines that exhibit specific behaviours. Positive selection can be done for example using FACS, conversely negative selection via a self-destruction mechanism.

4.3 The case study

4.3.1 The Boolean satisfiability problem

The NP-complete problem we aim to solve is the 3-SAT problem. This problem can be simply defined as:

Given any boolean formula in Conjunctive Normal Form (CNF), with at most 3 literals per clause, is there a valuation of the variables that satisfy the formula?

In other words, it asks whether the variables of a given Boolean formula can be consistently replaced by the values *true* or *false* in such a way that the formula evaluates to *true*. If it is the case, the formula is called satisfiable. The literals are either a variable (v) or the negation of a variable ($\neg v$); They are connected with the or operator (\vee) to form a clause; The clauses are connected with the and operator (\wedge) to obtain the formula in CNF. For example:

$$F(a, b, c) = (a \vee \neg b \vee c) \wedge (b \vee \neg c) \wedge (a \vee b) \quad (1)$$

is true when $a = \text{true}$, $b = \text{true}$, $c = \text{false}$, so the formula $F(a, b, c)$ is satisfiable. Conversely, the formula:

$$G(a, b, c) = (a \vee b \vee c) \wedge (\neg a \vee b) \wedge (b \vee \neg c) \wedge \neg b \quad (2)$$

is not satisfiable because all the eight possible valuations for a ; b ; c lead to $G = \text{false}$.

To find if a formula is satisfiable, we will build as many protocellular machines as there are combinations of valuations of the input variables. To do this, we will exploit the combinatorial power of ligand-receptor binding to link constant protocells (with *false* or *true* values) to the inputs of the protocellular machine to cover all the value space. A protocellular machine is dedicated to a specific formula, and therefore is not programmable in the sense an electronic computer is. The protocellular machines are self assembled according to the formula they have to check, so in our approach, the program is the process that directs the assembly of the machines. We will ascertain that there is at least one instance of a protocellular machine per possible valuation of the variables.

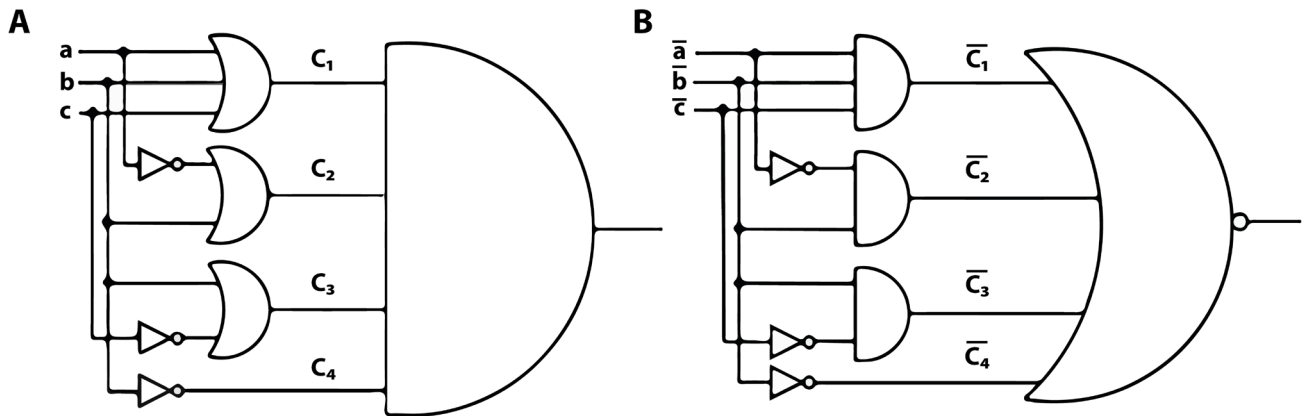


Figure 4.4: (A) Direct implementation of the G formula in standard Conjunctive Normal Form. (B) Using the De Morgan laws, the same Boolean function is rewritten using a NOR gate instead of the final AND gate, easier to build with a large number of inputs, and multiple 2- and 3-inputs AND gates fed with the complement of the original inputs.

An instance of the machine can be made using 2- and 3-inputs OR gates connected to a big AND gate with as much inputs as there are clauses in the formula. Each input of a clause is connected to a protocell representing a variable v sending *true* or *false* when a specific start signal is given, or to an inverter protocell sending the negation of v when the start signal is given. The output of the AND gate is connected to a protocell that fluoresces when the input value is *true*. For example, the protocellular machine corresponding to the G formula would be made of a 4-input AND gate, two 2-input OR gates, one 3-input OR gate and three inverters connected as in the equation above (**Figure 4.4A**).

As we have at least one (and probably more) instance of the machine for each possible valuation of the variables, if at least one of the protocellular machines fluoresces, the formula is satisfiable. Conversely, if there is no fluorescence at all then the formula is not satisfiable.

We can simplify the construction of the machines using the De Morgan laws to replace the big AND gate by a NOR gate, which is easier to build and also more efficient than an AND gate when there is a lot of inputs. Since the output of this NOR gate is the output of the whole machine, the final inverter can be made using an inhibitor of the fluorophores stored inside the protocell implementing the gate. We also need to feed the inputs of the AND gates with the complement of the variables, which could lead us to use a lot of inverters; But they can be avoided because these inputs are the inputs of the whole machine, and since we need to test all the valuations of the variables, these inputs will be fed with constant values. Therefore we can program the assembly of a machine with the constants already inverted (**Figure 4.4B**) and we will need no more inverters than negated variables specified in the original formula.

4.3.2 The assembly of the machines

To obtain one instance of a computing protocellular machine, we need to direct the self assembly of as many copies of AND gate protocells as there are clauses in the formula (except when a clause has only one literal), the output of each AND gate being connected to an input of a fluorescent NOR protocell. The inputs of each AND gate are also to be connected to the output of an inverter or to the output of a wiring protocell (representing the input variables of the formula). Then, to test a valuation of the variables of the formula, the input of each wiring protocell will be connected to special inputless protocells that output the constant value *true* or *false*. Once the machine and its inputs are assembled, when a start signal is given, after a few minutes, the NOR gate of this machine will fluoresce if the formula is true for this valuation of the variables, and therefore the formula is satisfiable.

We must ensure that correlated inputs of two (ore more) AND gates are fed with correlated values. In the previous formula (rewritten using a NOR of ANDs, with the complemented variables as shown in **Figure 4.4B**)

$$G(a, b, c) = \overline{(\bar{a} \wedge \bar{b} \wedge \bar{c}) \vee (\neg\bar{a} \wedge \bar{b}) \vee (\bar{b} \wedge \neg\bar{c}) \vee \neg\bar{b}} \quad (3)$$

The first input of the first clause, a , is always the opposite of the first input of the second clause ($\neg a$), and the second input of the two first clauses, \bar{b} have always the same value, etc. To achieve that we will use inverter protocells, and wire protocells that can transfer their input to two or more outputs.

In this example, since there are 3 variables, we must assemble 8 protocellular machines to test each of the 8 possible valuations. Each line of the table in **Table 4.1** shows the input values (0 for *false*, 1 for *true*) of one of the 8 different protocellular machines, the complemented value of each clause, and the value of the formula (3), which is always *false* (this formula is not satisfiable).

In order to have an efficient assembly mechanism, we split the process in two steps. The first one does not depend on a specific formula, but on the maximal numbers of variables (V_{max}) and of clauses (C_{max}) a formula can have. To be able to test any given formula within the limits of size we stated, we build a reservoir containing at most for one protocellular machine instance:

- one C_{max} -input NOR gate
- C_{max} 2- and 3-inputs AND gates.
- V_{max} inverter protocells
- $2 \cdot V_{max}$ types of inputless *constant protocells*, outputting the constant *false* or *true* to represent the two possible values of each variable.
- a formula dependent number of wiring protocells that duplicate their input to two (or more) outputs in order to cast each constant protocell output to the appropriate AND input or inverter.

\bar{a}	\bar{b}	\bar{c}	\bar{c}_1	\bar{c}_2	\bar{c}_3	\bar{c}_4	$G(a, b, c)$
0	0	0	0	0	0	1	0
0	0	1	0	0	0	1	0
0	1	0	0	1	1	0	0
0	1	1	0	1	0	0	0
1	0	0	0	0	0	1	0
1	0	1	0	0	0	1	0
1	1	0	0	0	1	0	0
1	1	1	1	0	0	0	0

Table 4.1: Complemented value of each clause for the eight possible valuations of the variables, and the corresponding value of the formula.

Of course we can have a larger number of copies of these building blocks if we want to test more than one instance of the formula. We can remark that depending on the formula we want to test, all the C_{max} inputs of the NOR gate are not used and will stay not connected to any output, which is equivalent to a false value and so these inputs will not interfere with the computation since we are certain that nothing can be bound to them.

To verify the satisfiability of a formula made of N/V_{max} variables and C/C_{max} clauses, we need to build 2^N protocellular machine instances, (at least) one per possible valuation of the input variables. The building of these protocellular machines constitutes the second step. Although this step is specific to a given formula, its principle is generic enough to be applied to any formula. This resembles to the compilation phase of a programme written in a high level programming language on a standard computer.

To assemble a machine we will program the binding of each input of one NOR gate to the output of a 2-inputs or a 3-inputs AND gate, or to one output of a wiring protocell, or to the output of an inverter. We will also need to program the binding of one wiring protocell per variable to some inverter, AND or NOR input, according to the formula. Then, to test a given valuation of the variables, we will need to bind the constant protocells corresponding to each variable of the formula to the inputs of this machine.

These programmed bindings are made possible because all the protocells in the reservoir have been built with specific tags on their inputs and outputs. These tags can be peptides/nucleic acids with a unique sequence to address them. The process of binding itself will be done by putting in the environment specific molecular attachment instructions that recognise and bind the tag on the output and the tag on the corresponding input. This will enable the binding of specific protocells together (**Figure 4.5**).

Each input of the NOR gate is labeled with a tag implementing the number of the corresponding clause (0 to $C_{max} - 1$). Similarly the output of each of the AND gate is labeled with the same number. Therefore, to connect an AND gate to the corresponding input of the NOR gate for one protocellular machine, we have to synthesise a molecular attachment that match at one end the tag labelling the output of the AND gate and at the other end, the tag labelling the input of the NOR gate.

The same mechanism is used for the input variables of the formula. The input of a wiring protocell that corresponds to a variable of the formula is labeled with a tag representing the variable number (0 to $V_{max} - 1$). The constant protocells used for each variable, whether their output is false or true, are labeled with a tag matching the corresponding wiring protocell of the machine. Since there is a high number of constant protocells in the medium, the *false* and *true* version for each variable will be randomly bound to the corresponding input of the machines, and after some time, all the possible valuations will be covered.

It is important to notice that we must use constant protocells that output the Boolean value false, even if a non connected input is equivalent, because when we want to test a valuation where some variable is *false*, we must be certain that no *true* constant protocells can be bound to this input.

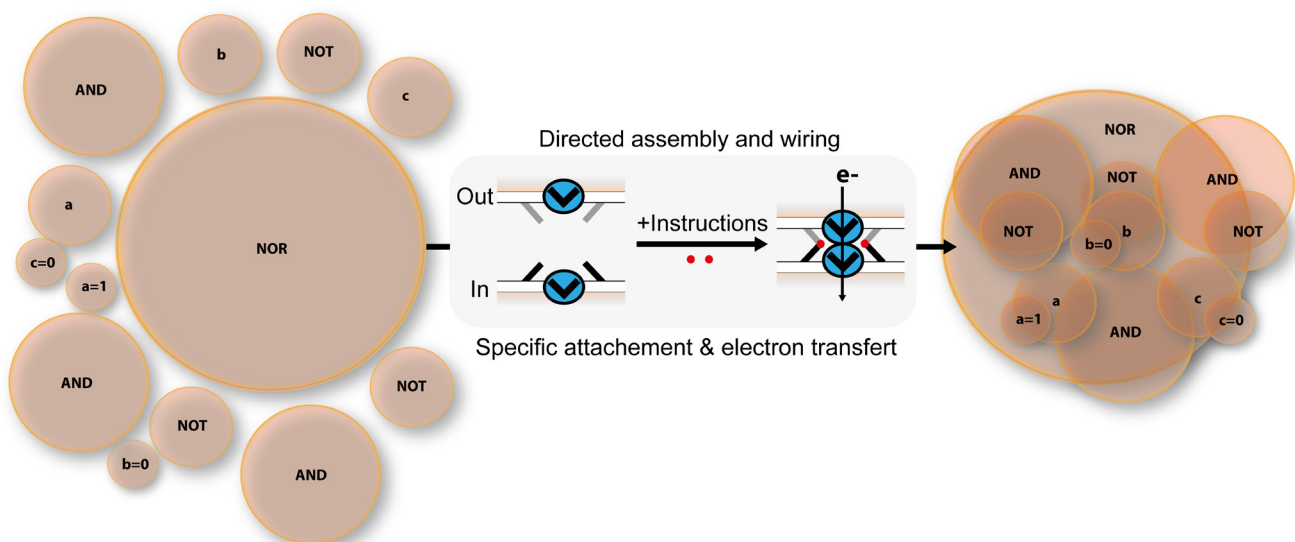


Figure 4.5: Directed assembly and wiring via specific attachments of one instance of a protocellular machine for the formula $G(a, b, c)$. The inputs a , b and c are implemented with wiring protocells (one input, one, two or more outputs) that distributes the values of the variables to inverters or to the NAND gates according to the formula (3), see **Figure 4.4B**. The NOR gate is a large protocell underneath the AND gates, where the outputs of the AND gates are bound. The small protocells a ; b ; $c = 0; 1$ are constant protocells for the input variables a , b and c (left). These input protocells will be randomly bound to constant *false* or *true* protocells to cover all the valuations of the variables. On the right side, the protocellular machine assembled tests the valuation $a = 1$; $b = 0$; $c = 0$

4.3.3 The computation process

The computation process may begin when we are certain that at least one copy of a protocellular machine is bound to each possible combination of input values. This process is started by remotely triggering the whole population of *true* constant protocells and inverter protocells using, for example, light switchable enzymes (e.g. a tryptophan dehydrogenase engineered to bear a photoswitch moiety)^{679 680 681}.

Since all the machines run concurrently to compute the value of the formula, the total computing time is the time needed either by the first one that outputs *true* (that become fluorescent) or when we can be certain that the slowest machine that outputs *false* has finished (in this case they all do). If there is a small number of protocellular machines that fluoresce, we could enhance the signal/noise ratio by scattering the solution into several parts such that the concentration of the fluorescent machines would appear higher, and so helps its detection. Another way to easily detect the first (and possibly only) protocellular machine that outputs *true* would be that this machine triggers the fluorescence of those in its neighbourhood, and so increase the global fluorescence. Independently of the formula we want to test, the maximal number of reactions needed from one input to the output is very small: one inverter, a small number of wiring protocells, one AND, and one NOR.

Considering the kinetics of enzymatic processes for these simple reactions, we could assume that the calculation time of a single protocell (i.e. the time required for effective electron transfers through the protocell) would be in the order of a few minutes. The computation time for one protocellular machine would then be proportional to the number of layers of this machine. The total computing time would not exceed 20 min, whatever the number of protocellular machines is needed to solve the problem. This is of course mainly because the computing process is massively parallel and to a lesser extent because each processor is dedicated to the specific problem we want to solve.

Since the size of a complete protocellular machine is of the order of magnitude of a micron-cube, even less, we can have more than 10^{12} machines in a few ml of solution. As 10^3 is approximately equal to 2^{10} , we could theoretically have about $2^{10(12/3)} = 2^{40}$ machines in a few ml. Therefore using this technique, we could potentially solve any 3-SAT problem involving up to 40 variables in a few minutes. If we suppose that an electronic computer needs 1 μ s to generate and test one valuation of the variables, the average computing time would be of the order of $10^{12} \cdot 10^{-6}$ s, which is more than 11 days and a half. Moreover, if we suppose we use a low power electronic computer, for example 20 watts, the energy consumed at the end of the 11.5 days would be $10^6 \cdot 20 = 2 \cdot 10^7$ J (~ 5.5 kWh), compared to a few joules for the protocellular machines.

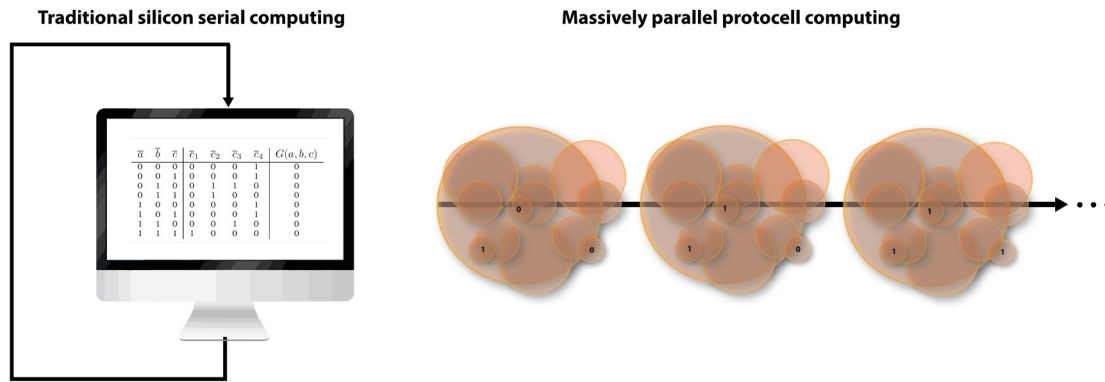


Figure 4.6: Comparison between traditional silicon based computers and proposed protocell computing.

4.4 Conclusion

The case studied here is an example of the kind of problem we could address with protocellular biocomputing machines. Here we proposed a machine assembly mechanism (i.e. computation units wiring or program instructions) that relies on highly derivatable and easily synthesizable biochemical substrate, such as nucleic acids or peptides bonds. However, thermodynamic, binding and kinetic validity of this approach needs to be investigated before reaching practical experimentation. In addition, for this reason making the very large number of instances of protocellular machines required to verify the satisfiability of a large formula is a bit speculative at the present day. Although highly theoretical, as we demonstrated in the previous chapter the governing mechanisms used to engineer protocell computation units are already under test in the lab. Many implementations of logic gates (much more than those shown in **Figure 4.5**) have been tested *in silico* using the HSIM³⁶⁸ simulation system and proven to be functioning *in vitro*.

Here, we propose to circumvent the classic silicon based approaches of serial computing, to massively parallel computing strategies taking advantage of the extremely small scale of protocellular computing devices. The computing time we claim, approximately one thousand times faster than a traditional electronic computer for a specific class and size of problem, is also a bit provocative, but the fact remains that this is an example of how to use the really massive parallelism of protocellular machines in order to solve dedicated problems (e.g. for N variables and a problem of size p the time required to evaluate all valuations: $t \sim p^N$ and $t \sim k \cdot p$, for a serial computer and a consortia of protocellular machines respectively).

Moreover, to our knowledge, this is the first described case where a synthetic biochemical computer could realistically compete with the speed of electronic computers, while being far less demanding in terms of energy. Nevertheless, in our opinion, the most exciting perspective of protocellular machines is that they are electronically and biologically interfaceable. Our approach allows us to design any given boolean function that can be connected and triggered by any biological and/or electronic input, and generate chosen outputs in a similar way. Thus they could be incorporated in living organisms, or into hybrid electronic/biological systems.

Another interesting application field involves cryptography (e.g. protocellular ciphers). This could be extended to new approaches to disease diagnosis, since a pathological state is in fact an unknown function of molecular patterns (i.e. *physiologically encrypted*), and elucidation of the specific cryptographic algorithm is analogous to finding a diagnostic algorithm.

*quas ob res ubi viderimus nil posse creari
de nihilo, tum quod sequimur iam rectius inde
perspiciemus, et unde queat res quaeque creari
et quo quaeque modo fiant opera sine divom.*

*When once we know from nothing still
Nothing can be created, we shall divine
More clearly what we seek: those elements
From which alone all things created are,
And how accomplished by no tool of Gods.*

Lucretius, *De rerum natura*, ~50 BC

Chapter 5

Final Remarks

5.1 Summary

As we have seen and discussed in Chapter 1, advances in synthetic biology enable more and more approaches to building biological devices operating with robustness and reliability. The success of synthetic biology is partly due to its exponential improvement in design capabilities capitalizing on standardized biological parts and modules and hierarchical abstraction of biological complexity. This lately enabled biotechnological bridging between medical and engineering disciplines. I envision that it will become increasingly useful as translational researchers become more familiar with the concepts and more engineering tools, modular parts and devices become widely accessible.

Building on these opportunities, in this thesis I established a foundation for engineering novel generations of cell-based and cell-like biosensors and biocomputers to tackle real world problems. These micrometer scale devices were rationally designed to interrogate their environment and process biological information according to medical rules, diagnostic accuracy, and clinical requirements. The most exciting is probably how these approaches permitted to bring rational design of biology to *in vitro* diagnostics with augmented capabilities.

In Chapter 2, I demonstrated how bioengineering solutions could bring synthetic cell-based biosensing technologies, or *bactosensors*, to operation. I demonstrated how the contribution of novel types of synthetic gene circuits, relying on the integrase based transcriptor architecture could be of outstanding interest to the field of biosensing, as it provided *true* digital processing of

biological information along with increased tunability and programmability. Transcriptor architectures thus allow for the bioengineer to focus on design and applications through the effective decoupling of design from gene circuit fabrication. Sequences of precise operations and information can be written in cells' DNA and stably stored until they are read or used. Arbitrary pathological biomarkers can be detected in clinical samples via robust and standardized engineered living cells that integrate medical expertise. While we proposed a direct application for the diagnosis of diabetes and sepsis, I envision that this will allow opportunities for the construction of ever more complex biosensors for medical diagnosis, covering an increasing range of clinical questions.

In Chapter 3, I set the basis for a totally new approach to biosensing, which relied on the biomolecular programming of specifically built protocells, *protosensors*, through integrated synthetic biochemical circuits. I demonstrated how a systematic bottom-up methodology relying on computer assisted design and microfluidics allowed for *de novo* construction of protocell devices according to specifications. This allowed for unprecedented versatility in engineering protocellular systems. From high quality biochemical standard parts, we are able to precisely program synthetic biochemical circuits to perform specific digital operation on molecular signals, using *in silico* prediction and automated exploration of design space. I validated mathematical models by constructing and testing *in vitro* several biochemical circuits in protocells. Medically programmed protosensors presented interesting analytical properties that defined them as accurate and expert diagnostic devices for the clinic.

This research constitutes a foundation to engineer more complex high order architectures from biochemical parts, but also opened up fascinating theoretical approaches to biocomputing that I presented in Chapter 4. The strategy I described takes advantage of the small scale of programmable circuits of protocellular logic gates to perform massively parallel computing, an approach that could theoretically come to competing with silicon based sequential computers on certain type of problems.

5.2 Discussion and Perspectives

Therefore, I suspect that the advances of synthetic biology could in a near future provide a new generation of expert biosensing diagnostic systems for the clinic. Indeed, diagnostics yield a great deal of information, which clinicians have to analyze and evaluate comprehensively in a short time. A few decades ago, computer sciences were first proposed to augment human reasoning in medicine⁶⁸² and permitted to enhance medical care by improving decision-making capabilities of diagnostic systems and clinicians^{683 684}. For instance, computer-aided detection and diagnosis is a procedure in medicine that assists practitioners in the interpretation of imaging techniques. Similarly, new diagnostic possibilities permitted by synthetic biology could improve clinician's ability to assess pathological states and monitor diseases and their prognosis. Diagnosis strategies fall into the definition of computing, and synthetic biology provides a modular substrate for computation and interfacing of physiology. This framework thus provides with low cost of

development, possible programming from person to person and various clinical disease presentations, and greatly shortens the concept-to-design-to-manufacturing process. This approach would be theoretically only limited by technological considerations (i.e. DNA synthesis in the case of bactorsensors, or biochemical species assembly for protosensors).

Although we did not address questions of maximal achievable sensitivities and specificities, which I assume would at the moment hardly compete with antibody based technologies, I believe that future work could be oriented towards exploiting signal amplification mechanisms discussed here for the detection of extremely low analyte concentrations. However, the rationale behind these novel devices is not to compete with gold standard clinical laboratory techniques (e.g. ELISA, PCR, MS, Chromatography, etc...), but better to provide with low resource and infrastructure requirements, portability and no apparatus requirement, ease of use, facilitated tailoring and programmability, and importantly integration of expertise (i.e. computation). In addition, while in this work we focused on portable and readily accessible diagnostic tools for non-expert usages, these devices could also be further integrated within heavy analytical machinery to multiply current assays capabilities.

As we previously discussed in the introduction, bactorsensors could benefit from new chassis with improved capabilities. Compared to *E. coli*, *Bacillus subtilis* is a promising, scalable and adaptable alternative chassis^{116 167} with a wide variety of exploitable two-component sensing systems or synthetic membrane receptors and the ability to generate spores, which can be used for long term storage and distribution. Other chassis should also be explored in order to implement all possible modalities of bactorsensors. This would permit to extend the bactorsensors approach to encompass not only *in vitro* diagnostics, but also for instance microbiome engineering, cheap and delocalized environmental monitoring, or cellular implantable theranostics. The promising bactorsensors based strategies for infectious disease diagnosis (i.e. sepsis) we presented, is the focus of ongoing work carried on to further investigate this promising application.

Nonetheless, while proving extremely valuable in certain circumstances, and benefitting from constant refinement and increase in robustness, synthetic cell-based biosensors pose intrinsic limitations such as the evolutionary barrier^{114 685} that hinder the translation into the clinics. At the moment, biosafety and regulatory concerns of self-replicating genetically engineered cells forbid their use out of a controlled *in vitro* context⁶⁸⁶, and pose biohazard risks of escape into open ecosystems⁶⁸⁷. However, expanding as an important domain of research⁶⁸⁸, environmental and health risks could be contained by rigorous risk assessment and management, and potentially reduced by methodologies such as genome minimization, metabolite dependency, encapsulation, orthogonal systems or new genetic biocontainment strategies^{689 690}. Safety, harmonized regulatory regimes, standardization, as well as appropriate future risk assessment methods⁶⁹¹ are essential catalysts to overcome these hurdles and obtain subsequent scaling up of the research⁵⁵⁴. Public, market and regulatory structure may not be ideologically ready for such dramatic change of concepts in medicine⁶⁹². Although we addressed reliability and reproducibility issue arising from the use of engineered biological systems in clinical settings, development has yet to address safety and regulatory issues prior to effective use in medical applications. In addition, commercial interest in biosensor technology remains hampered by the legislation controlling the application

of genetically engineered bacteria, and by the need to overcome these technical problems inherent to living organisms¹¹⁷.

In consequence, amongst the wide range of different biological substrates, cell-free devices could hold most clinical promises. Between the two approaches to synthetic biology, top-down and bottom-up, the latter is probably more relevant to diagnostics as it provides more flexibility and highest control on properties and could more easily be cleared for regulatory approval.

However, cell-based and minimal biosensors are not optimized to answer the same technological questions. While engineered cells prove very valuable and robust in certain contexts, additionally benefitting from self assembling/replicating mechanisms, protosensors could show poor operability, and *vice versa*. Nevertheless, a bottom-up approach ultimately may offer a greater design space to the price of having to deal with greater complexity. On a bigger picture, I propose that the two widely described synthetic biology approaches, so called top-down and bottom-up, should therefore not be seen as in opposition but more synergistic (**Figure 5.1**). Therefore, I believe that future interesting developments are constituted by *hybrid* approaches (i.e. so called *middle-out*).

In addition, this work gave perspectives on the advantages and inconvenient both brought by synthetic gene and biochemical circuits, which I believe are interesting to denote. The Transcriptor architecture we used for digital biosensing and information processing exploits the versatility of DNA synthesis to achieve high programmability. Writing information in DNA sequences has become incredibly straightforward. However, it can suffer from unfavorable kinetics and subsequent response time compared with biochemical circuits. While Transcriptor based circuits go through long transcription-translation processes, biochemical circuits rely on post-translational regulation only and thus display faster kinetics. Besides, while abstractions and circuit construction can be straightforward with synthetic gene circuits, intrinsic complexity of biochemical circuits make them technologically more demanding.

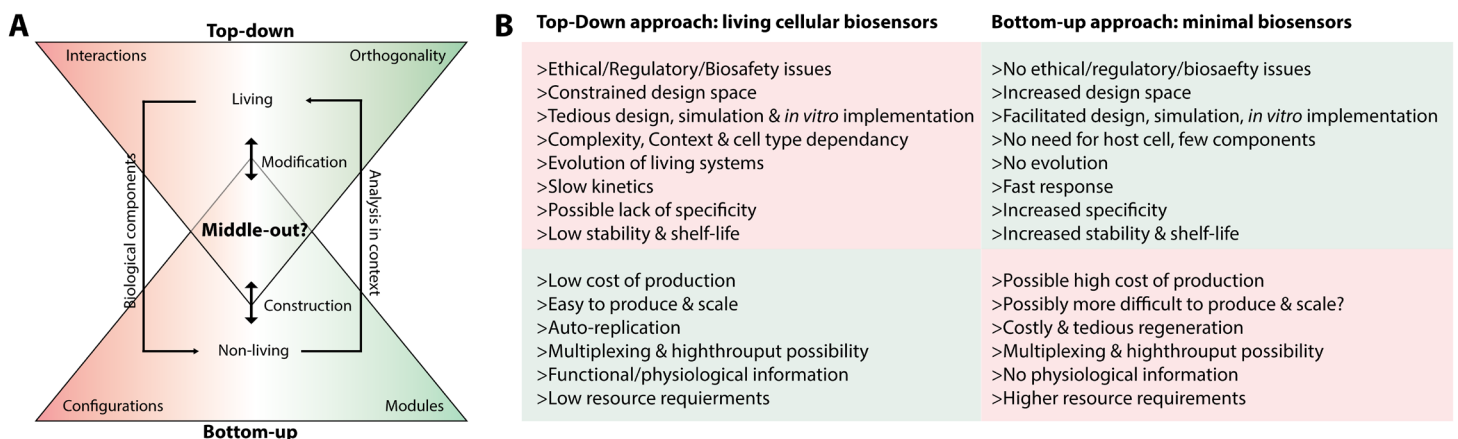


Figure 5.1: Considerations on the so called Top-down and Bottom-up approaches used in this thesis. (A) Abstract representation of the two approaches and their respective engineering process. They appear synergistic in understanding and engineering biological complexity (B) Advantages and drawbacks of top-down versus bottom-up approaches to synthetic biology for the developments of integrated biosensors for medical diagnosis.

Going further, an interesting perspective could be to use synthetic gene circuits *ex vivo*, as recently described¹⁰³. In fact, the transcriptor based architecture could be well accommodated in standardized purified minimal cell extracts, as unnecessary and complex cell machinery imposes an energetic burden on circuits functioning. Moreover, selective antibiotic pressures imposed on bactorsensors, as well as cultivation steps that pose problem towards application would become unnecessary. By that mean we could also bypass ethical and regulatory hurdles. *Ex vivo* operation of bactorsensors will thus be explored, and appears as a very promising approach for certain applications.

Another next interesting step would be to adapt synthetic circuits for sequence-based detection of nucleic acid input biomarkers, such as drug resistance genes, circulating tumor DNA, or viral charge parameters. This should be a relatively low-hanging fruit accessible for instance using strand displacement cascades²⁷⁸ coupled to Transcriptor operation or RNA/DNAzymes biochemical circuits.

The systematic approaches to synthetic biochemical circuit engineering coupled to computer assisted design offers tremendous possibilities to program protocellular machines. It will give the opportunity for future work to create increasingly sophisticated and complex systems. The approach we proposed here might constitute an interesting step forward for the investigation of *de novo* autopoietic mechanisms. Although a still immensely difficult and somehow provocative task lies ahead, I envision that the approaches discussed in Chapter 3 could benefit the field of protocell and origin of life research.

Likewise, imagination appears as the real limit to applications of protosensors, as we now seek to construct new prototypes for various disease diagnosis. We also explore the technological modalities of integrating these devices into novel biosensing formats such as bioMEMS. Surface functionalization of protosensors constitutes an important aspect of the next improvements to bring, since it would enable space patterning and *in vivo* addressing, hence giving access to a large landscape of applications. *In situ* biomolecular sensing and logic is thus an interesting perspective that will be explored.

While my work investigated biological signal processing to implement decision making mostly relied on digital Boolean architecture, synthetic biosensing devices could benefit from novel biocomputing frameworks. Indeed, increasing evidence tend to show that biological signal processing exploit hybrid analog-digital architectures with greater precision. Indeed, digital computation is a subset of analog computation that operates at its highest or lowest extremes⁴⁷⁴, and although digital design is straightforward and scalable on silicon substrate, this might not be the case on biological substrate. Analog computing enables finer and more efficient computational modes with lower energy, time, space, molecular concentration requirements. For example, biomarker concentration could be integrated by analog systems more efficiently and with wide-dynamic range biosensing for certain type of operation, such as multivaluation, addition, subtraction and log-domain analysis of pathological inputs, compared to digital assessment of presence/absence. Furthermore, not only coping with noise and stochasticity, biological systems are capable of exploiting these properties to their advantage, to achieve most efficient signal processing. Therefore, fully harnessing and understanding how biological

structures perform biocomputing in future projects, is likely to bring even more powerful and scalable synthetic biosystems for practical applications. The design of analog synthetic gene and biochemical logic circuits will, however, require novel standards and abstractions. In addition, other theoretical biocomputing frameworks relying on rationally designed protocells could be envisaged, and theoretical work is underway.

We explored in Chapter 2 the coupling of synthetic biology with microencapsulation technologies. This has already been described to generate innovative cell-based biomedical applications, such as *in vitro* diagnostic formats or smart implantable theranostics⁶⁹³. Cell-based biosensors encapsulation and immobilization is a promising technological evolution enabling sealing of engineered cells into portable, easy to handle formats, which provide suitable extracellular environment, semi-permeable and biocompatible microcapsule without the need of culture facilities. Microencapsulation can also be used to develop cell microarrays suitable for simultaneous measurement of a large number of samples. In the future, various polymeric materials could be engineered at the nanoscale with control on biophysical properties and spatial patterning to enhance robustness and reliability of encapsulated synthetic cellular systems⁶⁹⁴. New nanofabrication technologies and synthetic biology approaches are likely to lead to new prospects for developing devices with tailored functionalities¹⁷².

Last but not least, it is interesting to take into consideration the industrial landscape to fully comprehend the potential evolution of synthetic biosensors for diagnostics. While the global value of synthetic biology market is expected to expand and reach \$16 billion by 2018, the market in diagnostics and pharmaceutical industry has been evaluated around \$5 billion in 2016, appearing as the most important industry driving innovation amongst chemicals, R&D, agriculture, and energy⁶⁹⁵. The market growth for biosensors is exploding, with medical sensors global market is expected to reach \$15 billion in 2019, with a growth of 6.3% from 2013 to 2019⁶⁹⁶. These economic considerations will play an increasingly important role in a biomedical context, if synthetic biology is to offer simpler, more elegant and least expensive solutions more likely to be clinically successful.

Even though biocomputational versions of diagnosis using biological components have been proposed, there might still be a long way to go until synthetic biology based biomedical devices become a wide spread clinical reality. However, this work is another stepping stone towards biologically encoded medical tools, and will contribute to increase the pace at which synthetic biosystems can be built according to medical needs, and expands the role of synthetic biology for global health.

Chapter 6

Bibliography

1. Weber, W. & Fussenegger, M. Emerging biomedical applications of synthetic biology. *Nat. Rev. Genet.* (2011). doi:10.1038/nrg3094
2. Arkin, A. P. & Schaffer, D. V. Network News: Innovations in 21st Century Systems Biology. *Cell* **144**, 844–849 (2011).
3. Smolke, C. D. & Silver, P. A. Informing Biological Design by Integration of Systems and Synthetic Biology. *Cell* **144**, 855–859 (2011).
4. Jasny, B. R. & Zahn, L. M. A Celebration of the Genome, Part I. *Science* **331**, 546–546 (2011).
5. NEST High-Level Expert Group. *Synthetic Biology, Applying Engineering to Biology*. (European Commission Report). at <<http://www.synbiosafe.eu/uploads///pdf/EU-highlevel-syntheticbiology.pdf>>
6. Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* **10**, 410–422 (2009).
7. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **26**, 787–793 (2008).
8. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
9. Smolke, C. D. Building outside of the box: iGEM and the BioBricks Foundation. *Nat. Biotechnol.* **27**, 1099–1102 (2009).
10. Annaluru, N. *et al.* Total Synthesis of a Functional Designer Eukaryotic Chromosome. *Science* **344**, 55–58 (2014).
11. Gibson, D. G. *et al.* Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **329**, 52–56 (2010).
12. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
13. Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013).
14. Olson, E. J. & Tabor, J. J. Optogenetic characterization methods overcome key challenges in synthetic and systems biology. *Nat. Chem. Biol.* **10**, 502–511 (2014).

15. Barrett, C. L., Kim, T. Y., Kim, H. U., Palsson, B. Ø. & Lee, S. Y. Systems biology as a foundation for genome-scale synthetic biology. *Curr. Opin. Biotechnol.* **17**, 488–492 (2006).
16. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
17. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
18. Friedland, A. E. *et al.* Synthetic Gene Networks That Count. *Science* **324**, 1199–1202 (2009).
19. You, L., Cox, R. S., Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
20. Wang, L. *et al.* SynBioLGDB: a resource for experimentally validated logic gates in synthetic biology. *Sci. Rep.* **5**, 8090 (2015).
21. Yang, L. *et al.* Permanent genetic memory with >1-byte capacity. *Nat. Methods* **11**, 1261–1266 (2014).
22. Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C. & Voigt, C. A. Genetic programs constructed from layered logic gates in single cells. *Nature* **491**, 249–253 (2012).
23. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372 (2011).
24. Brophy, J. A. N. & Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520 (2014).
25. Dymond, J. S. *et al.* Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**, 471–476 (2011).
26. Mukherji, S. & van Oudenaarden, A. Synthetic biology: understanding biological design from synthetic circuits. *Nat. Rev. Genet.* (2009). doi:10.1038/nrg2697
27. Pinheiro, V. B. *et al.* Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **336**, 341–344 (2012).
28. Lang, K. & Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **114**, 4764–4806 (2014).
29. Johnson, R. Xeno-nucleic acids: Unnatural biocatalysts. *Nat. Chem.* **7**, 94–94 (2015).
30. Bornscheuer, U. T. *et al.* Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
31. Adamala, K. & Szostak, J. W. Nonenzymatic Template-Directed RNA Synthesis Inside Model Protocells. *Science* **342**, 1098–1100 (2013).
32. Zhang, S., Blain, J. C., Zielinska, D., Gryaznov, S. M. & Szostak, J. W. Fast and accurate nonenzymatic copying of an RNA-like synthetic genetic polymer. *Proc. Natl. Acad. Sci.* **110**, 17732–17737 (2013).
33. Lincoln, T. A. & Joyce, G. F. Self-Sustained Replication of an RNA Enzyme. *Science* **323**, 1229–1232 (2009).

34. Schmidt, C. W. Synthetic Biology: Environmental Health Implications of a New Field. *Environ. Health Perspect.* **118**, a118–a123 (2010).
35. Tyagi, A. *et al.* Synthetic Biology: Applications in Food Sector. *Crit. Rev. Food Sci. Nutr.* 00–00 (2014). doi:10.1080/10408398.2013.782534
36. Zhang, W. & Nielsen, D. R. Synthetic biology applications in industrial microbiology. *Front. Microbiol.* **5**, (2014).
37. Keasling, J. D. Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* **14**, 189–195 (2012).
38. Hörner, M., Reischmann, N. & Weber, W. Synthetic Biology: Programming Cells for Biomedical Applications. *Perspect. Biol. Med.* **55**, 490–502 (2012).
39. Benenson, Y. Biomolecular computing systems: principles, progress and potential. *Nat. Rev. Genet.* **13**, 455–468 (2012).
40. Keret, O. Biomedical synthetic biology: an overview for physicians. *Isr. Med. Assoc. J. IMAJ* **15**, 308–312 (2013).
41. Kotula, J. W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci.* **111**, 4838–4843 (2014).
42. De Gregorio, E. & Rappuoli, R. From empiricism to rational design: a personal perspective of the evolution of vaccine development. *Nat. Rev. Immunol.* **14**, 505–514 (2014).
43. Thaker, M. N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
44. DeLoache, W. C. *et al.* An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat. Chem. Biol.* (2015). doi:10.1038/nchembio.1816
45. Thaker, M. N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
46. Ajikumar, P. K. *et al.* Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **330**, 70–74 (2010).
47. Thodey, K., Galanie, S. & Smolke, C. D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* (2014). doi:10.1038/nchembio.1613
48. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
49. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095–1100 (2015).
50. Walsh, G. M. **Biotherapeutics: Recent Developments using Chemical and Molecular Biology**. Edited by Lyn H. Jones, Andrew J. McKnight. *ChemMedChem* **9**, 2814–2814 (2014).
51. Rim, Y. A. *et al.* Self in vivo production of a synthetic biological drug CTLA4Ig using a minicircle vector. *Sci. Rep.* **4**, 6935 (2014).

52. Bryksin, A. V., Brown, A. C., Baksh, M. M., Finn, M. G. & Barker, T. H. Learning from nature – Novel synthetic biology approaches for biomaterial design. *Acta Biomater.* **10**, 1761–1769 (2014).
53. Moore, R. *et al.* CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells. *Nucleic Acids Res.* **43**, 1297–1303 (2015).
54. Gabrieli, P., Smidler, A. & Catteruccia, F. Engineering the control of mosquito-borne infectious diseases. *Genome Biol.* **15**, 535 (2014).
55. Chakravarti, D. & Wong, W. W. Synthetic biology in cell-based cancer immunotherapy. *Trends Biotechnol.* **33**, 449–461 (2015).
56. Ruder, W. C., Lu, T. & Collins, J. J. Synthetic Biology Moving into the Clinic. *Science* **333**, 1248–1252 (2011).
57. Nielsen, J. *et al.* Engineering synergy in biotechnology. *Nat. Chem. Biol.* **10**, 319–322 (2014).
58. Kis, Z., Pereira, H. S., Homma, T., Pedrigi, R. M. & Krams, R. Mammalian synthetic biology: emerging medical applications. *J. R. Soc. Interface* **12**, 20141000–20141000 (2015).
59. Heng, B. C. & Fussenegger, M. in *Encyclopedia of Molecular Cell Biology and Molecular Medicine* (ed. Meyers, R. A.) 1–17 (Wiley-VCH Verlag GmbH & Co. KGaA, 2014). at <<http://doi.wiley.com/10.1002/3527600906.mcb.20120067>>
60. Planson, A.-G., Carbonell, P., Grigoras, I. & Faulon, J.-L. A retrosynthetic biology approach to therapeutics: from conception to delivery. *Curr. Opin. Biotechnol.* **23**, 948–956 (2012).
61. Chen, Y. Y. & Smolke, C. D. From DNA to Targeted Therapeutics: Bringing Synthetic Biology to the Clinic. *Sci. Transl. Med.* **3**, 106ps42–106ps42 (2011).
62. Elowitz, M. & Lim, W. A. Build life to understand it. *Nature* **468**, 889–890 (2010).
63. Bashor, C. J., Horwitz, A. A., Peisajovich, S. G. & Lim, W. A. Rewiring Cells: Synthetic Biology as a Tool to Interrogate the Organizational Principles of Living Systems. *Annu. Rev. Biophys.* **39**, 515–537 (2010).
64. Quo, C. F. *et al.* Reverse engineering biomolecular systems using -omic data: challenges, progress and opportunities. *Brief. Bioinform.* **13**, 430–445 (2012).
65. Sample, V., Mehta, S. & Zhang, J. Genetically encoded molecular probes to visualize and perturb signaling dynamics in living biological systems. *J. Cell Sci.* **127**, 1151–1160 (2014).
66. Yagi, H. *et al.* A Synthetic Biology Approach Reveals a CXCR4-G13-Rho Signaling Axis Driving Transendothelial Migration of Metastatic Breast Cancer Cells. *Sci. Signal.* **4**, ra60–ra60 (2011).
67. Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77–80 (2005).
68. Becker, M. M. *et al.* Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc. Natl. Acad. Sci.* **105**, 19944–19949 (2008).
69. Cello, J., Paul, A. V. & Wimmer, E. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297**, 1016–1018 (2002).

70. Berthet, N. *et al.* Reconstructed Ancestral Sequences Improve Pathogen Identification Using Resequencing DNA Microarrays. *PLoS ONE* **5**, e15243 (2010).
71. Burbelo, P. D., Ching, K. H., Bush, E. R., Han, B. L. & Iadarola, M. J. Antibody-profiling technologies for studying humoral responses to infectious agents. *Expert Rev. Vaccines* **9**, 567–578 (2010).
72. Burbelo, P. D. *et al.* LIPS arrays for simultaneous detection of antibodies against partial and whole proteomes of HCV, HIV and EBV. *Mol. Biosyst.* **7**, 1453 (2011).
73. Feng, H., Shuda, M., Chang, Y. & Moore, P. S. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319**, 1096–1100 (2008).
74. Burbelo, P. D., Ching, K. H., Bush, E. R., Han, B. L. & Iadarola, M. J. Antibody-profiling technologies for studying humoral responses to infectious agents. *Expert Rev. Vaccines* **9**, 567–578 (2010).
75. Gómara, M. J. & Haro, I. Synthetic peptides for the immunodiagnosis of human diseases. *Curr. Med. Chem.* **14**, 531–546 (2007).
76. Burbelo, P. D. *et al.* Rapid, Simple, Quantitative, and Highly Sensitive Antibody Detection for Lyme Disease. *Clin. Vaccine Immunol.* **17**, 904–909 (2010).
77. Meloen, R. H., Puijk, W. C., Langeveld, J. P. M., Langedijk, J. P. M. & Timmerman, P. Design of synthetic peptides for diagnostics. *Curr. Protein Pept. Sci.* **4**, 253–260 (2003).
78. Larman, H. B. *et al.* Autoantigen discovery with a synthetic human peptidome. *Nat. Biotechnol.* **29**, 535–541 (2011).
79. Xu, G. J. *et al.* Comprehensive serological profiling of human populations using a synthetic human virome. *Science* **348**, aaa0698–aaa0698 (2015).
80. Wine, Y., Horton, A. P., Ippolito, G. C. & Georgiou, G. Serology in the 21st century: the molecular-level analysis of the serum antibody repertoire. *Curr. Opin. Immunol.* **35**, 89–97 (2015).
81. Geering, B. & Fussenegger, M. Synthetic immunology: modulating the human immune system. *Trends Biotechnol.* **33**, 65–79 (2015).
82. Warren, A. D., Kwong, G. A., Wood, D. K., Lin, K. Y. & Bhatia, S. N. Point-of-care diagnostics for noncommunicable diseases using synthetic urinary biomarkers and paper microfluidics. *Proc. Natl. Acad. Sci.* **111**, 3671–3676 (2014).
83. Kwong, G. A. *et al.* Mass-encoded synthetic biomarkers for multiplexed urinary monitoring of disease. *Nat. Biotechnol.* **31**, 63–70 (2013).
84. Warren, A. D. *et al.* Disease detection by ultrasensitive quantification of microdosed synthetic urinary biomarkers. *J. Am. Chem. Soc.* **136**, 13709–13714 (2014).
85. Lin, K. Y., Kwong, G. A., Warren, A. D., Wood, D. K. & Bhatia, S. N. Nanoparticles That Sense Thrombin Activity As Synthetic Urinary Biomarkers of Thrombosis. *ACS Nano* **7**, 9001–9009 (2013).
86. Hay Burgess, D. C., Wasserman, J. & Dahl, C. A. Global health diagnostics. *Nature* **444**, 1–2 (2006).
87. Garcia, P. J., You, P., Fridley, G., Mabey, D. & Peeling, R. Point-of-care diagnostic tests for low-resource settings. *Lancet Glob. Health* **3**, e257–e258 (2015).

88. Walt, D. R. CHEMISTRY: Miniature Analytical Methods for Medical Diagnostics. *Science* **308**, 217–219 (2005).
89. Hood, L. Systems Biology and New Technologies Enable Predictive and Preventative Medicine. *Science* **306**, 640–643 (2004).
90. Giljohann, D. A. & Mirkin, C. A. Drivers of biodiagnostic development. *Nature* **462**, 461–464 (2009).
91. Turner, A. P. F. Biosensors: sense and sensibility. *Chem. Soc. Rev.* **42**, 3184 (2013).
92. Ruder, W. C., Lu, T. & Collins, J. J. Synthetic Biology Moving into the Clinic. *Science* **333**, 1248–1252 (2011).
93. Rooke, J. Synthetic biology as a source of global health innovation. *Syst. Synth. Biol.* **7**, 67–72 (2013).
94. Xie, M. & Fussenegger, M. Mammalian designer cells: Engineering principles and biomedical applications. *Biotechnol. J.* n/a–n/a (2015). doi:10.1002/biot.201400642
95. *Fundamentals of clinical practice.* (Kluwer Academic/Plenum Publishers, 2002).
96. Courbet, A., Endy, D., Renard, E., Molina, F. & Bonnet, J. Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci. Transl. Med.* **7**, 289ra83–289ra83 (2015).
97. Ausländer, D. *et al.* A designer cell-based histamine-specific human allergy profiler. *Nat. Commun.* **5**, (2014).
98. Danino, T. *et al.* Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med.* **7**, 289ra84–289ra84 (2015).
99. Lu, T. K., Bowers, J. & Koeris, M. S. Advancing bacteriophage-based microbial diagnostics with synthetic biology. *Trends Biotechnol.* **31**, 325–327 (2013).
100. Douglas, S. M., Bachelet, I. & Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **335**, 831–834 (2012).
101. You, M. *et al.* DNA ‘Nano-Claw’: Logic-Based Autonomous Cancer Targeting and Therapy. *J. Am. Chem. Soc.* **136**, 1256–1259 (2014).
102. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333**, 1307–1311 (2011).
103. Pardee, K. *et al.* Paper-Based Synthetic Gene Networks. *Cell* **159**, 940–954 (2014).
104. Griss, R. *et al.* Bioluminescent sensor proteins for point-of-care therapeutic drug monitoring. *Nat. Chem. Biol.* **10**, 598–603 (2014).
105. Barrett, D. M., Singh, N., Porter, D. L., Grupp, S. A. & June, C. H. Chimeric Antigen Receptor Therapy for Cancer. *Annu. Rev. Med.* **65**, 333–347 (2014).

106. Kojima, R., Aubel, D. & Fussenegger, M. Novel theranostic agents for next-generation personalized medicine: small molecules, nanoparticles, and engineered mammalian cells. *Curr. Opin. Chem. Biol.* **28**, 29–38 (2015).
107. Steinhubl, S. R., Muse, E. D. & Topol, E. J. The emerging field of mobile health. *Sci. Transl. Med.* **7**, 283rv3–283rv3 (2015).
108. Folcher, M. *et al.* Mind-controlled transgene expression by a wireless-powered optogenetic designer cell implant. *Nat. Commun.* **5**, 5392 (2014).
109. Salis, H., Tamsir, A. & Voigt, C. Engineering bacterial signals and sensors. *Contrib. Microbiol.* **16**, 194–225 (2009).
110. Raut, N., O'Connor, G., Pasini, P. & Daunert, S. Engineered cells as biosensing systems in biomedical analysis. *Anal. Bioanal. Chem.* **402**, 3147–3159 (2012).
111. Marchisio, M. A. & Rudolf, F. Synthetic biosensing systems. *Int. J. Biochem. Cell Biol.* **43**, 310–319 (2011).
112. Yagi, K. Applications of whole-cell bacterial sensors in biotechnology and environmental science. *Appl. Microbiol. Biotechnol.* **73**, 1251–1258 (2007).
113. Liu, Q. *et al.* Cell-Based Biosensors and Their Application in Biomedicine. *Chem. Rev.* **114**, 6423–6461 (2014).
114. Michelini, E. *et al.* Field-deployable whole-cell bioluminescent biosensors: so near and yet so far. *Anal. Bioanal. Chem.* **405**, 6155–6163 (2013).
115. Kobayashi, H. *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8414–8419 (2004).
116. *The science and applications of synthetic and systems biology: workshop summary.* (National Academies Press, 2011).
117. van der Meer, J. R. & Belkin, S. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nat. Rev. Microbiol.* **8**, 511–522 (2010).
118. Park, M., Tsai, S.-L. & Chen, W. Microbial Biosensors: Engineered Microorganisms as the Sensing Machinery. *Sensors* **13**, 5777–5795 (2013).
119. Wang, B., Barahona, M. & Buck, M. A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. *Biosens. Bioelectron.* **40**, 368–376 (2013).
120. Checa, S. K., Zurbriggen, M. D. & Soncini, F. C. Bacterial signaling systems as platforms for rational design of new generations of biosensors. *Curr. Opin. Biotechnol.* **23**, 766–772 (2012).
121. Daunert, S. *et al.* Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chem. Rev.* **100**, 2705–2738 (2000).
122. Hernandez Espitia, C. A. & Osma, J. F. in *Biosensors: Recent advances and mathematical challenges* (eds. Osma, J. F. & Stoytcheva, M.) 51–96 (OmniaScience, 2014). at <http://omniascience.com/monographs/index.php/monograficos/issue/view/15/showToc>

123. Campàs, M., Carpentier, R. & Rouillon, R. Plant tissue- and photosynthesis-based biosensors. *Biotechnol. Adv.* **26**, 370–378 (2008).
124. Brayner, R., Couté, A., Livage, J., Perrette, C. & Sicard, C. Micro-algal biosensors. *Anal. Bioanal. Chem.* **401**, 581–597 (2011).
125. Banerjee, P., Franz, B. & Bhunia, A. K. in *Whole Cell Sensing Systems I* (eds. Belkin, S. & Gu, M. B.) 21–55 (Springer Berlin Heidelberg, 2010). at <http://link.springer.com/10.1007/10_2009_21>
126. Ausländer, S. & Fussenegger, M. From gene switches to mammalian designer cells: present and future prospects. *Trends Biotechnol.* **31**, 155–168 (2013).
127. Walmsley, R. M. & Keenan, P. The eukaryote alternative: Advantages of using yeasts in place of bacteria in microbial biosensor development. *Biotechnol. Bioprocess Eng.* **5**, 387–394 (2000).
128. Su, L., Jia, W., Hou, C. & Lei, Y. Microbial biosensors: A review. *Biosens. Bioelectron.* **26**, 1788–1799 (2011).
129. Stenger, D. A. *et al.* Detection of physiologically active compounds using cell-based biosensors. *Trends Biotechnol.* **19**, 304–309 (2001).
130. Banerjee, P. & Bhunia, A. K. Cell-based biosensor for rapid screening of pathogens and toxins. *Biosens. Bioelectron.* **26**, 99–106 (2010).
131. Hung, P. J., Lee, P. J., Sabounchi, P., Lin, R. & Lee, L. P. Continuous perfusion microfluidic cell culture array for high-throughput cell-based assays. *Biotechnol. Bioeng.* **89**, 1–8 (2005).
132. Tani, H., Maehana, K. & Kamidate, T. Chip-based bioassay using bacterial sensor strains immobilized in three-dimensional microfluidic network. *Anal. Chem.* **76**, 6693–6697 (2004).
133. El-Ali, J., Sorger, P. K. & Jensen, K. F. Cells on chips. *Nature* **442**, 403–411 (2006).
134. Melamed, S., Elad, T. & Belkin, S. Microbial sensor cell arrays. *Curr. Opin. Biotechnol.* **23**, 2–8 (2012).
135. Ben-Yoav, H., Melamed, S., Freeman, A., Shacham-Diamand, Y. & Belkin, S. Whole-cell biochips for bio-sensing: integration of live cells and inanimate surfaces. *Crit. Rev. Biotechnol.* **31**, 337–353 (2011).
136. *Cell-based biosensors: principles and applications.* (Artech House, 2009).
137. Prindle, A. *et al.* A sensing array of radically coupled genetic ‘biopixels’. *Nature* **481**, 39–44 (2011).
138. Gast, T. Sensors with oscillating elements. *J. Phys. [E]* **18**, 783–789 (1985).
139. Turner, K. *et al.* Hydroxylated polychlorinated biphenyl detection based on a genetically engineered bioluminescent whole-cell sensing system. *Anal. Chem.* **79**, 5740–5745 (2007).
140. Watstein, D. M., McNerney, M. P. & Styczynski, M. P. Precise metabolic engineering of carotenoid biosynthesis in *Escherichia coli* towards a low-cost biosensor. *Metab. Eng.* **31**, 171–180 (2015).
141. Lewis, C. *et al.* Novel use of a whole cell *E. coli* bioreporter as a urinary exposure biomarker. *Environ. Sci. Technol.* **43**, 423–428 (2009).

142. Horswell, J. & Dickson, S. Use of biosensors to screen urine samples for potentially toxic chemicals. *J. Anal. Toxicol.* **27**, 372–376 (2003).
143. Bahl, M. I., Hansen, L. H., Licht, T. R. & Sørensen, S. J. In vivo detection and quantification of tetracycline by use of a whole-cell biosensor in the rat intestine. *Antimicrob. Agents Chemother.* **48**, 1112–1117 (2004).
144. Hansen, L. H., Aarestrup, F. & Sørensen, S. J. Quantification of bioavailable chlortetracycline in pig feces using a bacterial whole-cell biosensor. *Vet. Microbiol.* **87**, 51–57 (2002).
145. Kurittu, J., Lönnberg, S., Virta, M. & Karp, M. Qualitative detection of tetracycline residues in milk with a luminescence-based microbial method: the effect of milk composition and assay performance in relation to an immunoassay and a microbial inhibition assay. *J. Food Prot.* **63**, 953–957 (2000).
146. Paton, G. I., Reid, B. J. & Semple, K. T. Application of a luminescence-based biosensor for assessing naphthalene biodegradation in soils from a manufactured gas plant. *Environ. Pollut.* **157**, 1643–1648 (2009).
147. Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell* **159**, 925–939 (2014).
148. Muranaka, N., Sharma, V., Nomura, Y. & Yokobayashi, Y. Efficient Design Strategy for Whole-Cell and Cell-Free Biosensors based on Engineered Riboswitches. *Anal. Lett.* **42**, 108–122 (2009).
149. Marvin, J. S., Schreiter, E. R., Echevarría, I. M. & Looger, L. L. A genetically encoded, high-signal-to-noise maltose sensor. *Proteins Struct. Funct. Bioinforma.* **79**, 3025–3036 (2011).
150. Jeffery, C. J. Engineering periplasmic ligand binding proteins as glucose nanosensors. *Nano Rev.* **2**, (2011).
151. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185–190 (2003).
152. Jo, J.-J. & Shin, J.-S. Construction of intragenic synthetic riboswitches for detection of a small molecule. *Biotechnol. Lett.* **31**, 1577–1581 (2009).
153. Win, M. N. & Smolke, C. D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14283–14288 (2007).
154. Baker, J. L. *et al.* Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science* **335**, 233–235 (2012).
155. Qi, L., Lucks, J. B., Liu, C. C., Mutalik, V. K. & Arkin, A. P. Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals. *Nucleic Acids Res.* **40**, 5775–5786 (2012).
156. Lucks, J. B., Qi, L., Mutalik, V. K., Wang, D. & Arkin, A. P. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc. Natl. Acad. Sci.* **108**, 8617–8622 (2011).
157. Ninfa, A. J. *et al.* Using two-component systems and other bacterial regulatory factors for the fabrication of synthetic genetic devices. *Methods Enzymol.* **422**, 488–512 (2007).
158. de los Santos, E. L. C., Meyerowitz, J. T., Mayo, S. L. & Murray, R. M. *Engineering Transcriptional Regulator Effector Specificity using Computational Design and In Vitro Rapid Prototyping: Developing a Vanillin Sensor.* (2015). at <<http://biorxiv.org/lookup/doi/10.1101/015438>>

159. Wang, B., Barahona, M. & Buck, M. Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. *Nucleic Acids Res.* **43**, 1955–1964 (2015).
160. Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci.* **102**, 12678–12683 (2005).
161. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-Sequence Determinants of Gene Expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
162. Roda, A. *et al.* SENSITIVE DETERMINATION OF URINARY MERCURY(II) BY A BIOLUMINESCENT TRANSGENIC BACTERIA-BASED BIOSENSOR. *Anal. Lett.* **34**, 29–41 (2001).
163. Skerker, J. M. *et al.* Rewiring the specificity of two-component signal transduction systems. *Cell* **133**, 1043–1054 (2008).
164. Xu, J. & Lavan, D. A. Designing artificial cells to harness the biological ion concentration gradient. *Nat. Nanotechnol.* **3**, 666–670 (2008).
165. Booth, I. R., Edwards, M. D., Black, S., Schumann, U. & Miller, S. Mechanosensitive channels in bacteria: signs of closure? *Nat. Rev. Microbiol.* **5**, 431–440 (2007).
166. Danchin, A. Scaling up synthetic biology: Do not forget the chassis. *FEBS Lett.* **586**, 2129–2137 (2012).
167. Harwood, C. R., Pohl, S., Smith, W. & Wipat, A. in *Methods in Microbiology* **40**, 87–117 (Elsevier, 2013).
168. Nikel, P. I., Martínez-García, E. & de Lorenzo, V. Biotechnological domestication of pseudomonads using synthetic biology. *Nat. Rev. Microbiol.* **12**, 368–379 (2014).
169. Knecht, L. D., Pasini, P. & Daunert, S. Bacterial spores as platforms for bioanalytical and biomedical applications. *Anal. Bioanal. Chem.* **400**, 977–989 (2011).
170. Date, A., Pasini, P. & Daunert, S. Integration of spore-based genetically engineered whole-cell sensing systems into portable centrifugal microfluidic platforms. *Anal. Bioanal. Chem.* **398**, 349–356 (2010).
171. Rotman, B. & Cote, M. A. Application of a real-time biosensor to detect bacteria in platelet concentrates. *Biochem. Biophys. Res. Commun.* **300**, 197–200 (2003).
172. Michelini, E. & Roda, A. Staying alive: new perspectives on cell immobilization for biosensing purposes. *Anal. Bioanal. Chem.* **402**, 1785–1797 (2012).
173. Struss, A., Pasini, P., Ensor, C. M., Raut, N. & Daunert, S. Paper strip whole cell biosensors: a portable test for the semiquantitative detection of bacterial quorum signaling molecules. *Anal. Chem.* **82**, 4457–4463 (2010).
174. Park, S. J. *et al.* New paradigm for tumor theranostic methodology using bacteria-based microrobot. *Sci. Rep.* **3**, (2013).
175. Chen, A. Y. *et al.* Synthesis and patterning of tunable multiscale materials with engineered cells. *Nat. Mater.* **13**, 515–523 (2014).
176. Bacchus, W., Aubel, D. & Fussenegger, M. Biomedically relevant circuit-design strategies in mammalian synthetic biology. *Mol. Syst. Biol.* **9**, 691–691 (2014).

177. Lienert, F., Lohmueller, J. J., Garg, A. & Silver, P. A. Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nat. Rev. Mol. Cell Biol.* **15**, 95–107 (2014).
178. Stanton, B. C. *et al.* Systematic Transfer of Prokaryotic Sensors and Circuits to Mammalian Cells. *ACS Synth. Biol.* **3**, 880–891 (2014).
179. Daringer, N. M., Dudek, R. M., Schwarz, K. A. & Leonard, J. N. Modular Extracellular Sensor Architecture for Engineering Mammalian Cell-based Devices. *ACS Synth. Biol.* **3**, 892–902 (2014).
180. Fung, E. *et al.* A synthetic gene-metabolic oscillator. *Nature* **435**, 118–122 (2005).
181. Weber, W., Lienhart, C., Daoud-El Baba, M. & Fussenegger, M. A biotin-triggered genetic switch in mammalian cells and mice. *Metab. Eng.* **11**, 117–124 (2009).
182. Weber, W., Bacchus, W., Daoud-El Baba, M. & Fussenegger, M. Vitamin H-regulated transgene expression in mammalian cells. *Nucleic Acids Res.* **35**, e116 (2007).
183. Weber, W., Daoud-El Baba, M. & Fussenegger, M. Synthetic ecosystems based on airborne inter- and intrakingdom communication. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10435–10440 (2007).
184. Wang, W.-D., Chen, Z.-T., Kang, B.-G. & Li, R. Construction of an artificial intercellular communication network using the nitric oxide signaling elements in mammalian cells. *Exp. Cell Res.* **314**, 699–706 (2008).
185. Ausländer, D. *et al.* A synthetic multifunctional mammalian pH sensor and CO₂ transgene-control device. *Mol. Cell* **55**, 397–408 (2014).
186. Wright, C. M., Wright, R. C., Eshleman, J. R. & Ostermeier, M. A protein therapeutic modality founded on molecular regulation. *Proc. Natl. Acad. Sci.* **108**, 16206–16211 (2011).
187. Chang, A. L., Wolf, J. J. & Smolke, C. D. Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr. Opin. Biotechnol.* **23**, 679–688 (2012).
188. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
189. Culler, S. J., Hoff, K. G. & Smolke, C. D. Reprogramming Cellular Behavior with RNA Controllers Responsive to Endogenous Proteins. *Science* **330**, 1251–1255 (2010).
190. Slomovic, S. & Collins, J. J. DNA sense-and-respond protein modules for mammalian cells. *Nat. Methods* (2015). doi:10.1038/nmeth.3585
191. Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat. Struct. Mol. Biol.* **20**, 290–299 (2013).
192. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
193. Ehrhardt, K., Guinn, M. T., Quarton, T., Zhang, M. Q. & Bleris, L. Reconfigurable hybrid interface for molecular marker diagnostics and in-situ reporting. *Biosens. Bioelectron.* **74**, 744–750 (2015).
194. Kemmer, C. *et al.* Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nat. Biotechnol.* **28**, 355–360 (2010).

195. Han, J., McLane, B., Kim, E.-H., Yoon, J.-W. & Jun, H.-S. Remission of Diabetes by Insulin Gene Therapy Using a Hepatocyte-specific and Glucose-responsive Synthetic Promoter. *Mol. Ther.* **19**, 470–478 (2011).
196. Rössger, K., Charpin-El-Hamri, G. & Fussenegger, M. A closed-loop synthetic gene circuit for the treatment of diet-induced obesity in mice. *Nat. Commun.* **4**, (2013).
197. Nissim, L. & Bar-Ziv, R. H. A tunable dual-promoter integrator for targeting of cancer cells. *Mol. Syst. Biol.* **6**, (2010).
198. Benenson, Y., Gil, B., Ben-Dor, U., Adar, R. & Shapiro, E. An autonomous molecular computer for logical control of gene expression. *Nature* **429**, 423–429 (2004).
199. Liu, Y. *et al.* Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. *Nat. Commun.* **5**, 5393 (2014).
200. Wroblewska, L. *et al.* Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. *Nat. Biotechnol.* **33**, 839–841 (2015).
201. Lim, W. A. Designing customized cell signalling circuits. *Nat. Rev. Mol. Cell Biol.* **11**, 393–403 (2010).
202. Conklin, B. R. *et al.* Engineering GPCR signaling pathways with RASSLs. *Nat. Methods* **5**, 673–678 (2008).
203. Heng, B. C., Aubel, D. & Fussenegger, M. G Protein–Coupled Receptors Revisited: Therapeutic Applications Inspired by Synthetic Biology. *Annu. Rev. Pharmacol. Toxicol.* **54**, 227–249 (2014).
204. Grupp, S. A. *et al.* Chimeric Antigen Receptor–Modified T Cells for Acute Lymphoid Leukemia. *N. Engl. J. Med.* **368**, 1509–1518 (2013).
205. Kalos, M. *et al.* T Cells with Chimeric Antigen Receptors Have Potent Antitumor Effects and Can Establish Memory in Patients with Advanced Leukemia. *Sci. Transl. Med.* **3**, 95ra73–95ra73 (2011).
206. Brentjens, R. J. *et al.* CD19-Targeted T Cells Rapidly Induce Molecular Remissions in Adults with Chemotherapy-Refractory Acute Lymphoblastic Leukemia. *Sci. Transl. Med.* **5**, 177ra38–177ra38 (2013).
207. Adeniran, A., Sherer, M. & Tyo, K. E. J. Yeast-based biosensors: design and applications. *FEMS Yeast Res.* (2014). doi:10.1111/1567-1364.12203
208. Schofield, D., Sharp, N. J. & Westwater, C. Phage-based platforms for the clinical detection of human bacterial pathogens. *Bacteriophage* **2**, (2012).
209. Citorik, R. J., Mimee, M. & Lu, T. K. Bacteriophage-based synthetic biology for the study of infectious diseases. *Curr. Opin. Microbiol.* **19**, 59–69 (2014).
210. Sin, M. L. Y., Mach, K. E., Wong, P. K. & Liao, J. C. Advances and challenges in biosensor-based diagnosis of infectious diseases. *Expert Rev. Mol. Diagn.* **14**, 225–244 (2014).
211. Miyata, K., Nishiyama, N. & Kataoka, K. Rational design of smart supramolecular assemblies for gene delivery: chemical challenges in the creation of artificial viruses. *Chem. Soc. Rev.* **41**, 2562–2574 (2012).
212. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Syst.* **1**, 187–196 (2015).

213. Bartel, M. A., Weinstein, J. R. & Schaffer, D. V. Directed evolution of novel adeno-associated viruses for therapeutic gene delivery. *Gene Ther.* **19**, 694–700 (2012).
214. Yildiz, I., Shukla, S. & Steinmetz, N. F. Applications of viral nanoparticles in medicine. *Curr. Opin. Biotechnol.* **22**, 901–908 (2011).
215. Märtsch, S., Huber, A., Hallek, M., Büning, H. & Perabo, L. A novel directed evolution method to enhance cell-type specificity of adeno-associated virus vectors. *Comb. Chem. High Throughput Screen.* **13**, 807–812 (2010).
216. Abshire, T. G., Brown, J. E. & Ezzell, J. W. Production and Validation of the Use of Gamma Phage for Identification of Bacillus anthracis. *J. Clin. Microbiol.* **43**, 4780–4788 (2005).
217. McNerney, R., Kambashi, B. S., Kinkese, J., Tembwe, R. & Godfrey-Faussett, P. Development of a Bacteriophage Phage Replication Assay for Diagnosis of Pulmonary Tuberculosis. *J. Clin. Microbiol.* **42**, 2115–2120 (2004).
218. Reiman, R. W., Atchley, D. H. & Voorhees, K. J. Indirect detection of Bacillus anthracis using real-time PCR to detect amplified gamma phage DNA. *J. Microbiol. Methods* **68**, 651–653 (2007).
219. Sergueev, K. V., He, Y., Borschel, R. H., Nikolich, M. P. & Filippov, A. A. Rapid and Sensitive Detection of Yersinia pestis Using Amplification of Plague Diagnostic Bacteriophages Monitored by Real-Time PCR. *PLoS ONE* **5**, e11337 (2010).
220. Fujinami, Y., Hirai, Y., Sakai, I., Yoshino, M. & Yasuda, J. Sensitive detection of Bacillus anthracis using a binding protein originating from gamma-phage. *Microbiol. Immunol.* **51**, 163–169 (2007).
221. Funatsu, T., Taniyama, T., Tajima, T., Tadakuma, H. & Namiki, H. Rapid and sensitive detection method of a bacterium by using a GFP reporter phage. *Microbiol. Immunol.* **46**, 365–369 (2002).
222. Waddell, T. E. & Poppe, C. Construction of mini-Tn10luxABcam/Ptac-ATS and its use for developing a bacteriophage that transduces bioluminescence to Escherichia coli O157:H7. *FEMS Microbiol. Lett.* **182**, 285–289 (2000).
223. Namura, M., Hijikata, T., Miyanaga, K. & Tanji, Y. Detection of Escherichia coli with Fluorescent Labeled Phages That Have a Broad Host Range to E. coli in Sewage Water. *Biotechnol. Prog.* **24**, 481–486 (2008).
224. Dusthacker, A. *et al.* Construction and evaluation of luciferase reporter phages for the detection of active and non-replicating tubercle bacilli. *J. Microbiol. Methods* **73**, 18–25 (2008).
225. Piuri, M., Jacobs, W. R. & Hatfull, G. F. Fluoromycobacteriophages for Rapid, Specific, and Sensitive Antibiotic Susceptibility Testing of Mycobacterium tuberculosis. *PLoS ONE* **4**, e4870 (2009).
226. Jacobs, W. R. *et al.* Rapid assessment of drug susceptibilities of Mycobacterium tuberculosis by means of luciferase reporter phages. *Science* **260**, 819–822 (1993).
227. Carrière, C. *et al.* Conditionally replicating luciferase reporter phages: improved sensitivity for rapid detection and assessment of drug susceptibility of Mycobacterium tuberculosis. *J. Clin. Microbiol.* **35**, 3232–3239 (1997).

228. Loessner, M. J., Rees, C. E., Stewart, G. S. & Scherer, S. Construction of luciferase reporter bacteriophage A511::luxAB for rapid and sensitive detection of viable *Listeria* cells. *Appl. Environ. Microbiol.* **62**, 1133–1140 (1996).
229. Bhowmick, T. *et al.* Controlled Multicenter Evaluation of a Bacteriophage-Based Method for Rapid Detection of *Staphylococcus aureus* in Positive Blood Cultures. *J. Clin. Microbiol.* **51**, 1226–1230 (2013).
230. Smartt, A. E. *et al.* Pathogen detection using engineered bacteriophages. *Anal. Bioanal. Chem.* **402**, 3127–3146 (2012).
231. Pacheco-Gómez, R. *et al.* Detection of Pathogenic Bacteria Using a Homogeneous Immunoassay Based on Shear Alignment of Virus Particles and Linear Dichroism. *Anal. Chem.* **84**, 91–97 (2012).
232. Smartt, A. E. & Ripp, S. Bacteriophage reporter technology for sensing and detecting microbial targets. *Anal. Bioanal. Chem.* **400**, 991–1007 (2011).
233. Tawil, N., Sacher, E., Mandeville, R. & Meunier, M. Bacteriophages: biosensing tools for multi-drug resistant pathogens. *The Analyst* **139**, 1224 (2014).
234. Mani, V. *et al.* Emerging technologies for monitoring drug-resistant tuberculosis at the point-of-care. *Adv. Drug Deliv. Rev.* **78**, 105–117 (2014).
235. *Bacteriophages: biology and applications.* (CRC Press, 2005).
236. Mastrobattista, E., van der Aa, M. A. E. M., Hennink, W. E. & Crommelin, D. J. A. Artificial viruses: a nanotechnological approach to gene delivery. *Nat. Rev. Drug Discov.* **5**, 115–121 (2006).
237. Kim, E. *et al.* Ad-mTERT-delta19, a conditional replication-competent adenovirus driven by the human telomerase promoter, selectively replicates in and elicits cytopathic effect in a cancer cell-specific manner. *Hum. Gene Ther.* **14**, 1415–1428 (2003).
238. Nettelbeck, D. M., Rivera, A. A., Balagué, C., Alemany, R. & Curiel, D. T. Novel oncolytic adenoviruses targeted to melanoma: specific viral replication and cytolysis by expression of E1A mutants from the tyrosinase enhancer/promoter. *Cancer Res.* **62**, 4663–4670 (2002).
239. Manchester, M. & Singh, P. Virus-based nanoparticles (VNPs): platform technologies for diagnostic imaging. *Adv. Drug Deliv. Rev.* **58**, 1505–1522 (2006).
240. Adusumilli, P. S. *et al.* Intraoperative localization of lymph node metastases with a replication-competent herpes simplex virus. *J. Thorac. Cardiovasc. Surg.* **132**, 1179–1188 (2006).
241. Stritzker, J. *et al.* Vaccinia virus-mediated melanin production allows MR and optoacoustic deep tissue imaging and laser-induced thermotherapy of cancer. *Proc. Natl. Acad. Sci.* **110**, 3316–3320 (2013).
242. Haddad, D. & Fong, Y. Molecular imaging of oncolytic viral therapy. *Mol. Ther. — Oncolytics* **1**, 14007 (2015).
243. Hess, M. *et al.* Bacterial glucuronidase as general marker for oncolytic virotherapy or other biological therapies. *J. Transl. Med.* **9**, 172 (2011).
244. Park, J.-S. *et al.* A highly sensitive and selective diagnostic assay based on virus nanoparticles. *Nat. Nanotechnol.* **4**, 259–264 (2009).

245. Lee, J.-H. *et al.* A Three-Dimensional and Sensitive Bioassay Based on Nanostructured Quartz Combined with Viral Nanoparticles. *Adv. Funct. Mater.* **20**, 2004–2009 (2010).
246. Bromley, E. H. C., Channon, K., Moutevelis, E. & Woolfson, D. N. Peptide and Protein Building Blocks for Synthetic Biology: From Programming Biomolecules to Self-Organized Biomolecular Systems. *ACS Chem. Biol.* **3**, 38–50 (2008).
247. Hockenberry, A. J. & Jewett, M. C. Synthetic in vitro circuits. *Curr. Opin. Chem. Biol.* **16**, 253–259 (2012).
248. Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.* **14**, 261–269 (2012).
249. Breaker, R. R. & Joyce, G. F. The Expanding View of RNA and DNA Function. *Chem. Biol.* **21**, 1059–1065 (2014).
250. Linko, V. & Dietz, H. The enabled state of DNA nanotechnology. *Curr. Opin. Biotechnol.* **24**, 555–561 (2013).
251. Wang, F., Willner, B. & Willner, I. DNA nanotechnology with one-dimensional self-assembled nanostructures. *Curr. Opin. Biotechnol.* **24**, 562–574 (2013).
252. Padirac, A., Fujii, T. & Rondelez, Y. Nucleic acids for the rational design of reaction circuits. *Curr. Opin. Biotechnol.* **24**, 575–580 (2013).
253. Wang, F., Lu, C.-H. & Willner, I. From Cascaded Catalytic Nucleic Acids to Enzyme–DNA Nanostructures: Controlling Reactivity, Sensing, Logic Operations, and Assembly of Complex Structures. *Chem. Rev.* **114**, 2881–2941 (2014).
254. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-Free Nucleic Acid Logic Circuits. *Science* **314**, 1585–1588 (2006).
255. Famulok, M. & Mayer, G. Aptamers and SELEX in Chemistry & Biology. *Chem. Biol.* **21**, 1055–1058 (2014).
256. Yang, L. & Ellington, A. in *Fluorescence Sensors and Biosensors* (ed. Thompson, R.) 5–43 (CRC Press, 2005). at <<http://www.crcnetbase.com/doi/abs/10.1201/9781420028287.ch2>>
257. Famulok, M., Hartig, J. S. & Mayer, G. Functional Aptamers and Aptazymes in Biotechnology, Diagnostics, and Therapy. *Chem. Rev.* **107**, 3715–3743 (2007).
258. Ozer, A., Pagano, J. M. & Lis, J. T. New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Mol. Ther. Acids* **3**, e183 (2014).
259. Hong, P., Li, W. & Li, J. Applications of Aptasensors in Clinical Diagnostics. *Sensors* **12**, 1181–1193 (2012).
260. Toh, S. Y., Citartan, M., Gopinath, S. C. B. & Tang, T.-H. Aptamers as a replacement for antibodies in enzyme-linked immunosorbent assay. *Biosens. Bioelectron.* **64**, 392–403 (2015).
261. Zhou, M. *et al.* Aptamer-Controlled Biofuel Cells in Logic Systems and Used as Self-Powered and Intelligent Logic Aptasensors. *J. Am. Chem. Soc.* **132**, 2172–2174 (2010).

262. Zhou, W., Jimmy Huang, P.-J., Ding, J. & Liu, J. Aptamer-based biosensors for biomedical diagnostics. *The Analyst* **139**, 2627 (2014).
263. Serganov, A. & Nudler, E. A Decade of Riboswitches. *Cell* **152**, 17–24 (2013).
264. Tang, J. & Breaker, R. R. Rational design of allosteric ribozymes. *Chem. Biol.* **4**, 453–459 (1997).
265. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281–286 (2004).
266. Frommer, J., Appel, B. & Müller, S. Ribozymes that can be regulated by external stimuli. *Curr. Opin. Biotechnol.* **31**, 35–41 (2015).
267. Berens, C. & Suess, B. Riboswitch engineering — making the all-important second and third steps. *Curr. Opin. Biotechnol.* **31**, 10–15 (2015).
268. Ogawa, A. Rational design of artificial riboswitches based on ligand-dependent modulation of internal ribosome entry in wheat germ extract and their applications as label-free biosensors. *RNA* **17**, 478–488 (2011).
269. Olea, C. & Joyce, G. F. in *Methods in Enzymology* **550**, 23–39 (Elsevier, 2015).
270. Kahan-Hanum, M., Douek, Y., Adar, R. & Shapiro, E. A library of programmable DNAzymes that operate in a cellular environment. *Sci. Rep.* **3**, (2013).
271. Kong, R.-M. *et al.* Unimolecular Catalytic DNA Biosensor for Amplified Detection of L -Histidine via an Enzymatic Recycling Cleavage Strategy. *Anal. Chem.* **83**, 7603–7607 (2011).
272. Tang, L. *et al.* Colorimetric and Ultrasensitive Bioassay Based on a Dual-Amplification System Using Aptamer and DNAzyme. *Anal. Chem.* **84**, 4711–4717 (2012).
273. Pavlov, V. *et al.* Amplified chemiluminescence surface detection of DNA and telomerase activity using catalytic nucleic acid labels. *Anal. Chem.* **76**, 2152–2156 (2004).
274. Freeman, R. *et al.* DNAzyme-Like Activity of Hemin-Telomeric G-Quadruplexes for the Optical Analysis of Telomerase and its Inhibitors. *ChemBioChem* **11**, 2362–2367 (2010).
275. Qian, L. & Winfree, E. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science* **332**, 1196–1201 (2011).
276. Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **3**, 103–113 (2011).
277. Chen, S. X., Zhang, D. Y. & Seelig, G. Conditionally fluorescent molecular probes for detecting single base changes in double-stranded DNA. *Nat. Chem.* **5**, 782–789 (2013).
278. Jung, C. & Ellington, A. D. Diagnostic Applications of Nucleic Acid Circuits. *Acc. Chem. Res.* **47**, 1825–1835 (2014).
279. Eckhoff, G., Codrea, V., Ellington, A. D. & Chen, X. Beyond allostery: Catalytic regulation of a deoxyribozyme through an entropy-driven DNA amplifier. *J. Syst. Chem.* **1**, 13 (2010).
280. Li, B., Ellington, A. D. & Chen, X. Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Res.* **39**, e110–e110 (2011).

281. Li, B., Chen, X. & Ellington, A. D. Adapting Enzyme-Free DNA Circuits to the Detection of Loop-Mediated Isothermal Amplification Reactions. *Anal. Chem.* **84**, 8371–8377 (2012).
282. Zhang, B. *et al.* DNA-Based Hybridization Chain Reaction for Amplified Bioelectronic Signal and Ultrasensitive Detection of Proteins. *Anal. Chem.* **84**, 5392–5399 (2012).
283. Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.* **28**, 1208–1212 (2010).
284. Choi, H. M. T., Beck, V. A. & Pierce, N. A. Next-Generation *in Situ* Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability. *ACS Nano* **8**, 4284–4294 (2014).
285. Gil, B., Kahan-Hanum, M., Skirtenko, N., Adar, R. & Shapiro, E. Detection of Multiple Disease Indicators by an Autonomous Biomolecular Computer. *Nano Lett.* **11**, 2989–2996 (2011).
286. Zhou, G. *et al.* Multivalent Capture and Detection of Cancer Cells with DNA Nanostructured Biosensors and Multibranching Hybridization Chain Reaction Amplification. *Anal. Chem.* **86**, 7843–7848 (2014).
287. Rudchenko, M. *et al.* Autonomous molecular cascades for evaluation of cell surfaces. *Nat. Nanotechnol.* **8**, 580–586 (2013).
288. Allen, P. B., Arshad, S. A., Li, B., Chen, X. & Ellington, A. D. DNA circuits as amplifiers for the detection of nucleic acids on a paperfluidic platform. *Lab. Chip* **12**, 2951 (2012).
289. Chen, J., Zhou, X. & Zeng, L. Enzyme-free strip biosensor for amplified detection of Pb²⁺ based on a catalytic DNA circuit. *Chem Commun* **49**, 984–986 (2013).
290. Elbeik, T. *et al.* Multicenter Evaluation of the Performance Characteristics of the Bayer VERSANT HCV RNA 3.0 Assay (bDNA). *J. Clin. Microbiol.* **42**, 563–569 (2004).
291. Elbeik, T. *et al.* Simultaneous runs of the Bayer VERSANT HIV-1 version 3.0 and HCV bDNA version 3.0 quantitative assays on the system 340 platform provide reliable quantitation and improved work flow. *J. Clin. Microbiol.* **42**, 3120–3127 (2004).
292. Campolongo, M. J., Tan, S. J., Xu, J. & Luo, D. DNA nanomedicine: Engineering DNA as a polymer for therapeutic and diagnostic applications. *Adv. Drug Deliv. Rev.* **62**, 606–616 (2010).
293. Zhang, L., Lei, J., Liu, L., Li, C. & Ju, H. Self-Assembled DNA Hydrogel as Switchable Material for Aptamer-Based Fluorescent Detection of Protein. *Anal. Chem.* **85**, 11077–11082 (2013).
294. *Algorithmic bioprocesses.* (Springer, 2009).
295. Bath, J. & Turberfield, A. J. DNA nanomachines. *Nat. Nanotechnol.* **2**, 275–284 (2007).
296. Ben-Ishay, E., Abu-Horowitz, A. & Bachelet, I. Designing a Bio-responsive Robot from DNA Origami. *J. Vis. Exp.* (2013). doi:10.3791/50268
297. Wang, D. *et al.* Molecular Logic Gates on DNA Origami Nanostructures for MicroRNA Diagnostics. *Anal. Chem.* **86**, 1932–1936 (2014).
298. Amir, Y. *et al.* Universal computing by DNA origami robots in a living animal. *Nat. Nanotechnol.* **9**, 353–357 (2014).

299. Taylor, A. I., Arangundy-Franklin, S. & Holliger, P. Towards applications of synthetic genetic polymers in diagnosis and therapy. *Curr. Opin. Chem. Biol.* **22**, 79–84 (2014).
300. Kimoto, M., Yamashige, R., Matsunaga, K., Yokoyama, S. & Hirao, I. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat. Biotechnol.* **31**, 453–457 (2013).
301. Davies, D. R. *et al.* Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl. Acad. Sci.* **109**, 19971–19976 (2012).
302. Imaizumi, Y. *et al.* Efficacy of Base-Modification on Target Binding of Small Molecule DNA Aptamers. *J. Am. Chem. Soc.* **135**, 9412–9419 (2013).
303. Shi, H. *et al.* Locked nucleic acid/DNA chimeric aptamer probe for tumor diagnosis with improved serum stability and extended imaging window in vivo. *Anal. Chim. Acta* **812**, 138–144 (2014).
304. Olson, E. J. & Tabor, J. J. Post-translational tools expand the scope of synthetic biology. *Curr. Opin. Chem. Biol.* **16**, 300–306 (2012).
305. Foo, J. L., Ching, C. B., Chang, M. W. & Leong, S. S. J. The imminent role of protein engineering in synthetic biology. *Biotechnol. Adv.* **30**, 541–549 (2012).
306. Grunberg, R. & Serrano, L. Strategies for protein synthetic biology. *Nucleic Acids Res.* **38**, 2663–2675 (2010).
307. Gunnoo, S. B. *et al.* Creation of a gated antibody as a conditionally functional synthetic protein. *Nat. Commun.* **5**, (2014).
308. Strauch, E.-M., Fleishman, S. J. & Baker, D. Computational design of a pH-sensitive IgG binding protein. *Proc. Natl. Acad. Sci.* **111**, 675–680 (2014).
309. Byrne, H., Conroy, P. J., Whisstock, J. C. & O’Kennedy, R. J. A tale of two specificities: bispecific antibodies for therapeutic and diagnostic applications. *Trends Biotechnol.* **31**, 621–632 (2013).
310. Currin, A., Swainston, N., Day, P. J. & Kell, D. B. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem Soc Rev* **44**, 1172–1239 (2015).
311. Saito, K. *et al.* Luminescent proteins for high-speed single-cell and whole-body imaging. *Nat. Commun.* **3**, 1262 (2012).
312. Ballew, J. T. *et al.* Antibody biomarker discovery through in vitro directed evolution of consensus recognition epitopes. *Proc. Natl. Acad. Sci.* **110**, 19330–19335 (2013).
313. Gebauer, M. & Skerra, A. Engineered protein scaffolds as next-generation antibody therapeutics. *Curr. Opin. Chem. Biol.* **13**, 245–255 (2009).
314. Miao, Z., Levi, J. & Cheng, Z. Protein scaffold-based molecular probes for cancer molecular imaging. *Amino Acids* **41**, 1037–1047 (2011).
315. You, C. & Zhang, Y.-H. P. Self-Assembly of Synthetic Metabolons through Synthetic Protein Scaffolds: One-Step Purification, Co-immobilization, and Substrate Channeling. *ACS Synth. Biol.* **2**, 102–110 (2013).
316. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759 (2009).

317. Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855–862 (2009).
318. Kiss, G., Çelebi-Ölçüm, N., Moretti, R., Baker, D. & Houk, K. N. Computational Enzyme Design. *Angew. Chem. Int. Ed.* **52**, 5700–5725 (2013).
319. Kaplan, J. & DeGrado, W. F. De novo design of catalytic proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11566–11570 (2004).
320. Röthlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
321. Khoury, G. A., Smadbeck, J., Kieslich, C. A. & Floudas, C. A. Protein folding and de novo protein design for biotechnological applications. *Trends Biotechnol.* **32**, 99–109 (2014).
322. Feldmeier, K. & Höcker, B. Computational protein design of ligand binding and catalysis. *Curr. Opin. Chem. Biol.* **17**, 929–933 (2013).
323. Mandell, D. J. & Kortemme, T. Computer-aided design of functional protein interactions. *Nat. Chem. Biol.* **5**, 797–807 (2009).
324. Schreier, B., Stumpp, C., Wiesner, S. & Hocker, B. Computational design of ligand binding is not a solved problem. *Proc. Natl. Acad. Sci.* **106**, 18491–18496 (2009).
325. Tinberg, C. E. *et al.* Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501**, 212–216 (2013).
326. Kortemme, T. & Baker, D. Computational design of protein-protein interactions. *Curr. Opin. Chem. Biol.* **8**, 91–97 (2004).
327. Stein, V. & Alexandrov, K. Synthetic protein switches: design principles and applications. *Trends Biotechnol.* **33**, 101–110 (2015).
328. Ostermeier, M. Designing switchable enzymes. *Curr. Opin. Struct. Biol.* **19**, 442–448 (2009).
329. Dueber, J. E., Yeh, B. J., Bhattacharyya, R. P. & Lim, W. A. Rewiring cell signaling: the logic and plasticity of eukaryotic protein circuitry. *Curr. Opin. Struct. Biol.* **14**, 690–699 (2004).
330. Dueber, J. E., Yeh, B. J., Chak, K. & Lim, W. A. Reprogramming control of an allosteric signaling switch through modular recombination. *Science* **301**, 1904–1908 (2003).
331. Dueber, J. E., Mirsky, E. A. & Lim, W. A. Engineering synthetic signaling proteins with ultrasensitive input/output control. *Nat. Biotechnol.* **25**, 660–662 (2007).
332. Yeh, B. J., Rutigliano, R. J., Deb, A., Bar-Sagi, D. & Lim, W. A. Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* **447**, 596–600 (2007).
333. Guntas, G., Mansell, T. J., Kim, J. R. & Ostermeier, M. Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 11224–11229 (2005).
334. Guntas, G. & Ostermeier, M. Creation of an allosteric enzyme by domain insertion. *J. Mol. Biol.* **336**, 263–273 (2004).

335. Dagliyan, O. *et al.* Rational design of a ligand-controlled protein conformational switch. *Proc. Natl. Acad. Sci.* **110**, 6800–6804 (2013).
336. Karginov, A. V., Ding, F., Kota, P., Dokholyan, N. V. & Hahn, K. M. Engineered allosteric activation of kinases in living cells. *Nat. Biotechnol.* **28**, 743–747 (2010).
337. Sallee, N. A., Yeh, B. J. & Lim, W. A. Engineering modular protein interaction switches by sequence overlap. *J. Am. Chem. Soc.* **129**, 4606–4611 (2007).
338. Ibraheem, A. & Campbell, R. E. Designs and applications of fluorescent protein-based biosensors. *Curr. Opin. Chem. Biol.* **14**, 30–36 (2010).
339. Binkowski, B., Fan, F. & Wood, K. Engineered luciferases for molecular sensing in living cells. *Curr. Opin. Biotechnol.* **20**, 14–18 (2009).
340. Banala, S., Aper, S. J. A., Schalk, W. & Merckx, M. Switchable Reporter Enzymes Based on Mutually Exclusive Domain Interactions Allow Antibody Detection Directly in Solution. *ACS Chem. Biol.* **8**, 2127–2132 (2013).
341. Nirantar, S. R., Yeo, K. S., Chee, S., Lane, D. P. & Ghadessy, F. J. A generic scaffold for conversion of peptide ligands into homogenous biosensors. *Biosens. Bioelectron.* **47**, 421–428 (2013).
342. Ashkenasy, G. & Ghadiri, M. R. Boolean logic functions of a synthetic peptide network. *J. Am. Chem. Soc.* **126**, 11140–11141 (2004).
343. Niazov, T., Baron, R., Katz, E., Lioubashevski, O. & Willner, I. Concatenated logic gates using four coupled biocatalysts operating in series. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17160–17163 (2006).
344. Sugita, M. Functional analysis of chemical systems in vivo using a logical circuit equivalent. II. The idea of a molecular automation. *J. Theor. Biol.* **4**, 179–192 (1963).
345. Arkin, A. & Ross, J. Computational functions in biochemical reaction networks. *Biophys. J.* **67**, 560–578 (1994).
346. Wang, J. & Katz, E. Digital biosensors with built-in logic for biomedical applications—biosensors based on a biocomputing concept. *Anal. Bioanal. Chem.* **398**, 1591–1603 (2010).
347. Katz, E., Minko, S., Halámek, J., MacVittie, K. & Yancey, K. Electrode interfaces switchable by physical and chemical signals for biosensing, biofuel, and biocomputing applications. *Anal. Bioanal. Chem.* **405**, 3659–3672 (2013).
348. Poghossian, A. *et al.* Integration of biomolecular logic gates with field-effect transducers. *Electrochimica Acta* **56**, 9661–9665 (2011).
349. Krämer, M. *et al.* Coupling of Biocomputing Systems with Electronic Chips: Electronic Interface for Transduction of Biochemical Information. *J. Phys. Chem. C* **113**, 2573–2579 (2009).
350. Katz, E. & Minko, S. Enzyme-based logic systems interfaced with signal-responsive materials and electrodes. *Chem Commun* **51**, 3493–3500 (2015).
351. Privman, M., Tam, T. K., Pita, M. & Katz, E. Switchable Electrode Controlled by Enzyme Logic Network System: Approaching Physiologically Regulated Bioelectronics. *J. Am. Chem. Soc.* **131**, 1314–1321 (2009).

352. Katz, E., Wang, J., Privman, M. & Halámek, J. Multianalyte Digital Enzyme Biosensors with Built-in Boolean Logic. *Anal. Chem.* **84**, 5463–5469 (2012).
353. Zhou, J., Halámek, J., Bocharova, V., Wang, J. & Katz, E. Bio-logic analysis of injury biomarker patterns in human serum samples. *Talanta* **83**, 955–959 (2011).
354. Halámek, J. *et al.* Multi-enzyme logic network architectures for assessing injuries: digital processing of biomarkers. *Mol. Biosyst.* **6**, 2554 (2010).
355. Pita, M. *et al.* Enzyme logic gates for assessing physiological conditions during an injury: Towards digital sensors and actuators. *Sens. Actuators B Chem.* **139**, 631–636 (2009).
356. Halámek, J. *et al.* Multiplexing of injury codes for the parallel operation of enzyme logic gates. *The Analyst* **135**, 2249 (2010).
357. Zhou, N. *et al.* Enzyme-based NAND gate for rapid electrochemical screening of traumatic brain injury in serum. *Anal. Chim. Acta* **703**, 94–100 (2011).
358. Mailloux, S., Halámek, J. & Katz, E. A model system for targeted drug release triggered by biomolecular signals logically processed through enzyme logic networks. *The Analyst* **139**, 982–986 (2014).
359. Tokarev, I. *et al.* Stimuli-responsive hydrogel membranes coupled with biocatalytic processes. *ACS Appl. Mater. Interfaces* **1**, 532–536 (2009).
360. Mailloux, S., Guz, N., Gamella Carballo, M., Pingarrón, J. M. & Katz, E. Model system for targeted drug release triggered by immune-specific signals. *Anal. Bioanal. Chem.* **406**, 4825–4829 (2014).
361. Guz, N., Halámek, J., Rusling, J. F. & Katz, E. A biocatalytic cascade with several output signals—towards biosensors with different levels of confidence. *Anal. Bioanal. Chem.* **406**, 3365–3370 (2014).
362. Privman, V., Domanskyi, S., Mailloux, S., Holade, Y. & Katz, E. Kinetic Model for a Threshold Filter in an Enzymatic System for Bioanalytical and Biocomputing Applications. *J. Phys. Chem. B* **118**, 12435–12443 (2014).
363. Bakshi, S., Zavalov, O., Halámek, J., Privman, V. & Katz, E. Modularity of Biochemical Filtering for Inducing Sigmoid Response in Both Inputs in an Enzymatic AND Gate. *J. Phys. Chem. B* **117**, 9857–9865 (2013).
364. Melnikov, D., Strack, G., Pita, M., Privman, V. & Katz, E. Analog Noise Reduction in Enzymatic Logic Gates. *J. Phys. Chem. B* **113**, 10472–10479 (2009).
365. Privman, V., Strack, G., Solenov, D., Pita, M. & Katz, E. Optimization of Enzymatic Biochemical Logic for Noise Reduction and Scalability: How Many Biocomputing Gates Can Be Interconnected in a Circuit? *J. Phys. Chem. B* **112**, 11777–11784 (2008).
366. Hucka, M. *et al.* The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinforma. Oxf. Engl.* **19**, 524–531 (2003).
367. Rialle, S. *et al.* BioNetCAD: design, simulation and experimental validation of synthetic biochemical networks. *Bioinformatics* **26**, 2298–2304 (2010).
368. Amar, P. *et al.* A stochastic automaton shows how enzyme assemblies may contribute to metabolic efficiency. *BMC Syst. Biol.* **2**, 27 (2008).

369. Rizk, A., Batt, G., Fages, F. & Soliman, S. A general computational method for robustness analysis with applications to synthetic gene networks. *Bioinformatics* **25**, i169–i178 (2009).
370. Liu, C. C. & Schultz, P. G. Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).
371. O'Donoghue, P., Ling, J., Wang, Y.-S. & Söll, D. Upgrading protein synthesis for synthetic biology. *Nat. Chem. Biol.* **9**, 594–598 (2013).
372. Acevedo-Rocha, C. G. *et al.* Non-canonical amino acids as a useful synthetic biological tool for lipase-catalysed reactions in hostile environments. *Catal. Sci. Technol.* **3**, 1198 (2013).
373. Lemke, E. A., Summerer, D., Geierstanger, B. H., Brittain, S. M. & Schultz, P. G. Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nat. Chem. Biol.* **3**, 769–772 (2007).
374. Wang, K. *et al.* Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **6**, 393–403 (2014).
375. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108 (2014).
376. Castillo-León, J., Belotti, Y. & Svendsen, W. in *Nanomedicine in Diagnostics* (ed. Rozlosnik, N.) 50–67 (Science Publishers, 2012). at <<http://www.crcnetbase.com/doi/abs/10.1201/b11929-4>>
377. Men, D. *et al.* Seeding-Induced Self-assembling Protein Nanowires Dramatically Increase the Sensitivity of Immunoassays. *Nano Lett.* **9**, 2246–2250 (2009).
378. Leng, Y. *et al.* Integration of a Fluorescent Molecular Biosensor into Self-Assembled Protein Nanowires: A Large Sensitivity Enhancement. *Angew. Chem. Int. Ed.* **49**, 7243–7246 (2010).
379. Domigan, L. J. in *Protein Nanotechnology* (ed. Gerrard, J. A.) **996**, 131–152 (Humana Press, 2013).
380. Lee, J.-H. *et al.* Proteinticle Engineering for Accurate 3D Diagnosis. *ACS Nano* **7**, 10879–10886 (2013).
381. Chou, D. H.-C. *et al.* Glucose-responsive insulin activity by covalent modification with aliphatic phenylboronic acid conjugates. *Proc. Natl. Acad. Sci.* **112**, 2401–2406 (2015).
382. Matsubayashi, H. & Ueda, T. Purified cell-free systems as standard parts for synthetic biology. *Curr. Opin. Chem. Biol.* **22**, 158–162 (2014).
383. Jewett, M. C. & Forster, A. C. Update on designing and building minimal cells. *Curr. Opin. Biotechnol.* **21**, 697–703 (2010).
384. Stanó, P. & Luisi, P. L. Semi-synthetic minimal cells: origin and recent developments. *Curr. Opin. Biotechnol.* **24**, 633–638 (2013).
385. Doktycz, M. J. & Simpson, M. L. Nano-enabled synthetic biology. *Mol. Syst. Biol.* **3**, 125 (2007).
386. Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17669–17674 (2004).

387. Wu, F. & Tan, C. The engineering of artificial cellular nanosystems using synthetic biology approaches: Artificial cellular nanosystems using synthetic biology approaches. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **6**, 369–383 (2014).
388. Stano, P. in *Progress in Molecular and Environmental Bioengineering - From Analysis and Modeling to Technology Applications* (ed. Carpi, A.) (InTech, 2011). at <<http://www.intechopen.com/books/progress-in-molecular-and-environmental-bioengineering-from-analysis-and-modeling-to-technology-applications/advances-in-minimal-cell-models-a-new-approach-to-synthetic-biology-and-origin-of-life>>
389. Zhang, Y., Ruder, W. C. & LeDuc, P. R. Artificial cells: building bioinspired systems using small-scale biology. *Trends Biotechnol.* **26**, 14–20 (2008).
390. Courbet, A., Molina, F. & Amar, P. Computing with Synthetic Protocells. *Acta Biotheor.* (2015). doi:10.1007/s10441-015-9258-8
391. Smaldon, J. *et al.* A computational study of liposome logic: towards cellular computing from the bottom up. *Syst. Synth. Biol.* **4**, 157–179 (2010).
392. Broz, P. *et al.* Toward intelligent nanosize bioreactors: a pH-switchable, channel-equipped, functional polymer nanocontainer. *Nano Lett.* **6**, 2349–2353 (2006).
393. Pasparakis, G., Krasnogor, N., Cronin, L., Davis, B. G. & Alexander, C. Controlled polymer synthesis—from biomimicry towards synthetic biology. *Chem Soc Rev* **39**, 286–300 (2010).
394. Christensen, S. M. & Stamou, D. G. Sensing-Applications of Surface-Based Single Vesicle Arrays. *Sensors* **10**, 11352–11368 (2010).
395. Tanner, P. *et al.* Polymeric vesicles: from drug carriers to nanoreactors and artificial organelles. *Acc. Chem. Res.* **44**, 1039–1049 (2011).
396. Fallah-Araghi, A. *et al.* Enhanced chemical synthesis at soft interfaces: a universal reaction-adsorption mechanism in microcompartments. *Phys. Rev. Lett.* **112**, 028301 (2014).
397. Moghimi, S. M., Hunter, A. C. & Murray, J. C. Nanomedicine: current status and future prospects. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **19**, 311–330 (2005).
398. Elani, Y., Law, R. V. & Ces, O. Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat. Commun.* **5**, 5305 (2014).
399. Kim, J. P., Park, C. H. & Sim, S. J. Aptamer biosensors for label-free colorimetric detection of human IgE based on polydiacetylene (PDA) supramolecules. *J. Nanosci. Nanotechnol.* **11**, 4269–4274 (2011).
400. Zhu, C., Liu, L., Yang, Q., Lv, F. & Wang, S. Water-soluble conjugated polymers for imaging, diagnosis, and therapy. *Chem. Rev.* **112**, 4687–4735 (2012).
401. Ho, R. J., Rouse, B. T. & Huang, L. Interactions of target-sensitive immunoliposomes with herpes simplex virus. The foundation of a sensitive immunoliposome assay for the virus. *J. Biol. Chem.* **262**, 13979–13984 (1987).
402. Kolusheva, S., Molt, O., Herm, M., Schrader, T. & Jelinek, R. Selective detection of catecholamines by synthetic receptors embedded in chromatic polydiacetylene vesicles. *J. Am. Chem. Soc.* **127**, 10000–10001 (2005).
403. Liu, Q. & Boyd, B. J. Liposomes in biosensors. *The Analyst* **138**, 391–409 (2013).

404. Haque, F., Geng, J., Montemagno, C. & Guo, P. Incorporation of a viral DNA-packaging motor channel in lipid bilayers for real-time, single-molecule sensing of chemicals and double-stranded DNA. *Nat. Protoc.* **8**, 373–392 (2013).
405. Baumann, P., Tanner, P., Onaca, O. & Palivan, C. G. Bio-Decorated Polymer Membranes: A New Approach in Diagnostics and Therapeutics. *Polymers* **3**, 173–192 (2011).
406. Martino, C. *et al.* Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures. *Angew. Chem. Int. Ed Engl.* **51**, 6416–6420 (2012).
407. Discher, D. E. Polymer Vesicles. *Science* **297**, 967–973 (2002).
408. Kamat, N. P., Katz, J. S. & Hammer, D. A. Engineering Polymersome Protocells. *J. Phys. Chem. Lett.* **2**, 1612–1623 (2011).
409. Peters, R. J. R. W. *et al.* Cascade Reactions in Multicompartmentalized Polymersomes. *Angew. Chem. Int. Ed.* **53**, 146–150 (2014).
410. Caldorera-Moore, M. E., Liechty, W. B. & Peppas, N. A. Responsive Theranostic Systems: Integration of Diagnostic Imaging Agents and Responsive Controlled Release Drug Delivery Carriers. *Acc. Chem. Res.* **44**, 1061–1070 (2011).
411. Phillips, M. A., Gran, M. L. & Peppas, N. A. Targeted nanodelivery of drugs and diagnostics. *Nano Today* **5**, 143–159 (2010).
412. Ren, T. *et al.* Multifunctional polymer vesicles for ultrasensitive magnetic resonance imaging and drug delivery. *J. Mater. Chem.* **22**, 12329 (2012).
413. Ben-Haim, N., Broz, P., Marsch, S., Meier, W. & Hunziker, P. Cell-Specific Integration of Artificial Organelles Based on Functionalized Polymer Vesicles. *Nano Lett.* **8**, 1368–1373 (2008).
414. Choi, J. *et al.* Biocompatible heterostructured nanoparticles for multimodal biological detection. *J. Am. Chem. Soc.* **128**, 15982–15983 (2006).
415. Cutler, J. I., Auyeung, E. & Mirkin, C. A. Spherical Nucleic Acids. *J. Am. Chem. Soc.* **134**, 1376–1391 (2012).
416. Wilner, O. I. *et al.* Enzyme cascades activated on topologically programmed DNA scaffolds. *Nat. Nanotechnol.* **4**, 249–254 (2009).
417. Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *J. Am. Chem. Soc.* **134**, 5516–5519 (2012).
418. Piperberg, G., Wilner, O. I., Yehezkeli, O., Tel-Vered, R. & Willner, I. Control of bioelectrocatalytic transformations on DNA scaffolds. *J. Am. Chem. Soc.* **131**, 8724–8725 (2009).
419. Wang, Z.-G., Wilner, O. I. & Willner, I. Self-Assembly of Aptamer–Circular DNA Nanostructures for Controlled Biocatalysis. *Nano Lett.* **9**, 4098–4102 (2009).
420. Golub, E., Freeman, R. & Willner, I. A hemin/G-quadruplex acts as an NADH oxidase and NADH peroxidase mimicking DNzyme. *Angew. Chem. Int. Ed Engl.* **50**, 11710–11714 (2011).

421. von Maltzahn, G. *et al.* Nanoparticle self-assembly gated by logical proteolytic triggers. *J. Am. Chem. Soc.* **129**, 6064–6065 (2007).
422. Konry, T. & Walt, D. R. Intelligent Medical Diagnostics via Molecular Logic. *J. Am. Chem. Soc.* **131**, 13232–13233 (2009).
423. Janssen, B. M. G., van Rosmalen, M., van Beek, L. & Merkx, M. Antibody Activation using DNA-Based Logic Gates. *Angew. Chem. Int. Ed.* **54**, 2530–2533 (2015).
424. Lai, Y.-H. *et al.* Biomolecular logic gate for analysis of the New Delhi metallo- β -lactamase (NDM)-coding gene with concurrent determination of its drug resistance-encoding fragments. *Chem Commun* **50**, 12018–12021 (2014).
425. Motornov, M. *et al.* ‘Chemical Transformers’ from Nanoparticle Ensembles Operated with Logic. *Nano Lett.* **8**, 2993–2997 (2008).
426. Motornov, M. *et al.* An Integrated Multifunctional Nanosystem from Command Nanoparticles and Enzymes. *Small* **5**, 817–820 (2009).
427. Motornov, M., Roiter, Y., Tokarev, I. & Minko, S. Stimuli-responsive nanoparticles, nanogels and capsules for integrated multifunctional intelligent systems. *Prog. Polym. Sci.* **35**, 174–211 (2010).
428. Lehner, R., Wang, X., Marsch, S. & Hunziker, P. Intelligent nanomaterials for medicine: Carrier platforms and targeting strategies in the context of clinical application. *Nanomedicine Nanotechnol. Biol. Med.* **9**, 742–757 (2013).
429. Wu, W. *et al.* Multi-functional core-shell hybrid nanogels for pH-dependent magnetic manipulation, fluorescent pH-sensing, and drug delivery. *Biomaterials* **32**, 9876–9887 (2011).
430. Naskar, J., Roy, S., Joardar, A., Das, S. & Banerjee, A. Self-assembling dipeptide-based nontoxic vesicles as carriers for drugs and other biologically important molecules. *Org. Biomol. Chem.* **9**, 6610 (2011).
431. Edmundson, M. C., Capeness, M. & Horsfall, L. Exploring the potential of metallic nanoparticles within synthetic biology. *New Biotechnol.* **31**, 572–578 (2014).
432. Zhou, M. & Dong, S. Bioelectrochemical Interface Engineering: Toward the Fabrication of Electrochemical Biosensors, Biofuel Cells, and Self-Powered Logic Biosensors. *Acc. Chem. Res.* **44**, 1232–1243 (2011).
433. Katz, E. & Pita, M. Biofuel Cells Controlled by Logically Processed Biochemical Signals: Towards Physiologically Regulated Bioelectronic Devices. *Chem. - Eur. J.* **15**, 12554–12564 (2009).
434. Tam, T. K., Pita, M., Ornatska, M. & Katz, E. Biofuel cell controlled by enzyme logic network — Approaching physiologically regulated devices. *Bioelectrochemistry* **76**, 4–9 (2009).
435. Zhou, M. & Wang, J. Biofuel Cells for Self-Powered Electrochemical Biosensing and Logic Biosensing: A Review. *Electroanalysis* **24**, 197–209 (2012).
436. Zebda, A. *et al.* Single Glucose Biofuel Cells Implanted in Rats Power Electronic Devices. *Sci. Rep.* **3**, (2013).
437. Falk, M., Andoralov, V., Silow, M., Toscano, M. D. & Shleev, S. Miniature Biofuel Cell as a Potential Power Source for Glucose-Sensing Contact Lenses. *Anal. Chem.* **85**, 6342–6348 (2013).

438. Zhou, M. *et al.* A Self-Powered ‘Sense-Act-Treat’ System that is Based on a Biofuel Cell and Controlled by Boolean Logic. *Angew. Chem. Int. Ed.* **51**, 2686–2689 (2012).
439. Liu, A. P. & Fletcher, D. A. Biology under construction: in vitro reconstitution of cellular function. *Nat. Rev. Mol. Cell Biol.* **10**, 644–650 (2009).
440. LeDuc, P. R. *et al.* Towards an in vivo biologically inspired nanofactory. *Nat. Nanotechnol.* **2**, 3–7 (2007).
441. Lehner, R., Wang, X., Wolf, M. & Hunziker, P. Designing switchable nanosystems for medical application. *J. Controlled Release* **161**, 307–316 (2012).
442. in *IUPAC Compendium of Chemical Terminology* (eds. Nič, M., Jiráť, J., Košata, B., Jenkins, A. & McNaught, A.) (IUPAC, 2009). at <<http://goldbook.iupac.org/B00663.html>>
443. Ferrell, J. E. & Ha, S. H. Ultrasensitivity part III: cascades, bistable switches, and oscillators. *Trends Biochem. Sci.* **39**, 612–618 (2014).
444. Monod, J. & Jacob, F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 389–401 (1961).
445. Regev, A. & Shapiro, E. Cellular abstractions: Cells as computation. *Nature* **419**, 343–343 (2002).
446. Haynes, K. A. *et al.* Engineering bacteria to solve the Burnt Pancake Problem. *J. Biol. Eng.* **2**, 8 (2008).
447. Benenson, Y. *et al.* Programmable and autonomous computing machine made of biomolecules. *Nature* **414**, 430–434 (2001).
448. Ausländer, S., Ausländer, D., Müller, M., Wieland, M. & Fussenegger, M. Programmable single-cell mammalian biocomputers. *Nature* (2012). doi:10.1038/nature11149
449. *Biomolecular information processing: from logic systems to smart sensors and actuators.* (Wiley-VCH, 2012).
450. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333**, 1307–1311 (2011).
451. Sarpeshkar, R. *Ultra low power bioelectronics: fundamentals, biomedical applications, and bio-inspired systems.* (Cambridge University Press, 2010).
452. Miyamoto, T., Razavi, S., DeRose, R. & Inoue, T. Synthesizing Biomolecule-Based Boolean Logic Gates. *ACS Synth. Biol.* **2**, 72–82 (2013).
453. Shoval, O. *et al.* Fold-change detection and scalar symmetry of sensory input fields. *Proc. Natl. Acad. Sci.* **107**, 15995–16000 (2010).
454. Bradley, R. W. & Wang, B. Designer cell signal processing circuits for biotechnology. *New Biotechnol.* (2015). doi:10.1016/j.nbt.2014.12.009
455. Roquet, N. & Lu, T. K. Digital and analog gene circuits for biotechnology. *Biotechnol. J.* **9**, 597–608 (2014).
456. Bonnet, J. & Endy, D. Switches, Switches, Every Where, In Any Drop We Drink. *Mol. Cell* **49**, 232–233 (2013).

457. Siuti, P., Yazbek, J. & Lu, T. K. Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* **31**, 448–452 (2013).
458. Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying Genetic Logic Gates. *Science* **340**, 599–603 (2013).
459. Bonnet, J., Subsoontorn, P. & Endy, D. Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl. Acad. Sci.* **109**, 8884–8889 (2012).
460. Khalil, A. S. & Collins, J. J. Synthetic biology: applications come of age. *Nat. Rev. Genet.* **11**, 367–379 (2010).
461. Lu, T. K., Khalil, A. S. & Collins, J. J. Next-generation synthetic gene networks. *Nat. Biotechnol.* **27**, 1139–1150 (2009).
462. Ang, J., Harris, E., Hussey, B. J., Kil, R. & McMillen, D. R. Tuning Response Curves for Synthetic Biology. *ACS Synth. Biol.* **2**, 547–567 (2013).
463. Arkin, A. P. A wise consistency: engineering biology for conformity, reliability, predictability. *Curr. Opin. Chem. Biol.* **17**, 893–901 (2013).
464. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
465. Lou, C., Stanton, B., Chen, Y.-J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
466. Cambray, G. *et al.* Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res.* **41**, 5139–5148 (2013).
467. Nielsen, A. A., Segall-Shapiro, T. H. & Voigt, C. A. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr. Opin. Chem. Biol.* **17**, 878–892 (2013).
468. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–866 (2012).
469. Mirasoli, M., Feliciano, J., Michelini, E., Daunert, S. & Roda, A. Internal response correction for fluorescent whole-cell biosensors. *Anal. Chem.* **74**, 5948–5953 (2002).
470. Bugaj, L. J. & Schaffer, D. V. Bringing next-generation therapeutics to the clinic through synthetic biology. *Curr. Opin. Chem. Biol.* **16**, 355–361 (2012).
471. Kwok, R. Five hard truths for synthetic biology. *Nature* **463**, 288–290 (2010).
472. Lapique, N. & Benenson, Y. Digital switching in a biosensor circuit via programmable timing of gene availability. *Nat. Chem. Biol.* **10**, 1020–1027 (2014).
473. Daniel, R., Rubens, J. R., Sarpeshkar, R. & Lu, T. K. Synthetic analog computation in living cells. *Nature* **497**, 619–623 (2013).
474. Sarpeshkar, R. Analog synthetic biology. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **372**, 20130110–20130110 (2014).

475. Zhang, H. *et al.* Rational design of a biosensor circuit with semi-log dose-response function in *Escherichia coli*. *Quant. Biol.* **1**, 209–220 (2013).
476. Purcell, O. & Lu, T. K. Synthetic analog and digital circuits for cellular computation and memory. *Curr. Opin. Biotechnol.* **29**, 146–155 (2014).
477. Mootz, H. D., Blum, E. S., Tyszkiewicz, A. B. & Muir, T. W. Conditional protein splicing: a new tool to control protein structure and function in vitro and in vivo. *J. Am. Chem. Soc.* **125**, 10561–10569 (2003).
478. Selgrade, D. F., Lohmueller, J. J., Lienert, F. & Silver, P. A. Protein Scaffold-Activated Protein Trans-Splicing in Mammalian Cells. *J. Am. Chem. Soc.* **135**, 7713–7719 (2013).
479. Park, S.-H., Zarrinpar, A. & Lim, W. A. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061–1064 (2003).
480. Bongor, K. M., Chen, L., Liu, C. W. & Wandless, T. J. Small-molecule displacement of a cryptic degron causes conditional protein degradation. *Nat. Chem. Biol.* **7**, 531–537 (2011).
481. Prindle, A. *et al.* Rapid and tunable post-translational coupling of genetic circuits. *Nature* **508**, 387–391 (2014).
482. O’Shaughnessy, E. C., Palani, S., Collins, J. J. & Sarkar, C. A. Tunable Signal Processing in Synthetic MAP Kinase Cascades. *Cell* **144**, 119–131 (2011).
483. Wang, B., Barahona, M., Buck, M. & Schumacher, J. Rewiring cell signalling through chimaeric regulatory protein engineering. *Biochem. Soc. Trans.* **41**, 1195–1200 (2013).
484. Lienert, F. *et al.* Two- and three-input TALE-based AND logic computation in embryonic stem cells. *Nucleic Acids Res.* **41**, 9967–9975 (2013).
485. Nielsen, A. A. & Voigt, C. A. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.* **10**, 763–763 (2014).
486. Lohmueller, J. J., Armel, T. Z. & Silver, P. A. A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res.* **40**, 5180–5187 (2012).
487. Zalatan, J. G. *et al.* Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* **160**, 339–350 (2015).
488. Khalil, A. S. *et al.* A Synthetic Biology Framework for Programming Eukaryotic Transcription Functions. *Cell* **150**, 647–658 (2012).
489. Kim, J., Khetarpal, I., Sen, S. & Murray, R. M. Synthetic circuit for exact adaptation and fold-change detection. *Nucleic Acids Res.* **42**, 6078–6089 (2014).
490. Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nat. Nanotechnol.* **8**, 755–762 (2013).
491. Katz, E. & Privman, V. Enzyme-based logic systems for information processing. *Chem. Soc. Rev.* **39**, 1835 (2010).
492. Pischel, U. Chemical approaches to molecular logic elements for addition and subtraction. *Angew. Chem. Int. Ed Engl.* **46**, 4026–4040 (2007).

493. Baron, R., Lioubashevski, O., Katz, E., Niazov, T. & Willner, I. Logic gates and elementary computing by enzymes. *J. Phys. Chem. A* **110**, 8548–8553 (2006).
494. Baron, R., Lioubashevski, O., Katz, E., Niazov, T. & Willner, I. Elementary arithmetic operations by enzymes: a model for metabolic pathway based computing. *Angew. Chem. Int. Ed Engl.* **45**, 1572–1576 (2006).
495. Hart, Y. & Alon, U. The Utility of Paradoxical Components in Biological Circuits. *Mol. Cell* **49**, 213–221 (2013).
496. Tam, T. K. *et al.* Biochemically Controlled Bioelectrocatalytic Interface. *J. Am. Chem. Soc.* **130**, 10888–10889 (2008).
497. Poghossian, A. *et al.* Interfacing of biocomputing systems with silicon chips: Enzyme logic gates based on field-effect devices. *Procedia Chem.* **1**, 682–685 (2009).
498. Wagler, P. F., Tangen, U., Maeke, T. & McCaskill, J. S. Field programmable chemistry: Integrated chemical and electronic processing of informational molecules towards electronic chemical cells. *Biosystems* **109**, 2–17 (2012).
499. Fu, E., Yager, P., Floriano, P. N., Christodoulides, N. & McDevitt, J. T. Perspective on Diagnostics for Global Health. *IEEE Pulse* **2**, 40–50 (2011).
500. de la Rica, R. & Stevens, M. M. Plasmonic ELISA for the ultrasensitive detection of disease biomarkers with the naked eye. *Nat. Nanotechnol.* **7**, 821–824 (2012).
501. Chi, X. *et al.* Nanoprobes for in vitro diagnostics of cancer and infectious diseases. *Biomaterials* **33**, 189–206 (2012).
502. Martinez, A. W., Phillips, S. T., Whitesides, G. M. & Carrilho, E. Diagnostics for the Developing World: Microfluidic Paper-Based Analytical Devices. *Anal. Chem.* **82**, 3–10 (2010).
503. Yager, P. *et al.* Microfluidic diagnostic technologies for global public health. *Nature* **442**, 412–418 (2006).
504. Ahmed, M. U., Saaem, I., Wu, P. C. & Brown, A. S. Personalized diagnostics and biosensors: a review of the biology and technology needed for personalized medicine. *Crit. Rev. Biotechnol.* **34**, 180–196 (2014).
505. Gubala, V., Harris, L. F., Ricco, A. J., Tan, M. X. & Williams, D. E. Point of Care Diagnostics: Status and Future. *Anal. Chem.* **84**, 487–515 (2012).
506. Su, L., Jia, W., Hou, C. & Lei, Y. Microbial biosensors: A review. *Biosens. Bioelectron.* **26**, 1788–1799 (2011).
507. American Society for Microbiology. *Manual of industrial microbiology and biotechnology*. (ASM Press, 2010).
508. Date, A., Pasini, P., Sangal, A. & Daunert, S. Packaging Sensing Cells in Spores for Long-Term Preservation of Sensors: A Tool for Biomedical and Environmental Analysis. *Anal. Chem.* **82**, 6098–6103 (2010).
509. Cortés-Salazar, F., Beggah, S., van der Meer, J. R. & Girault, H. H. Electrochemical As(III) whole-cell based biochip sensor. *Biosens. Bioelectron.* **47**, 237–242 (2013).

510. Michellini, E. *et al.* Recombinant cell-based bioluminescence assay for androgen bioactivity determination in clinical samples. *Clin. Chem.* **51**, 1995–1998 (2005).
511. Alloush, H. M. *et al.* A bioluminescent microbial biosensor for in vitro pretreatment assessment of cytarabine efficacy in leukemia. *Clin. Chem.* **56**, 1862–1870 (2010).
512. Duan, F., Curtis, K. L. & March, J. C. Secretion of Insulinotropic Proteins by Commensal Bacteria: Rewiring the Gut To Treat Diabetes. *Appl. Environ. Microbiol.* **74**, 7437–7438 (2008).
513. Chen, Y. Y., Jensen, M. C. & Smolke, C. D. Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proc. Natl. Acad. Sci.* **107**, 8531–8536 (2010).
514. Ham, T. S., Lee, S. K., Keasling, J. D. & Arkin, A. P. Design and Construction of a Double Inversion Recombination Switch for Heritable Sequential Genetic Memory. *PLoS ONE* **3**, e2815 (2008).
515. Groth, A. C. & Calos, M. P. Phage integrases: biology and applications. *J. Mol. Biol.* **335**, 667–678 (2004).
516. Brown, W. R. A., Lee, N. C. O., Xu, Z. & Smith, M. C. M. Serine recombinases as tools for genome engineering. *Methods* **53**, 372–379 (2011).
517. Kelly, J. R. *et al.* Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3**, 4 (2009).
518. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
519. Kittleson, J. T., Wu, G. C. & Anderson, J. C. Successes and failures in modular genetic engineering. *Curr. Opin. Chem. Biol.* **16**, 329–336 (2012).
520. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
521. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
522. Sharma, J. N., Al-Omran, A. & Parvathy, S. S. Role of nitric oxide in inflammatory diseases. *Inflammopharmacology* **15**, 252–259 (2007).
523. Lin, H.-Y., Bledsoe, P. J. & Stewart, V. Activation of *yeaR-yoaG* Operon Transcription by the Nitrate-Responsive Regulator NarL Is Independent of Oxygen- Responsive Regulator Fnr in *Escherichia coli* K-12. *J. Bacteriol.* **189**, 7539–7548 (2007).
524. Schaerli, Y. & Isalan, M. Building synthetic gene circuits from combinatorial libraries: screening and selection strategies. *Mol. Biosyst.* **9**, 1559 (2013).
525. Farrell, C. M., Grossman, A. D. & Sauer, R. T. Cytoplasmic degradation of *ssrA*-tagged proteins. *Mol. Microbiol.* **57**, 1750–1761 (2005).
526. Andersen, J. B. *et al.* New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.* **64**, 2240–2246 (1998).
527. Arpino, J. A. J. *et al.* Tuning the dials of Synthetic Biology. *Microbiology* **159**, 1236–1253 (2013).

528. Cheetham, P. S. J., Blunt, K. W. & Bocke, C. Physical Studies on Cell Immobilization Using Calcium Alginate Gels. *Biotechnol. Bioeng.* **21**, 2155–2168 (1979).
529. Takei, T., Ikeda, K., Ijima, H. & Kawakami, K. Fabrication of poly(vinyl alcohol) hydrogel beads crosslinked using sodium sulfate for microorganism immobilization. *Process Biochem.* **46**, 566–571 (2011).
530. Zain, N. A. M., Suhaimi, M. S. & Idris, A. Development and modification of PVA–alginate as a suitable immobilization matrix. *Process Biochem.* **46**, 2122–2129 (2011).
531. American Diabetes Association. Standards of Medical Care in Diabetes--2014. *Diabetes Care* **37**, S14–S80 (2014).
532. Lima, B. P. *et al.* Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter: Acetylation of RNA polymerase. *Mol. Microbiol.* **81**, 1190–1204 (2011).
533. Wolfe, A. J., Parikh, N., Lima, B. P. & Zemaitaitis, B. Signal Integration by the Two-Component Signal Transduction Response Regulator CpxR. *J. Bacteriol.* **190**, 2314–2322 (2008).
534. Bauer, M. & Reinhart, K. Molecular diagnostics of sepsis--where are we today? *Int. J. Med. Microbiol. IJMM* **300**, 411–413 (2010).
535. Lever, A. & Mackenzie, I. Sepsis: definition, epidemiology, and diagnosis. *BMJ* **335**, 879–883 (2007).
536. Nguyen, H. B. *et al.* Severe sepsis and septic shock: review of the literature and emergency department management guidelines. *Ann. Emerg. Med.* **48**, 28–54 (2006).
537. Kumar, A. *et al.* Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit. Care Med.* **34**, 1589–1596 (2006).
538. Mancini, N. *et al.* The Era of Molecular and Other Non-Culture-Based Methods in Diagnosis of Sepsis. *Clin. Microbiol. Rev.* **23**, 235–251 (2010).
539. Carrigan, S. D. Toward Resolving the Challenges of Sepsis Diagnosis. *Clin. Chem.* **50**, 1301–1314 (2004).
540. Gonsalves, M. D. & Sakr, Y. Early identification of sepsis. *Curr. Infect. Dis. Rep.* **12**, 329–335 (2010).
541. Rådström, P., Knutsson, R., Wolffs, P., Lövenklev, M. & Löfström, C. Pre-PCR processing: strategies to generate PCR-compatible samples. *Mol. Biotechnol.* **26**, 133–146 (2004).
542. Pierrakos, C. & Vincent, J.-L. Sepsis biomarkers: a review. *Crit. Care* **14**, R15 (2010).
543. van der Poll, T. & Opal, S. M. Host-pathogen interactions in sepsis. *Lancet Infect. Dis.* **8**, 32–43 (2008).
544. Miller, M. B. & Bassler, B. L. Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **55**, 165–199 (2001).
545. Kumari, A. *et al.* Biosensing systems for the detection of bacterial quorum signaling molecules. *Anal. Chem.* **78**, 7603–7609 (2006).

546. Kumari, A., Pasini, P. & Daunert, S. Detection of bacterial quorum sensing N-acyl homoserine lactones in clinical samples. *Anal. Bioanal. Chem.* **391**, 1619–1627 (2008).
547. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
548. Winson, M. K. *et al.* Construction and analysis of *luxCDABE* -based plasmid sensors for investigating N -acyl homoserine lactone-mediated quorum sensing. *FEMS Microbiol. Lett.* **163**, 185–192 (1998).
549. Lindsay, A. & Ahmer, B. M. M. Effect of *sdiA* on Biosensors of N-Acylhomoserine Lactones. *J. Bacteriol.* **187**, 5054–5058 (2005).
550. Swift, S. *et al.* Quorum sensing in *Aeromonas hydrophila* and *Aeromonas salmonicida*: identification of the LuxRI homologs AhyRI and AsaRI and their cognate N-acylhomoserine lactone signal molecules. *J. Bacteriol.* **179**, 5271–5281 (1997).
551. Khajanchi, B. K., Kirtley, M. L., Brackman, S. M. & Chopra, A. K. Immunomodulatory and Protective Roles of Quorum-Sensing Signaling Molecules N-Acyl Homoserine Lactones during Infection of Mice with *Aeromonas hydrophila*. *Infect. Immun.* **79**, 2646–2657 (2011).
552. Win, M. N., Liang, J. C. & Smolke, C. D. Frameworks for Programming Biological Function through RNA Parts and Devices. *Chem. Biol.* **16**, 298–310 (2009).
553. Sadelain, M., Brentjens, R. & Rivière, I. The promise and potential pitfalls of chimeric antigen receptors. *Curr. Opin. Immunol.* **21**, 215–223 (2009).
554. Archana Chugh, P. B. Synthetic Biology Based Biosensors and the Emerging Governance Issues. *Curr. Synth. Syst. Biol.* **01**, (2013).
555. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA Assembly Design Automation Software. *ACS Synth. Biol.* **1**, 14–21 (2012).
556. van Roekel, H. W. H. *et al.* Programmable chemical reaction networks: emulating regulatory functions in living cells using a bottom-up approach. *Chem Soc Rev* (2015). doi:10.1039/C5CS00361J
557. Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A. & Weiss, R. Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell Biol.* **15**, 289–294 (2014).
558. Shetty, R. P., Endy, D. & Knight, T. F. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 5 (2008).
559. Breitling, R. & Takano, E. Synthetic biology advances for pharmaceutical production. *Curr. Opin. Biotechnol.* **35**, 46–51 (2015).
560. Ye, H. & Fussenegger, M. Synthetic therapeutic gene circuits in mammalian cells. *FEBS Lett.* **588**, 2537–2544 (2014).
561. Semenov, S. N. *et al.* Rational design of functional and tunable oscillating enzymatic networks. *Nat. Chem.* **7**, 160–165 (2015).
562. Caschera, F. & Noireaux, V. Integration of biological parts toward the synthesis of a minimal cell. *Curr. Opin. Chem. Biol.* **22**, 85–91 (2014).

563. van Roekel, H. W. H. *et al.* Automated Design of Programmable Enzyme-Driven DNA Circuits. *ACS Synth. Biol.* **4**, 735–745 (2015).
564. Marchisio, M. A. & Stelling, J. Automatic Design of Digital Synthetic Gene Circuits. *PLoS Comput. Biol.* **7**, e1001083 (2011).
565. Alexander Ivanovich Oparin. *The Origin of Life*. (1923).
566. Erwin Schrödinger. *What is Life?* (Cambridge University Press, 1944).
567. Luisi, P. L. & Stano, P. *The minimal cell the biophysics of cell compartment and the origin of cell functionality*. (Springer, 2011). at <<http://site.ebrary.com/id/10427796>>
568. *Protocells: bridging nonliving and living matter*. (MIT Press, 2009).
569. Bahadur, K. Synthesis of Jeewanu, the Protocell. *Zentralblatt Für Bakteriologie, Parasitenkunde, Infektionshygiene und Antimikrobiologie*. **121**, 291–319 (1967).
570. Noireaux, V., Maeda, Y. T. & Libchaber, A. Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc. Natl. Acad. Sci.* **108**, 3473–3480 (2011).
571. Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454**, 122–125 (2008).
572. Huang, X., Patil, A. J., Li, M. & Mann, S. Design and Construction of Higher-Order Structure and Function in Proteinosome-Based Protocells. *J. Am. Chem. Soc.* **136**, 9225–9234 (2014).
573. Cooper, G. J. T. *et al.* Modular Redox-Active Inorganic Chemical Cells: iCHELLs. *Angew. Chem. Int. Ed.* **50**, 10373–10376 (2011).
574. Li, M., Green, D. C., Anderson, J. L. R., Binks, B. P. & Mann, S. In vitro gene expression and enzyme catalysis in bio-inorganic protocells. *Chem. Sci.* **2**, 1739 (2011).
575. Forster, A. C. & Church, G. M. Synthetic biology projects in vitro. *Genome Res.* **17**, 1–6 (2007).
576. Sole, R. V., Munteanu, A., Rodriguez-Caso, C. & Macia, J. Synthetic protocell biology: from reproduction to computation. *Philos. Trans. R. Soc. B Biol. Sci.* **362**, 1727–1739 (2007).
577. Schwille, P. Bottom-Up Synthetic Biology: Engineering in a Tinkerer's World. *Science* **333**, 1252–1254 (2011).
578. Walleczek, J. *Self-organized biological dynamics & nonlinear control toward understanding complexity, chaos, and emergent function in living systems*. (Cambridge University Press, 2006).
579. Sugita, M. Functional analysis of chemical systems in vivo using a logical circuit equivalent. *J. Theor. Biol.* **1**, 415–430 (1961).
580. Nakajima, M. Reconstitution of Circadian Oscillation of Cyanobacterial KaiC Phosphorylation in Vitro. *Science* **308**, 414–415 (2005).
581. Padirac, A., Fujii, T. & Rondelez, Y. Bottom-up construction of in vitro switchable memories. *Proc. Natl. Acad. Sci.* **109**, E3212–E3220 (2012).

582. Tyson, J. J. & Novák, B. Functional Motifs in Biochemical Reaction Networks. *Annu. Rev. Phys. Chem.* **61**, 219–240 (2010).
583. Stojanovic, M. N., Stefanovic, D. & Rudchenko, S. Exercises in Molecular Computing. *Acc. Chem. Res.* **47**, 1845–1852 (2014).
584. Adleman, L. M. Molecular computation of solutions to combinatorial problems. *Science* **266**, 1021–1024 (1994).
585. Faulhammer, D., Cukras, A. R., Lipton, R. J. & Landweber, L. F. Molecular computation: RNA solutions to chess problems. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 1385–1389 (2000).
586. Bray, D. Protein molecules as computational elements in living cells. *Nature* **376**, 307–312 (1995).
587. Mehta, P., Lang, A. H. & Schwab, D. J. *Landauer in the age of synthetic biology: energy consumption and information processing in biochemical networks.* (2015). at <<http://biorxiv.org/lookup/doi/10.1101/020594>>
588. Chandran, D., Bergmann, F. T., Sauro, H. M. & Densmore, D. in *Design and Analysis of Biomolecular Circuits* (eds. Koeppl, H., Setti, G., di Bernardo, M. & Densmore, D.) 203–224 (Springer New York, 2011). at <http://link.springer.com/10.1007/978-1-4419-6766-4_10>
589. Koeppl, H. *Design and analysis of bio-molecular circuits.* (Springer, 2011).
590. Shinar, G. & Feinberg, M. Structural Sources of Robustness in Biochemical Reaction Networks. *Science* **327**, 1389–1391 (2010).
591. Marchisio, M. A. & Stelling, J. Computational design tools for synthetic biology. *Curr. Opin. Biotechnol.* **20**, 479–485 (2009).
592. Mendes, P. *et al.* in *Systems Biology* (ed. Maly, I. V.) **500**, 17–59 (Humana Press, 2009).
593. Matsuoka, Y., Funahashi, A., Ghosh, S. & Kitano, H. in *Transcription Factor Regulatory Networks* (eds. Miyamoto-Sato, E., Ohashi, H., Sasaki, H., Nishikawa, J. & Yanagawa, H.) **1164**, 121–145 (Springer New York, 2014).
594. Kaznessis, Y. N. in *Methods in Enzymology* **498**, 137–152 (Elsevier, 2011).
595. Chaouiya, C. *et al.* SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools. *BMC Syst. Biol.* **7**, 135 (2013).
596. Zheng, Y. & Sriram, G. Mathematical Modeling: Bridging the Gap between Concept and Realization in Synthetic Biology. *J. Biomed. Biotechnol.* **2010**, 1–16 (2010).
597. Kaznessis, Y. N. Models for synthetic biology. *BMC Syst. Biol.* **1**, 47 (2007).
598. Lewis, D. D., Villarreal, F. D., Wu, F. & Tan, C. Synthetic Biology Outside the Cell: Linking Computational Tools to Cell-Free Systems. *Front. Bioeng. Biotechnol.* **2**, (2014).
599. Amar, P. HSIM: a simulation programme to study large assemblies of proteins. *J. Biol. Phys. Chem.* **4**, 79–84 (2002).
600. Amar, P. & Paulevé, L. HSIM: A Hybrid Stochastic Simulation System for Systems Biology. *Electron. Notes Theor. Comput. Sci.* **313**, 3–21 (2015).

601. Calzone, L., Fages, F. & Soliman, S. BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge. *Bioinforma. Oxf. Engl.* **22**, 1805–1807 (2006).
602. Gay, S., Soliman, S. & Fages, F. A graphical method for reducing and relating models in systems biology. *Bioinformatics* **26**, i575–i581 (2010).
603. Fages, F. & Soliman, S. in *Formal Methods for Computational Systems Biology* (eds. Bernardo, M., Degano, P. & Zavattaro, G.) **5016**, 54–80 (Springer Berlin Heidelberg, 2008).
604. Gulati, S. *et al.* Opportunities for microfluidic technologies in synthetic biology. *J. R. Soc. Interface* **6**, S493–S506 (2009).
605. Vinuselvi, P. *et al.* Microfluidic Technologies for Synthetic Biology. *Int. J. Mol. Sci.* **12**, 3576–3593 (2011).
606. Beebe, D. J., Mensing, G. A. & Walker, G. M. Physics and Applications of Microfluidics in Biology. *Annu. Rev. Biomed. Eng.* **4**, 261–286 (2002).
607. Ng, J. M. K., Gitlin, I., Stroock, A. D. & Whitesides, G. M. Components for integrated poly(dimethylsiloxane) microfluidic systems. *Electrophoresis* **23**, 3461–3473 (2002).
608. Quake, S. R. & Scherer, A. From micro- to nanofabrication with soft materials. *Science* **290**, 1536–1540 (2000).
609. Korir, G. & Prakash, M. Punch Card Programmable Microfluidics. *PLOS ONE* **10**, e0115993 (2015).
610. Sackmann, E. K., Fulton, A. L. & Beebe, D. J. The present and future role of microfluidics in biomedical research. *Nature* **507**, 181–189 (2014).
611. White, A. K., Heyries, K. A., Doolin, C., VanInsberghe, M. & Hansen, C. L. High-Throughput Microfluidic Single-Cell Digital Polymerase Chain Reaction. *Anal. Chem.* **85**, 7182–7190 (2013).
612. Hansen, C. & Quake, S. R. Microfluidics in structural biology: smaller, faster em leader better. *Curr. Opin. Struct. Biol.* **13**, 538–544 (2003).
613. Abate, A. R. *et al.* DNA sequence analysis with droplet-based microfluidics. *Lab. Chip* **13**, 4864 (2013).
614. Sanchez-Freire, V., Ebert, A. D., Kalisky, T., Quake, S. R. & Wu, J. C. Microfluidic single-cell real-time PCR for comparative analysis of gene expression patterns. *Nat. Protoc.* **7**, 829–838 (2012).
615. Lee, C.-Y., Chang, C.-L., Wang, Y.-N. & Fu, L.-M. Microfluidic Mixing: A Review. *Int. J. Mol. Sci.* **12**, 3263–3287 (2011).
616. Andrea Biral. MICROFLUIDIC NETWORKING: MODELLING AND ANALYSIS. (UNIVERSITA DEGLI STUDI DI PADOVA, 2012).
617. Bouffar, M., Molina, F. & Amar, P. Extracting logic gates from a metabolic network. in pp. 13 (2015).
618. Hansen, N. & Ostermeier, A. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **9**, 159–195 (2001).

619. Beal, J. Signal-to-Noise Ratio Measures Efficacy of Biological Computing Devices and Circuits. *Front. Bioeng. Biotechnol.* **3**, (2015).
620. Walde, P., Cosentino, K., Engel, H. & Stano, P. Giant Vesicles: Preparations and Applications. *ChemBioChem* **11**, 848–865 (2010).
621. Reeves, J. P. & Dowben, R. M. Formation and properties of thin-walled phospholipid vesicles. *J. Cell. Physiol.* **73**, 49–60 (1969).
622. Pautot, S., Frisken, B. J. & Weitz, D. A. Engineering asymmetric vesicles. *Proc. Natl. Acad. Sci.* **100**, 10718–10721 (2003).
623. van Swaay, D. & deMello, A. Microfluidic methods for forming liposomes. *Lab. Chip* **13**, 752 (2013).
624. Jahn, A., Vreeland, W. N., Gaitan, M. & Locascio, L. E. Controlled vesicle self-assembly in microfluidic channels with hydrodynamic focusing. *J. Am. Chem. Soc.* **126**, 2674–2675 (2004).
625. Richmond, D. L. *et al.* Forming giant vesicles with controlled membrane composition, asymmetry, and contents. *Proc. Natl. Acad. Sci.* **108**, 9431–9436 (2011).
626. Tan, Y.-C., Hettiarachchi, K., Siu, M., Pan, Y.-R. & Lee, A. P. Controlled microfluidic encapsulation of cells, proteins, and microbeads in lipid vesicles. *J. Am. Chem. Soc.* **128**, 5656–5658 (2006).
627. Matosevic, S. & Paegel, B. M. Stepwise Synthesis of Giant Unilamellar Vesicles on a Microfluidic Assembly Line. *J. Am. Chem. Soc.* **133**, 2798–2800 (2011).
628. Teh, S.-Y., Khnouf, R., Fan, H. & Lee, A. P. Stable, biocompatible lipid vesicle generation by solvent extraction-based droplet microfluidics. *Biomicrofluidics* **5**, 044113 (2011).
629. Jahn, A. *et al.* Microfluidic Mixing and the Formation of Nanoscale Lipid Vesicles. *ACS Nano* **4**, 2077–2087 (2010).
630. Mijajlovic, M., Wright, D., Zivkovic, V., Bi, J. X. & Biggs, M. J. Microfluidic hydrodynamic focusing based synthesis of POPC liposomes for model biological systems. *Colloids Surf. B Biointerfaces* **104**, 276–281 (2013).
631. Jahn, A., Vreeland, W. N., DeVoe, D. L., Locascio, L. E. & Gaitan, M. Microfluidic directed formation of liposomes of controlled size. *Langmuir ACS J. Surf. Colloids* **23**, 6289–6293 (2007).
632. Shum, H. C., Lee, D., Yoon, I., Kodger, T. & Weitz, D. A. Double Emulsion Templated Monodisperse Phospholipid Vesicles. *Langmuir* **24**, 7651–7653 (2008).
633. Shum, H. C., Kim, J.-W. & Weitz, D. A. Microfluidic Fabrication of Monodisperse Biocompatible and Biodegradable Polymersomes with Controlled Permeability. *J. Am. Chem. Soc.* **130**, 9543–9549 (2008).
634. Günther, A. & Jensen, K. F. Multiphase microfluidics: from flow characteristics to chemical and materials synthesis. *Lab. Chip* **6**, 1487–1503 (2006).
635. Utada, A. S. Monodisperse Double Emulsions Generated from a Microcapillary Device. *Science* **308**, 537–541 (2005).

636. Kozlov, M., Quarmyne, M., Chen, W. & McCarthy, T. J. Adsorption of Poly(vinyl alcohol) onto Hydrophobic Substrates. A General Approach for Hydrophilizing and Chemically Activating Surfaces. *Macromolecules* **36**, 6054–6059 (2003).
637. *Liposomes as tools in basic research and industry*. (CRC Press, 1995).
638. Williams, M. S., Longmuir, K. J. & Yager, P. A practical guide to the staggered herringbone mixer. *Lab. Chip* **8**, 1121 (2008).
639. Gullapalli, R. R., Demirel, M. C. & Butler, P. J. Molecular dynamics simulations of DiI-C18(3) in a DPPC lipid bilayer. *Phys. Chem. Chem. Phys.* **10**, 3548 (2008).
640. Siegel, J. B. *et al.* Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **329**, 309–313 (2010).
641. Maziere, P., Parisey, N., Beurton-Aimar, M. & Molina, F. Formal TCA cycle description based on elementary actions. *J. Biosci.* **32**, 145–155 (2007).
642. Mazière, P., Granier, C. & Molina, F. A description scheme of biological processes based on elementary bricks of action. *J. Mol. Biol.* **339**, 77–88 (2004).
643. Nagle JF & Tristram-Nagle S. Structure of lipid bilayers. *Biochim Biophys Acta.* 159–95. (2000).
644. Aksimentiev, A. & Schulten, K. Imaging α -Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability, and the Electrostatic Potential Map. *Biophys. J.* **88**, 3745–3761 (2005).
645. Watanabe, R. *et al.* Arrayed lipid bilayer chambers allow single-molecule analysis of membrane transporter activity. *Nat. Commun.* **5**, (2014).
646. Planchet, E. & Kaiser, W. M. Nitric oxide (NO) detection by DAF fluorescence and chemiluminescence: a comparison using abiotic and biotic NO sources. *J. Exp. Bot.* **57**, 3043–3055 (2006).
647. Dean, J. V. & Harper, J. E. The Conversion of Nitrite to Nitrogen Oxide(s) by the Constitutive NAD(P)H-Nitrate Reductase Enzyme from Soybean. *Plant Physiol.* **88**, 389–395 (1988).
648. Yokota, K. & Yamazaki, I. Analysis and computer simulation of aerobic oxidation of reduced nicotinamide adenine dinucleotide catalyzed by horseradish peroxidase. *Biochemistry (Mosc.)* **16**, 1913–1920 (1977).
649. Singh, R. *et al.* Catalase-peroxidases (KatG) exhibit NADH oxidase activity. *J. Biol. Chem.* **279**, 43098–43106 (2004).
650. Afanasyeva, M. S., Taraban, M. B., Purtov, P. A., Leshina, T. V. & Grissom, C. B. Magnetic spin effects in enzymatic reactions: radical oxidation of NADH by horseradish peroxidase. *J. Am. Chem. Soc.* **128**, 8651–8658 (2006).
651. Yorita, K. *et al.* Conversion of L-lactate oxidase to a long chain alpha-hydroxyacid oxidase by site-directed mutagenesis of alanine 95 to glycine. *J. Biol. Chem.* **271**, 28300–28305 (1996).
652. Liu, X. *et al.* Application of carbon fiber composite minielectrodes for measurement of kinetic constants of nitric oxide decay in solution. *Nitric Oxide* **23**, 311–318 (2010).
653. Li, H., Kundu, T. K. & Zweier, J. L. Characterization of the Magnitude and Mechanism of Aldehyde Oxidase-mediated Nitric Oxide Production from Nitrite. *J. Biol. Chem.* **284**, 33850–33858 (2009).

654. Kojima, H. *et al.* Detection and imaging of nitric oxide with novel fluorescent indicators: diaminofluoresceins. *Anal. Chem.* **70**, 2446–2453 (1998).
655. Espey, M. G., Miranda, K. M., Thomas, D. D. & Wink, D. A. Distinction between nitrosating mechanisms within human cells and aqueous solution. *J. Biol. Chem.* **276**, 30085–30091 (2001).
656. Arita, N. O., Cohen, M. F., Tokuda, G. & Yamasaki, H. in *Nitric Oxide in Plant Growth, Development and Stress Physiology* (eds. Lamattina, L. & Polacco, J. C.) **5**, 269–280 (Springer Berlin Heidelberg, 2007).
657. E.F. Olasehinde & *et al.* Reaction Kinetics for Nitrosation of DAF-2 in Air Saturated Nitric Oxide Solution. *Nature & Science* p129 (2012).
658. Stillwell, W. *An introduction to biological membranes: from bilayers to rafts.* (Elsevier/Academic Press, 2013).
659. Karp, R. M. in *Complexity of Computer Computations* (eds. Miller, R. E., Thatcher, J. W. & Bohlinger, J. D.) 85–103 (Springer US, 1972). at <http://link.springer.com/10.1007/978-1-4684-2001-2_9>
660. Cook, S. A. The complexity of theorem-proving procedures. in 151–158 (ACM Press, 1971). doi:10.1145/800157.805047
661. Chaize, B., Colletier, J.-P., Winterhalter, M. & Fournier, D. Encapsulation of enzymes in liposomes: high encapsulation efficiency and control of substrate permeability. *Artif. Cells. Blood Substit. Immobil. Biotechnol.* **32**, 67–75 (2004).
662. Yoshimoto, M. in *Enzyme Stabilization and Immobilization* (ed. Minteer, S. D.) **679**, 9–18 (Humana Press, 2011).
663. Sunami, T., Hosoda, K., Suzuki, H., Matsuura, T. & Yomo, T. Cellular Compartment Model for Exploring the Effect of the Lipidic Membrane on the Kinetics of Encapsulated Biochemical Reactions. *Langmuir* **26**, 8544–8551 (2010).
664. Thiele, J. *et al.* Fabrication of Polymersomes using Double-Emulsion Templates in Glass-Coated Stamped Microfluidic Devices. *Small* **6**, 1723–1727 (2010).
665. Duncanson, W. J. *et al.* Microfluidic synthesis of advanced microparticles for encapsulation and controlled release. *Lab. Chip* **12**, 2135 (2012).
666. Stanish, I. & Singh, A. Highly stable vesicles composed of a new chain-terminus acetylenic photopolymeric phospholipid. *Chem. Phys. Lipids* **112**, 99–108 (2001).
667. Osyczka, A., Moser, C. C., Daldal, F. & Dutton, P. L. Reversible redox energy coupling in electron transfer chains. *Nature* **427**, 607–612 (2004).
668. Katzen, F., Deshmukh, M., Daldal, F. & Beckwith, J. Evolutionary domain fusion expanded the substrate specificity of the transmembrane electron transporter DsbD. *EMBO J.* **21**, 3960–3969 (2002).
669. Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* **402**, 47–52 (1999).
670. Wakeham, M. C. & Jones, M. R. Rewiring photosynthesis: engineering wrong-way electron transfer in the purple bacterial reaction centre. *Biochem. Soc. Trans.* **33**, 851–857 (2005).

671. Hermann, T. & Patel, D. J. Adaptive recognition by nucleic acid aptamers. *Science* **287**, 820–825 (2000).
672. Šmuc, T., Ahn, I.-Y. & Ulrich, H. Nucleic acid aptamers as high affinity ligands in biotechnology and biosensorics. *J. Pharm. Biomed. Anal.* **81-82**, 210–217 (2013).
673. Falciani, C., Lozzi, L., Pini, A. & Bracci, L. Bioactive peptides from libraries. *Chem. Biol.* **12**, 417–426 (2005).
674. Stoltenburg, R., Reinemann, C. & Strehlitz, B. SELEX--a (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol. Eng.* **24**, 381–403 (2007).
675. Hanes, J., Schaffitzel, C., Knappik, A. & Plückthun, A. Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat. Biotechnol.* **18**, 1287–1292 (2000).
676. Binz, H. K., Amstutz, P. & Plückthun, A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **23**, 1257–1268 (2005).
677. Song, S., Kole, S. & Bernier, M. A chemical cross-linking method for the analysis of binding partners of heat shock protein-90 in intact cells. *BioTechniques* (2012). doi:10.2144/000113856
678. Xiang, Z. *et al.* Proximity-Enabled Protein Crosslinking through Genetically Encoding Haloalkane Unnatural Amino Acids. *Angew. Chem. Int. Ed.* **53**, 2190–2193 (2014).
679. Strickland, D. *et al.* TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* **9**, 379–384 (2012).
680. Rakhit, R., Navarro, R. & Wandless, T. J. Chemical biology strategies for posttranslational control of protein function. *Chem. Biol.* **21**, 1238–1252 (2014).
681. Riggsbee, C. W. & Deiters, A. Recent advances in the photochemical control of protein function. *Trends Biotechnol.* **28**, 468–475 (2010).
682. Schwartz, W. B. Medicine and the computer. The promise and problems of change. *N. Engl. J. Med.* **283**, 1257–1264 (1970).
683. Szolovits, P. & Pauker, S. G. Categorical and probabilistic reasoning in medical diagnosis☆. *Artif. Intell.* **11**, 115–144 (1978).
684. Weiss, S. M., Kulikowski, C. A., Amarel, S. & Safir, A. A model-based method for computer-aided medical decision-making. *Artif. Intell.* **11**, 145–172 (1978).
685. Zakeri, B. & Carr, P. A. The limits of synthetic biology. *Trends Biotechnol.* **33**, 57–58 (2015).
686. König, H., Frank, D., Heil, R. & Coenen, C. Synthetic Genomics and Synthetic Biology Applications Between Hopes and Concerns. *Curr. Genomics* **14**, 11–24 (2013).
687. Redford, K. H., Adams, W., Carlson, R., Mace, G. M. & Ceccarelli, B. Synthetic biology and the conservation of biodiversity. *Oryx* **48**, 330–336 (2014).
688. Dana, G. V., Kuiken, T., Rejeski, D. & Snow, A. A. Synthetic biology: Four steps to avoid a synthetic-biology disaster. *Nature* **483**, 29–29 (2012).

689. Mandell, D. J. *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).
690. Wright, O., Stan, G.-B. & Ellis, T. Building-in biosafety for synthetic biology. *Microbiology* **159**, 1221–1235 (2013).
691. Breitling, R., Takano, E. & Gardner, T. S. Judging synthetic biology risks. *Science* **347**, 107–107 (2015).
692. Kelley, N. J. *et al.* Engineering Biology to Address Global Problems: Synthetic Biology Markets, Needs, and Applications. *Ind. Biotechnol.* **10**, 140–149 (2014).
693. Ausländer, S., Wieland, M. & Fussenegger, M. Smart medication through combination of synthetic biology and cell microencapsulation. *Metab. Eng.* **14**, 252–260 (2012).
694. Seliktar, D. Designing Cell-Compatible Hydrogels for Biomedical Applications. *Science* **336**, 1124–1128 (2012).
695. Synthetic Biology Market - Global Industry Analysis, Size, Growth, Share And Forecast, 2012 - 2018. at <http://www.researchandmarkets.com/reports/2225060/synthetic_biology_market_global_industry>
696. Medical Sensors Market- Global Industry Analysis, Size, Share and Forecast 2013 - 2019. at <<http://www.transparencymarketresearch.com/medical-sensors-market.html>>
697. Sollis, K. A. *et al.* Systematic Review of the Performance of HIV Viral Load Technologies on Plasma Samples. *PLoS ONE* **9**, e85869 (2014).
698. Liu, C.-H. *et al.* Comparison of Abbott RealTime HCV Genotype II with Versant Line Probe Assay 2.0 for Hepatitis C Virus Genotyping. *J. Clin. Microbiol.* **53**, 1754–1757 (2015).
699. Windmiller, J. R. *et al.* Boolean-format biocatalytic processing of enzyme biomarkers for the diagnosis of soft tissue injury. *Sens. Actuators B Chem.* **150**, 285–290 (2010).
700. Manesh, K. M. *et al.* Enzyme logic gates for the digital analysis of physiological level upon injury. *Biosens. Bioelectron.* **24**, 3569–3574 (2009).
701. Rantala, A. *et al.* Luminescent bacteria-based sensing method for methylmercury specific determination. *Anal. Bioanal. Chem.* **400**, 1041–1049 (2011).
702. Rössger, K., Charpin-El-Hamri, G. & Fussenegger, M. Bile acid-controlled transgene expression in mammalian cells and mice. *Metab. Eng.* **21**, 81–90 (2014).
703. Gupta, S., Bram, E. E. & Weiss, R. Genetically Programmable Pathogen Sense and Destroy. *ACS Synth. Biol.* **2**, 715–723 (2013).

Selection of recent advances in synthetic biology of interest to the field of diagnostics

•, ••, ••• indicate increasing study importance and medical relevance

Designation /Importance	Technology and approach	Input biomarkers	Output/Readout	Device format	Targeted pathology/indication	Clinics /Lab	Ref
VERSANT HCV, HBV, HIV-1 RNA 3.0 Assay (bDNA) •••	Target RNA is detected via hybridisation using a series of capture probes. The target RNA then complexes with a fluorescent label probe through a series of hybridisation events involving the target probe, preamplifier and amplifier nucleic acids. The use of non-canonical nucleosides in the amplifier, preamplifier and label probe increases the specificity of the assay, as non-target DNA present in the sample cannot nonspecifically hybridise with the artificial DNA.	Nucleic acids	Fluorescence	Microplate	Detect HCV in infected patients, measurement of HCV, HBV and HIV-1 viral loads during and after antiviral therapy	Clinics	697 698 291
Bactosensors as a programmable platform for cell-based diagnostics •••	Bacterial biosensors with genetically encoded digital amplifying genetic switches can detect clinically relevant biomarkers in human urine and serum. They perform signal digitization and amplification, multiplexed signal processing with the use of Boolean logic gates, and data storage. We also provide a framework to quantify robustness in clinical samples and a method for easily reprogramming the sensor module for distinct medical detection agendas. First demonstration that bactosensors can be used to detect pathological signals.	Any molecular signal sensed by bacteria	Fluorescent, colorimetric	Polymer beads	Diabetes	Lab	115
Programmable probiotics for detection of cancer in urine •••	First example of orally administered diagnostic <i>in vivo</i> that can noninvasively indicate the presence of cancerogenesis by producing easily detectable signals in urine. No deleterious health effects on the mice bearing engineered bacteria where detected. They demonstrate that probiotics can be programmed to safely and selectively deliver synthetic gene circuits to diseased tissue microenvironments <i>in vivo</i> .		LacZ reporter in urines, colorimetric	Orally administered probiotic	Liver cancer	Lab (mice)	128
Intelligent Logic via aptasensors based Biofuel cells ••	First example of controlled power release of biofuel cells by aptamer-based biochemical signals processed according to the Boolean logic operations, to generate self-powered smart medical diagnostics "programmed" into biocomputing systems.	Thrombin and lysozyme	Electrochemical	On-Chip (microfluidics)	Proof of concept	Lab	261
"Sense-Act-treat" Biofuel cell ••	Self-powered biocomputing logic-controlled intelligent integrated "Sense-Act-Treat" system based on a BFC	Lactic acid and lactate dehydrogenase (LDH)	Release of therapeutic drug (acetaminophen)		Abdominal Trauma	Lab	438
Boolean-format biocatalytic processing of enzyme biomarkers ••	Enzymatically-processed biochemical information presented in the form of a NAND truth table allowed for high-fidelity discrimination between normal (physiological) and abnormal (pathological)	Creatine kinase (CK) and lactate dehydrogenase (LDH)	Electrochemical	Point of care	soft tissue injury	Lab	699
Biocomputing enzyme logic system •	Biocomputing system composed of a combination of enzyme logic gates designed to process biochemical information related to pathophysiological conditions originating from various injuries.	lactate, norepinephrine and glucose	Optical and electrochemical		traumatic brain injury and hemorrhagic shock	Lab	700
Cell-based allergy profiler •••	Mammalian cell-based biosensors that scores the allergen-triggered release of histamine from whole-blood-derived human basophils. A synthetic signalling cascade engineered within the allergy profiler rewires histamine input to the production of reporter protein, thereby integrating histamine levels in whole-blood samples.	various allergens	Fluorescence /enzyme assay	Liquid phase	Allergic disorders	Lab	97
Programmed engineered genetic circuit in cells that respond to biological signals •	Modular design strategy to create <i>Escherichia coli</i> strains where a genetic toggle switch is interfaced with: (i) the SOS signaling pathway responding to DNA damage, and (ii) a transgenic quorum sensing signaling pathway from <i>Vibrio fischeri</i> .	DNA damage, QS molecules (AHL)	Fluorescent protein (GFP)		Proof-of-concept	Lab	115

Designation /Importance	Technology and approach	Input biomarkers	Output/Readout	Device format	Targeted pathology/indication	Clinics /Lab	Ref
Delivery of exogenous synthetic agents for noninvasive disease monitoring ●●●	Low-cost, non-invasive method that relies on nanoscale agents that are administered to reveal the presence of diseased tissues by producing a biomarker via proteolytic release, in the urine that can be detected using different methods	Synthetic biomarkers	Molecular signatures of biomarkers in blood and urine, readable by MS, single molecule array, or lateral flow assay		Cardiovascular diseases, liver fibrosis and cancer	Lab (mice)	82 83 84
Programmable autonomous biomolecular computing device ●●●	Context-sensing mechanism of a biomolecular automaton that can simultaneously sense different types of molecules	mRNAs, miRNAs, proteins, and small molecules	nucleotide quantification using PAGE analysis		Proof-of-concept	Lab	285
Paper Strip cell-base biosensors for detection of QS signals ●●	Development of a fast, inexpensive, and portable filter-paper-based strip biosensor for the detection of bacterial quorum sensing signaling molecules, N-acylhomoserine lactones from Gram- pathogens in physiological samples.	AHLs	β -Galactosidase reporter: visual monitoring of a colorimetric signal	Paper strip	Gram- bacterial infectious diseases	Lab	173
Luciferase-based indicators of drugs (LUCIDs) ●●●	Semisynthetic bioluminescent protein sensors approach proposed as an entirely new mechanism for inexpensive point-of-care biosensors. That permits quantification of specific drugs in patient's samples by spotting minimal volumes on paper and recording the signal using a simple point-and-shoot camera.	MTX, Tacrolimus, Sirolimus, cyclosporin, topiramate, digoxin	Luminescence signal recorded by a digital camera	Paper strip	Companion diagnostics	Lab	104
Paper-Based Synthetic Gene Networks ●●●	Toehold RNA switches biosensors, in vitro paper-based platform that provides an alternate, versatile venue for synthetic biologists to operate and a much-needed medium for the safe deployment of engineered gene circuits beyond the lab. Commercially available cell-free systems are freeze dried onto paper, enabling the inexpensive, sterile, and abiotic distribution of synthetic-biology-based technologies for the clinic, global health, industry, research, and education.	diverse small molecules analytes (glucose), nucleic acids (mRNA)	Colorimetric human readable signal	Paper strip	Wide range of pathologies, proof-of-concept for Ebola virus diagnosis	Lab	103
Conditionally fluorescent dsDNA probe ●	Double-stranded toehold exchange: novel programmable mechanism in which each single nucleotide polymorphism generates two thermodynamically destabilizing mismatch bubbles rather than the single mismatch formed during typical hybridization-based assays. Up to a 12,000-fold excess of a target that contains a single nucleotide polymorphism is required to generate the same fluorescence as one equivalent of the intended target, and detection works reliably over a wide range of conditions.	Small variations in nucleic acid sequences and point mutations	fluorescence		SNP, proof of concept with bacterial antibiotic resistance genes	Lab	277
Logic gates that respond to the presence of both protein and DNA in a sample ●	Microarray sensor technology with logic capability for screening combinations of proteins and DNA in a biological sample.	combinations of proteins and DNA	fluorescence		Chronic obstructive pulmonary disease (COPD)	Lab	422
Bacteriophage-based microbial diagnostics ●●●	Engineering bacteriophages as near-real time microbial diagnostics by using them to transform target specific viable bacteria into factories for detectable molecules	bacterial pathogens	fluorescence, luminescence, colorimetric signals, phage/protein amplification	Cultivation in complex clinical sample	Detection of B. anthracis, Y. pestis, M. tuberculosis, S. aureus, L monocytogenes, Salmonella, E coli, and antibiotic susceptibility	Clinics	208 230 99 209
Boolean gated antibodies for logic detection ●●	Site-specific, chemical phosphorylation of a recognition domain creates boolean 'gated' antibodies. Binding is induced in an enzyme AND-antigen dependent manner. This 'AND-Ab' is active only in the presence of two biomarker inputs. Bivalent antibody-DNA conjugates as generic, noncovalent, and easily applicable molecular locks that allow the logic gated control of antibody activity using toehold-mediated strand displacement reactions.	Cell surface antigen and secreted enzyme, any epitope	Fluorescent/colorimetric output	Liquid phase	Immunoassays	Lab	307 423
Bacterial Quorum sensing biosensors for the clinics ●●	Bacterial biosensing systems to evaluate QSMs in physiological samples (stool, saliva) of patients	QS molecules	Bioluminescence/colorimetric	Paper based	Inflammatory bowel disease, Ulcerative colitis, Crohn's disease	Lab	173 545
Microbial biosensor for <i>in vitro</i> pretreatment assessment of Cytarabine efficacy in leukemia ●●	Microbial cell-based biosensor for the fast, <i>in vitro</i> prediction of leukemic cells response to the anticancer drug Ara-C (cytosine arabinoside)	Ara-C	Bioluminescence	Liquid phase	Leukemia	Lab	511
Bacterial biosensing system to monitor Methyl mercury poisoning ●	Bacterial biosensing system that can rapidly detect bioavailable MeHg	MeHg	Bioluminescence	Liquid phase	Methylmercury poisoning	Lab	701
Engineered virus nanoparticles based immunoassays ●●	The authors demonstrate that by combining viral nanoparticles, which are engineered to have dual affinity for troponin antibodies and nickel, with three-dimensional nanostructures they could detect troponin levels in human serum samples that are six to seven orders of magnitude lower than those detectable using conventional enzyme linked immunosorbent assays. The viral nanoparticle helps to orient the antibodies for maximum capture of biomarkers. High densities of antibodies on the surfaces of the nanoparticles lead to greater binding of the biomarkers, which enhances detection sensitivities.	Troponin I	Fluorescent, luminescent, electrochemical, enzymatic and colorimetric signals	Liquid phase	Acute myocardial infarction	Lab	244 245

Designation /Importance	Technology and approach	Input biomarkers	Output/Readout	Device format	Targeted pathology/indication	Clinics /Lab	Ref
Spore-based genetically engineered cell-based sensing systems ●●	Incorporated spore-based cell-based sensing systems into	Zinc and arsenite	fluorescence, luminescence	miniaturized microfluidic format (μTAS)	Measurements of seric zinc and arsenite levels	Lab	170
Nucleic Acid Circuits ●●●	Toehold mediated strand displacement mechanism alone have permitted to develop novel enzyme free nucleic acid amplification circuits for different diagnostic detection strategies, such as entropy-driven catalysis (EDC) circuits, seesaw gates, catalytic hairpin assembly (CHA) reactions and hybridization chain reactions (HCR)	Wide rand of analytes	fluorescent, luminescent, electrochemical, enzymatic and colorimetric signals	Liquid or solid phase		Lab	278
Logic-Based Autonomous cell surface profiling ●●	Cell types, both healthy and diseased, can be classified by inventories of their cell-surface markers using aptamers or antibodies. DNA nanorobots for programmable analysis of multiple surface markers to enable the clinical disease profile on whole cells. They engineered a device combining structure-switching DNA aptamers, or antibodies coupled with DNA devices with toehold-mediated strand displacement reactions to perform autonomous logic-based analysis of cell-surface markers.	Cell surface markers, Cluster of differentiation (CDs)	Fluorescence/ targeted therapeutics	Liquid phase	human cancer cell models	Lab	101 287
Biolecular logic gates that detect MDR bacteria ●●	Biochemical reaction networks exploiting enzymes and oligonucleotides with a computing functionality applied to the identification of bacteria exhibiting multi-drug resistance. This approach enables to identify the NDM-1-encoding gene (blaNDM-1) and concurrently to screen, by a tailor-designed biomolecular logical gate, two genetic fragments encoding the active sites bound to carbapenem.	Nucleic acids related to antibiotic resistance (NDM-1)	Electrochemical	Liquid phase/electrodes	MDR resistance of gram negative bacteria	Lab	424
Antibody diagnostics via evolution of peptides ●	Antibody diagnostics via evolution of peptides (ADEPt) to evolve diagnostically efficient peptides for de novo discovery and detection of antibody biomarkers without knowledge of disease pathophysiology. As pathological antibodies repertoire are known to change in diverse diseases, this methods has proven useful to create diagnostics for early disease detection, stratification, and therapeutic monitoring, and enabled effective identification of a critical environmental agent involved in celiac disease. Bacterial cell-displayed peptide libraries were quantitatively screened for binders to serum antibodies from patients with celiac disease.	Disease associated antibodies	Fluorescence	Liquid Phase	Celiac disease, theoretically many diseases	Lab	312
Synthetic genetic polymers XNA aptamers ●●	Novel synthetic nucleobases and their genetic polymers, known as XNA (xenonucleic acids) increase the chemical and structural diversity of nucleic acids, and open up the way for increased affinity and stability against enzymatic cleavage, expanded functionality such as enzymatic activity, and improved synthesis and selection procedures	PDGF, HIV RNA, Thrombin, Camptothecin, VEGF, Glucagon, IL-6, Cancerous cells	Various	Liquid phase	Various diseases	Lab	299
Prosthetic circuit to monitor and treat diet induced obesity ●●●	Mice transplanted with engineered cells bearing synthetic genetic circuit that constantly monitors blood fatty acid levels in the setting of diet-associated hyperlipidemia and coordinates reversible and adjustable expression of the clinically licensed appetite-suppressing peptide hormone.	fatty acid levels in blood	Appetite-suppressing peptide hormone Pramlintide	Microcapsule	Hyperlipidemia/Diet induced obesity	Lab	196
Biomolecular computer for diagnosis and therapy ●●●	Biomolecular computer performs in vitro the identification of a combination of cancer mRNA marker molecules at specific levels and generates a therapeutically active molecule	mRNAs	Therapeutic nucleic acid		models of small-cell lung cancer and prostate cancer	Lab	198
Bile acid-controlled prosthetic circuit ●●	Biosensor based on orthogonal synthetic gene switches that combine's bile acid-specific sensor capacity with dose-dependent expression of a specific transgene in mammalian cells and in mice.	pathological metabolites (Bile acids)	Therapeutic responses		Metabolic disorders	Lab	702
RNA control devices monitor signaling pathways and reprogram cellular fate ●	Protein-responsive RNAbased regulatory device integrating RNA aptamers that bind to disease associated protein ligands in key intronic locations of an alternatively spliced transcript linking intracellular protein concentrations to gene-expression events, and triggering apoptosis	Wnt and NF-κB pathway	Targeted apoptosis	intracellular RNA device	Cancer	Lab	189
Multi-input cancer cell classifier ●●●	Scalable synthetic genetic circuit works as a cell type classifier <i>in cellulo</i> by detecting customizable sets of endogenous pathological miRNAs and triggers apoptosis in HeLa cells	cancer specific endogenous miRNAs patterns	Apoptosis of cancer cells	intracellular genetic circuits	Cancer	Lab	102
Genetically Programmable platform to detect pathogens and trigger destruction ●	Proof-of-principle towards detection of <i>Pseudomonas aeruginosa</i> using quorum sensing signals and in situ destruction by an engineering E. coli secreting an engineered specific bacteriocin.	<i>P. aeruginosa</i> QS molecules (3OC12HSL)	Secretion of CoPy bacteriocin	in situ	Urinary tract and nosocomial infections	Lab	70 3
E. coli engineered into living diagnostics to probe the mammalian gut. ●●●	Engineered E coli that survive in mice gut gut and sense, remember, and report molecular signals thanks to a genetic circuits with a "trigger element" in which the lambda Cro gene is transcribed from a tetracycline-inducible promoter and a "memory element" derived from the cI/Cro region of phage lambda.	aTc	β-galactosidase reporter	Orally administered engineered bacterium, probiotic?	Proof-of-concept	Lab (mice)	41

Designation /Importance	Technology and approach	Input biomarkers	Output/Readout	Device format	Targeted pathology/indication	Clinics /Lab	Ref
Synthetic uric acid-responsive mammalian sensor circuit ●●●	Synthetic mammalian circuit to maintain uric acid homeostasis in the bloodstream. Modified <i>Deinococcus radiodurans</i> -derived protein that senses uric acids levels and triggers dose-dependent derepression of a secretion-engineered <i>Aspergillus flavus</i> urate oxidase that eliminates uric acid <i>in vivo</i> in mice	Uric acid	urate oxidase enzyme	Intraperitoneous implantation of microcapsules containing engineered cells	Proof-of-concept	Lab (mice)	194
Multifunctional Mammalian pH Sensor ●●●	The authors rewired the human proton-activated cell-surface receptor TDAG8 to chimeric promoters, creating a synthetic signaling cascade that monitors extracellular pH within the physiological range. The synthetic pH sensor was linked to production of insulin and implanted into type 1 diabetic mice developing diabetic ketoacidosis, creating a prosthetic network capable of automatically scoring acidic pH and coordinating an insulin expression response that corrected ketoacidosis.	pH, CO2	Fluorescence/ Insulin	Intraperitoneous implantation of microcapsules containing engineered cells	Proof-of-concept	Lab (mice)	185
Synthetic gene networks that detect bladder cancer cells ●●●	Synthetic gene network build using CRISPR-Cas9 technology in mammalian cells, that integrate cellular pathophysiological information from two cancer specific promoters as inputs and activate an output gene following a AND Boolean operation. When using a luciferase output, the authors could detect bladder cancer cells. The authors could also induce cell death using functional genes as outputs.	cancer specific intracellular transcriptional signals (human telomerase reverse transcriptase, human uroplakin II)	Luminescence, apoptosis	Intracellular gene circuits	Proof-of-concept	Lab	199
Protein switches that detect cancer and treats	The authors propose a strategy for designing protein therapeutics that link activation of a chosen therapeutic function to a specific cancer marker of choice. We demonstrate this strategy by creating a protein switch that renders cells susceptible to the in response to the cancer marker.	hypoxia-inducible factor 1 α (HIF-1 α)	Activation of the prodrug 5-fluorocytosine (5FC)	Intracellular protein switch	Human colon and breast cancer	Lab	186

Nature offers nothing that can be called this man's rather than another's; but under nature everything belongs to all.

Baruch Spinoza, *Tractatus Theologico-Politicus*, 1677

Chapter 8

Annexes

1. Plasmid maps

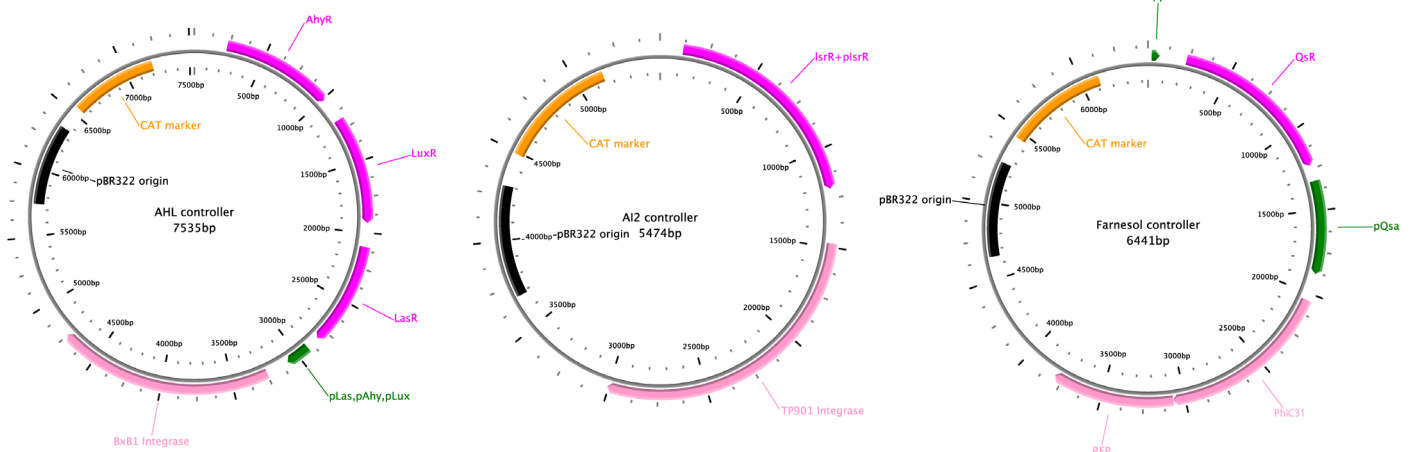


Figure S1: Prototype plasmid constructions for bactosensors mediated sepsis detection. From left to right: AHL-controller-BxB1 (AhyR, LuxR, LasR), AI2 controller-TP901 (lsrR) and Farnesol controller-Cre-RFP (QsR).





Figure S2: Plasmid used in biosensor development. (a) pYeaR-GFP measurement plasmid (GenBank accession number: KM234313). (b) pCpxP-GFP measurement plasmid (KM234314) (c) pCpxP-BxB1 controller (KM234315) (d) pYeaR-TP901 Controller 3 (KM234316) (e) pYeaR-BxB1 Controller (KM234317) (f) pYeaR-BxB1 Controller 2 (KM234318) (g) pYeaR-BxB1 controller 1 fusion with RFP (KM234319) (h) (i) (j) (k): AND, NAND, NOR and XOR BIL gates where the GFP output was replaced by RFP (mKate2) (KM234321, KM234322, KM234323) (l) pYeaR-TP901/pCpxP-BxB1 dual controller plasmid (KM347896)

2. Microfluidic chips

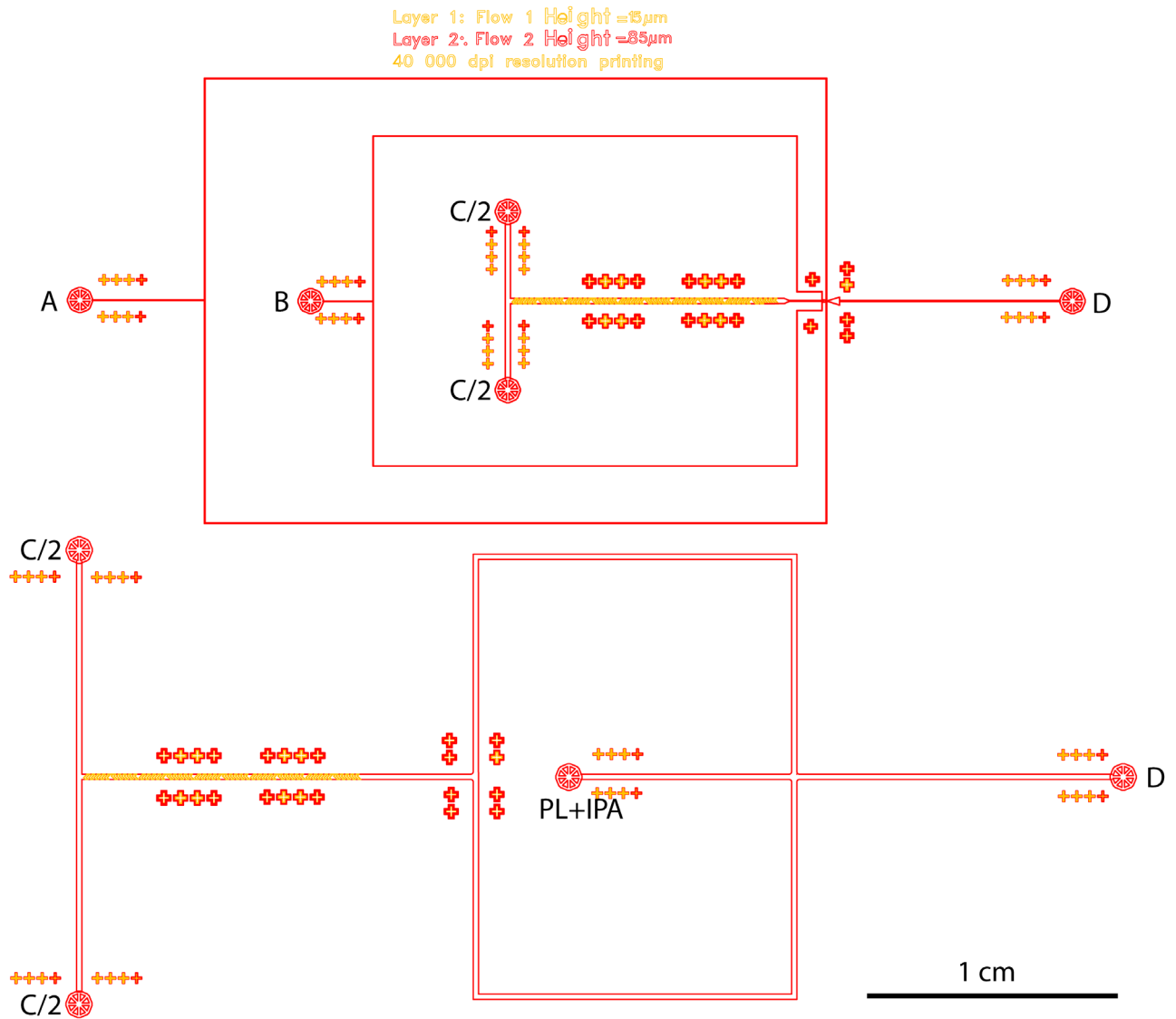


Figure S3: Microfluidic chips used in this study to generate protosensors. (A) Buffer (10% v/v methanol, 15% w/v glycerol, 3% w/v pluronic F68 in PBS), 1 μ l/min (B) DPPC dissolved in oleic acid, 0.4 μ l/min (C) Enzymes in PBS, 0.4 μ l/min (D) Out. **Top:** Double emulsion templating device, all experiments in this study were performed using this device. **Bottom:** Hydrodynamic flow focusing device. In this device, phospholipids are dissolved in an Isopropyl Alcohol buffer.

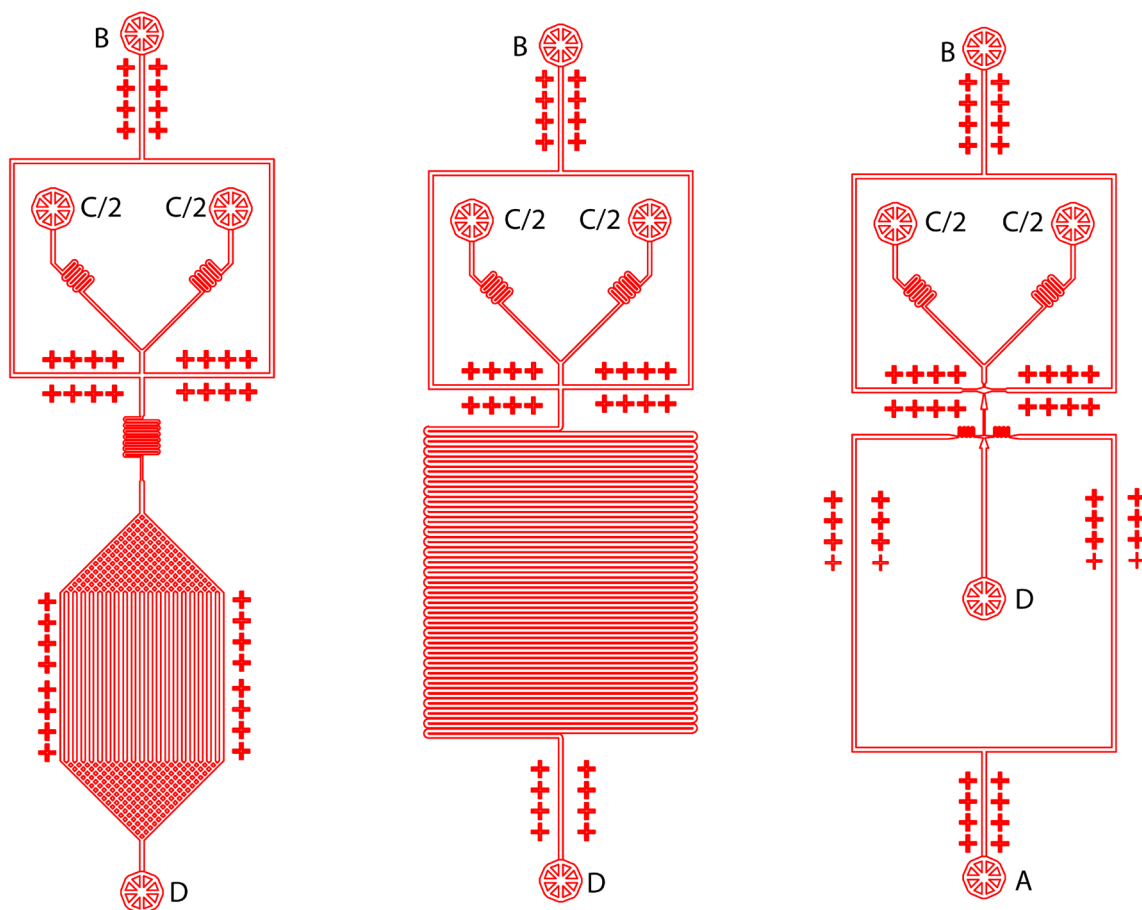


Figure S4: Miscellaneous microfluidic chips used in this study. (A) Buffer (10% v/v methanol, 15% w/v glycerol, 3% w/v pluronic F68 in PBS), 1 $\mu\text{l}/\text{min}$ (B) DPPC dissolved in oleic acid, 0.4 $\mu\text{l}/\text{min}$ (C) Enzymes in PBS, 0.4 $\mu\text{l}/\text{min}$ (D) Out. **Left:** Simple emulsion templating device. Used to generate and visualize the stability of simple emulsion in a built-in *tank*. This chip can also be used for two step vesicle preparation, where the output of the chip is transferred in a solution containing buffer A. **Middle:** Same as previously, but the architecture of the *tank* differs. **Right:** Alternative two-step *on-chip* to produce vesicles. This device works as expected to produce protocells but its flow rates modes are more difficult to operate than the main chip we used in this study.

3. HSIM et BIOCHAM code for models used in this work

3.1. GluONE batch mode

HSIM model code:

```
title = "GluONE_Batch";
geometry = 1000:1000;

metabolite
    'glucose',
    'gluconolacrone',
    'G_1DH',
    'acetone',
    'isopropanol',
    'ADH',
    'NADH',
    'NAD',
    'resazurin',
    'resorufin',
    'HRP',
    'H_2O_2',
    'AO',
    'NADN',
    'HRP2';

display (
    'NADN');

init (1 mM, 'glucose');
init (1 mM, 'acetone');

//      R 5:

metabolite Cfa5 hide, Cia5 hide, Cib5 hide;
'HRP' + 'H_2O_2' -> Cia5 [0.814815];
Cia5 -> 'HRP' + 'H_2O_2' [0.0024];
Cia5 + 'resazurin' -> Cib5 [0.55];
Cib5 -> Cia5 + 'resazurin' [0.0001];
Cib5 -> 'HRP' + 'resorufin' [0.024];

//      R 4:

metabolite Cf4 hide;
'HRP2' + 'NADH' -> Cf4 [0.00020625];
Cf4 -> 'HRP2' + 'NADH' [9e-008];
Cf4 -> 'HRP2' + 'NADN' [9e-007];

//      R 3:

metabolite Cf3 hide;
'AO' + 'isopropanol' -> Cf3 [0.004125];
Cf3 -> 'AO' + 'isopropanol' [0.0015];
metabolite Cio3 hide;
Cf3 -> Cio3 + 'H_2O_2' [0.015];
Cio3 -> 'AO' + 'HRP2' [1];

//      R 2:

metabolite Cfa2 hide, Cia2 hide, Cib2 hide;
'ADH' + 'NADH' -> Cia2 [6.72222e-007];
Cia2 -> 'ADH' + 'NADH' [3.3e-006];
Cia2 + 'acetone' -> Cib2 [0.077];
Cib2 -> Cia2 + 'acetone' [7e-006];
Cib2 -> Cfa2 + 'NAD' [7e-005];
Cfa2 -> 'ADH' + 'isopropanol' [3.3e-005];

//      R 1:
```



```

metabolite Cfa1 hide, Cia1 hide, Cib1 hide;
'G_1DH' + 'NAD' -> Cia1 [0.127907];
Cia1 -> 'G_1DH' + 'NAD' [0.004];
Cia1 + 'glucose' -> Cib1 [0.0006875];
Cib1 -> Cia1 + 'glucose' [0.002];
Cib1 -> Cfa1 + 'NADH' [0.02];
Cfa1 -> 'G_1DH' + 'gluconolacrone' [0.04];

```

```

init (0.354uM, 'G_1DH');
init (14.6uM, 'ADH');
init (250uM, 'NAD');
init (50 uM, 'resazurin');
init (0.001041uM, 'HRP');
init (0.02775uM, 'AO');
init (0.001041uM, 'HRP2');

```

Biocham model code:

```

parameter (k1, 1.15157e-005).
parameter (k2, 24).
parameter (k3, 7.77313e-006).
parameter (k4, 1).
parameter (k5, 240).
parameter (k6, 2.91492e-009).
parameter (k7, 0.0009).
parameter (k8, 0.009).
parameter (k9, 5.82985e-008).
parameter (k10, 15).
parameter (k11, 150).
parameter (k12, 10000).
parameter (k13, 9.50049e-012).
parameter (k14, 0.033).
parameter (k15, 1.08824e-006).
parameter (k16, 0.07).
parameter (k17, 0.7).
parameter (k18, 0.33).
parameter (k19, 1.8077e-006).
parameter (k20, 40).
parameter (k21, 9.71641e-009).
parameter (k22, 20).
parameter (k23, 200).
parameter (k24, 400).

```

```

MA(k1) for HRP + H_2O_2 => Cia5.
MA(k2) for Cia5 => HRP + H_2O_2.
MA(k3) for Cia5 + resazurin => Cib5.
MA(k4) for Cib5 => Cia5 + resazurin.
MA(k5) for Cib5 => HRP + resorufin.
MA(k6) for HRP2 + NADH => Cf4.
MA(k7) for Cf4 => HRP2 + NADH.
MA(k8) for Cf4 => HRP2 + NADN.
MA(k9) for AO + isopropanol => Cf3.
MA(k10) for Cf3 => AO + isopropanol.
MA(k11) for Cf3 => Cio3 + H_2O_2.
MA(k12) for Cio3 => AO + HRP2.
MA(k13) for ADH + NADH => Cia2.
MA(k14) for Cia2 => ADH + NADH.
MA(k15) for Cia2 + acetone => Cib2.
MA(k16) for Cib2 => Cia2 + acetone.
MA(k17) for Cib2 => Cfa2 + NAD.
MA(k18) for Cfa2 => ADH + isopropanol.
MA(k19) for G_1DH + NAD => Cia1.
MA(k20) for Cia1 => G_1DH + NAD.
MA(k21) for Cia1 + glucose => Cib1.
MA(k22) for Cib1 => Cia1 + glucose.
MA(k23) for Cib1 => Cfa1 + NADH.
MA(k24) for Cfa1 => G_1DH + gluconolacrone.

```

```

present(glucose, d).
present(acetone, e).

```

```

present(gluconolacrone, 0).
present(G_1DH, c).
present(isopropanol, 0).
present(ADH, a).
present(NADH, 0).
present(NAD, f).
present(resazurin, 315338394).
present(resorufin, 0).
present(HRP, 6565).
present(H_2O_2, 0).
present(AO, b).
present(NADN, 0).
present(HRP2, 6565).
present(Cfa5, 0).
present(Cia5, 0).
present(Cib5, 0).
present(Cf4, 0).
present(Cf3, 0).
present(Cio3, 0).
present(Cfa2, 0).
present(Cia2, 0).
present(Cib2, 0).
present(Cfa1, 0).
present(Cia1, 0).

parameter(b, 175012).
parameter(c, 2232595).
parameter(a, 92078816).
parameter(f, 1576691971).
parameter(d, 438319820).
parameter(e, 5360746).

hide_molecules(?).
show_molecules({NADN, acetone, glucose, resorufin, NADH}).

```

3.2 GluONE protosensor mode

HSIM model code:

```

title = "GluONE_protosensor";
geometry = 1500:1500;

metabolite
    'glucose',
    'gluconolacrone',
    'G_1DH',
    'acetone',
    'isopropanol',
    'ADH',
    'NADH',
    'NAD',
    'resazurin',
    'resorufin',
    'HRP',
    'H_2O_2',
    'AO',
    'NADN',
    'HRP2';

compartment {
    geometry = 1000:1000+0+0+0; // 10000 nm de long, 10000 nm de diametre
    init (7.08uM, 'G_1DH');
    init (292uM, 'ADH');
    init (5000uM, 'NAD');
    init (1000 uM, 'resazurin');
    init (0.02082uM, 'HRP');
    init (0.555uM, 'AO');
    init (0.02082uM, 'HRP2');
}

```

```

display (
    'NADN');

init (1 mM, 'glucose');
init (1 mM, 'acetone');

diffusion (glucose) = 5e-3; // Arrayed lipid bilayer chambers allow single-molecule analysis of membrane transporter activity
diffusion (acetone) = 1e-2; // Permeable ~ethanol/H2O, Molecular Biology of the Cell. 4th edition.

//      R 5:

metabolite Cfa5 hide, Cia5 hide, Cib5 hide;
'HRP' + 'H_2O_2' -> Cia5 [0.814815];
Cia5 -> 'HRP' + 'H_2O_2' [0.0024];
Cia5 + 'resazurin' -> Cib5 [0.55];
Cib5 -> Cia5 + 'resazurin' [0.0001];
Cib5 -> 'HRP' + 'resorufin' [0.024];

//      R 4:

metabolite Cf4 hide;
'HRP2' + 'NADH' -> Cf4 [0.00020625];
Cf4 -> 'HRP2' + 'NADH' [9e-008];
Cf4 -> 'HRP2' + 'NADN' [9e-007];

//      R 3:

metabolite Cf3 hide;
'AO' + 'isopropanol' -> Cf3 [0.004125];
Cf3 -> 'AO' + 'isopropanol' [0.0015];
metabolite Cio3 hide;
Cf3 -> Cio3 + 'H_2O_2' [0.015];
Cio3 -> 'AO' + 'HRP2' [1];

//      R 2:

metabolite Cfa2 hide, Cia2 hide, Cib2 hide;
'ADH' + 'NADH' -> Cia2 [6.72222e-007];
Cia2 -> 'ADH' + 'NADH' [3.3e-006];
Cia2 + 'acetone' -> Cib2 [0.077];
Cib2 -> Cia2 + 'acetone' [7e-006];
Cib2 -> Cfa2 + 'NAD' [7e-005];
Cfa2 -> 'ADH' + 'isopropanol' [3.3e-005];

//      R 1:

metabolite Cfa1 hide, Cia1 hide, Cib1 hide;
'G_1DH' + 'NAD' -> Cia1 [0.127907];
Cia1 -> 'G_1DH' + 'NAD' [0.004];
Cia1 + 'glucose' -> Cib1 [0.0006875];
Cib1 -> Cia1 + 'glucose' [0.002];
Cib1 -> Cfa1 + 'NADH' [0.02];
Cfa1 -> 'G_1DH' + 'gluconolacrone' [0.04];

```

Biocham model code:

```

parameter (k1, 1.15157e-005).
parameter (k2, 24).
parameter (k3, 7.77313e-006).
parameter (k4, 1).
parameter (k5, 240).
parameter (k6, 2.91492e-009).
parameter (k7, 0.0009).
parameter (k8, 0.009).
parameter (k9, 5.82985e-008).
parameter (k10, 15).
parameter (k11, 150).
parameter (k12, 10000).
parameter (k13, 9.50049e-012).
parameter (k14, 0.033).
parameter (k15, 1.08824e-006).

```

parameter (k16, 0.07).
parameter (k17, 0.7).
parameter (k18, 0.33).
parameter (k19, 1.8077e-006).
parameter (k20, 40).
parameter (k21, 9.71641e-009).
parameter (k22, 20).
parameter (k23, 200).
parameter (k24, 400).
parameter (k25, 5e-3).
parameter (k26, 1e-2).

MA(k25) for glucoseext => glucose.
MA(k26) for acetoneext => acetone.

MA(k1) for HRP + H_2O_2 => Cia5.
MA(k2) for Cia5 => HRP + H_2O_2.
MA(k3) for Cia5 + resazurin => Cib5.
MA(k4) for Cib5 => Cia5 + resazurin.
MA(k5) for Cib5 => HRP + resorufin.
MA(k6) for HRP2 + NADH => Cf4.
MA(k7) for Cf4 => HRP2 + NADH.
MA(k8) for Cf4 => HRP2 + NADN.
MA(k9) for AO + isopropanol => Cf3.
MA(k10) for Cf3 => AO + isopropanol.
MA(k11) for Cf3 => Cio3 + H_2O_2.
MA(k12) for Cio3 => AO + HRP2.
MA(k13) for ADH + NADH => Cia2.
MA(k14) for Cia2 => ADH + NADH.
MA(k15) for Cia2 + acetone => Cib2.
MA(k16) for Cib2 => Cia2 + acetone.
MA(k17) for Cib2 => Cfa2 + NAD.
MA(k18) for Cfa2 => ADH + isopropanol.
MA(k19) for G_1DH + NAD => Cia1.
MA(k20) for Cia1 => G_1DH + NAD.
MA(k21) for Cia1 + glucose => Cib1.
MA(k22) for Cib1 => Cia1 + glucose.
MA(k23) for Cib1 => Cfa1 + NADH.
MA(k24) for Cfa1 => G_1DH + gluconolacrone.

present(glucoseext, d).
present(acetoneext, e).
present(glucose, 0).
present(gluconolacrone, 0).
present(G_1DH, c).
present(acetone, 0).
present(isopropanol, 0).
present(ADH, a).
present(NADH, 0).
present(NAD, f).
present(resazurin, 315338394).
present(resorufin, 0).
present(HRP, 6565).
present(H_2O_2, 0).
present(AO, b).
present(NADN, 0).
present(HRP2, 6565).
present(Cfa5, 0).
present(Cia5, 0).
present(Cib5, 0).
present(Cf4, 0).
present(Cf3, 0).
present(Cio3, 0).
present(Cfa2, 0).
present(Cia2, 0).
present(Cib2, 0).
present(Cfa1, 0).
present(Cia1, 0).

parameter(b, 175012).
parameter(c, 2232595).


```

//      R 4:

metabolite Cia4 hide;
'ABTSox' + 'NADH' -> Cia4 [1e-004];
Cia4 -> 'ABTS' + 'NAD' [0.01];

//      R 3:

metabolite Cfa3 hide, Cfb3 hide, Cia3 hide, Cib3 hide;
'POD' + 'H2O2' -> Cia3 [41.8];
Cia3 -> 'POD' + 'H2O2' [0.0076];
'POD' + 'ABTS' -> Cib3 [1.16111];
Cib3 -> 'POD' + 'ABTS' [0.0076];
Cia3 + 'ABTS' -> Cfa3 [1.16111];
Cib3 + 'H2O2' -> Cfb3 [41.8];
Cfa3 -> Cia3 + 'ABTS' [0.0076];
Cfb3 -> Cib3 + 'H2O2' [0.0076];
Cfa3 -> 'POD' + 'ABTSox' [0.076];
Cfb3 -> 'POD' + 'ABTSox' [0.076];

//      R 2:

metabolite Cf2 hide;
'LO' + 'Lactate' -> Cf2 [0.327662];
Cf2 -> 'LO' + 'Lactate' [0.002383];
metabolite Cio2 hide;
Cf2 -> Cio2 + 'H2O2' [0.02383];
Cio2 -> 'LO' + 'Pyruvate' [1];

//      R 1:

metabolite Cfa1 hide, Cfb1 hide, Cia1 hide, Cib1 hide;
'ADH' + 'EtOH' -> Cia1 [0.0344309];
Cia1 -> 'ADH' + 'EtOH' [0.00308];
'ADH' + 'NAD' -> Cib1 [8.47];
Cib1 -> 'ADH' + 'NAD' [0.00308];
Cia1 + 'NAD' -> Cfa1 [8.47];
Cib1 + 'EtOH' -> Cfb1 [0.0344309];
Cfa1 -> Cia1 + 'NAD' [0.00308];
Cfb1 -> Cib1 + 'EtOH' [0.00308];
metabolite Cio1 hide;
Cfa1 -> Cio1 + 'acetaldehyde' [0.0308];
Cfb1 -> Cio1 + 'acetaldehyde' [0.0308];
Cio1 -> 'ADH' + 'NADH' [1];

init (20 mM, 'EtOH');
init (500 uM, 'Lactate');
init (1.12uM, 'LO');
init (100uM, 'ABTS');
init (0.00347uM, 'POD');
init (250 uM, 'NAD');
init (14.06 uM, 'ADH');

```

Biocham model code:

```

parameter (k1, 3.88656e-006).
parameter (k2, 10).
parameter (k3, 100).
parameter (k4, 4.63083e-006).
parameter (k5, 23.83).
parameter (k6, 238.3).
parameter (k7, 10000).
parameter (k8, 4.8661e-007).
parameter (k9, 30.8).
parameter (k10, 0.000119706).
parameter (k11, 30.8).
parameter (k12, 0.000119706).
parameter (k13, 4.8661e-007).
parameter (k14, 30.8).
parameter (k15, 30.8).
parameter (k16, 308).

```

parameter (k17, 308).
parameter (k18, 10000).
parameter (k19, 0.000590758).
parameter (k20, 76).
parameter (k21, 1.64099e-005).
parameter (k22, 76).
parameter (k23, 760).

MA(k1) for NADH + ABTSOX => Cf3.
MA(k2) for Cf3 => NADH + ABTSOX.
MA(k3) for Cf3 => NAD + ABTS.
MA(k4) for LO + Lactate => Cf2.
MA(k5) for Cf2 => LO + Lactate.
MA(k6) for Cf2 => Cio2 + H2O2.
MA(k7) for Cio2 => LO + Pyruvate.
MA(k8) for ADH + EtOH => Cia1.
MA(k9) for Cia1 => ADH + EtOH.
MA(k10) for ADH + NAD => Cib1.
MA(k11) for Cib1 => ADH + NAD.
MA(k12) for Cia1 + NAD => Cfa1.
MA(k13) for Cib1 + EtOH => Cfb1.
MA(k14) for Cfa1 => Cia1 + NAD.
MA(k15) for Cfb1 => Cib1 + EtOH.
MA(k16) for Cfa1 => Cio1 + acetaldehyde.
MA(k17) for Cfb1 => Cio1 + acetaldehyde.
MA(k18) for Cio1 => ADH + NADH.
MA(k19) for POD + ABTS => Cia5.
MA(k20) for Cia5 => POD + ABTS.
MA(k21) for Cia5 + H2O2 => Cib5.
MA(k22) for Cib5 => Cia5 + H2O2.
MA(k23) for Cib5 => POD + ABTSOX.

present(EtOH, e).
present(acetaldehyde, 0).
present(NADH, 0).
present(Lactate, d).
present(H2O2, 0).
present(LO, a).
present(Pyruvate, 0).
present(ABTS, 630676000).
present(ABTSOX, 0).
present(POD, b).
present(NAD, f).
present(ADH, c).
present(Cf3, 0).
present(Cf2, 0).
present(Cio2, 0).
present(Cfa1, 0).
present(Cfb1, 0).
present(Cia1, 0).
present(Cib1, 0).
present(Cio1, 0).
present(Cia5, 0).
present(Cib5, 0).

parameter(b, 21884).
parameter(c, 92078696).
parameter(a, 7063571).
parameter(f, 1576690000).
parameter(d, 3153380).
parameter(e, 5486881).

hide_molecules(?).
show_molecules({NADH, NAD, H2O2, ABTS, ABTSOX}).

3.4. LacOH protosensor mode

HSIM model code:

```
title = "LacOH_protosensor";
geometry = 1500:1500;
```

```
metabolite
```

```
'EtOH',
'acetaldehyde',
'NADH',
'Lactate',
'H2O2',
'LO',
'Pyruvate',
'ABTS',
'ABTSox',
'POD',
'NAD',
'ADH';
```

```
compartment {
```

```
  geometry = 1000:1000+0+0+0; // 10000 nm de long, 10000 nm de diametre
  init (22.4uM, 'LO');
  init (5000uM, 'NAD');
  init (2000 uM, 'ABTS');
  init (0.0694uM, 'POD');
  init (281.2uM, 'ADH');
}
```

```
diffusion (Lactate) = 5e-3; // Arrayed lipid bilayer chambers allow single-molecule analysis of membrane transporter activity
diffusion (EtOH) = 1e-2; // Permeable ~ethanol/H2O, Molecular Biology of the Cell. 4th edition.
```

```
// R 4:
```

```
metabolite Cia4 hide;
```

```
'ABTSox' + 'NADH' -> Cia4 [1e-004];
```

```
Cia4 -> 'ABTS' + 'NAD' [0.01];
```

```
// R 3:
```

```
metabolite Cfa3 hide, Cfb3 hide, Cia3 hide, Cib3 hide;
```

```
'POD' + 'H2O2' -> Cia3 [41.8];
```

```
Cia3 -> 'POD' + 'H2O2' [0.0076];
```

```
'POD' + 'ABTS' -> Cib3 [1.16111];
```

```
Cib3 -> 'POD' + 'ABTS' [0.0076];
```

```
Cia3 + 'ABTS' -> Cfa3 [1.16111];
```

```
Cib3 + 'H2O2' -> Cfb3 [41.8];
```

```
Cfa3 -> Cia3 + 'ABTS' [0.0076];
```

```
Cfb3 -> Cib3 + 'H2O2' [0.0076];
```

```
Cfa3 -> 'POD' + 'ABTSox' [0.076];
```

```
Cfb3 -> 'POD' + 'ABTSox' [0.076];
```

```
// R 2:
```

```
metabolite Cf2 hide;
```

```
'LO' + 'Lactate' -> Cf2 [0.327662];
```

```
Cf2 -> 'LO' + 'Lactate' [0.002383];
```

```
metabolite Cio2 hide;
```

```
Cf2 -> Cio2 + 'H2O2' [0.02383];
```

```
Cio2 -> 'LO' + 'Pyruvate' [1];
```

```
// R 1:
```

```
metabolite Cfa1 hide, Cfb1 hide, Cia1 hide, Cib1 hide;
```

```
'ADH' + 'EtOH' -> Cia1 [0.0344309];
```

```
Cia1 -> 'ADH' + 'EtOH' [0.00308];
```

```
'ADH' + 'NAD' -> Cib1 [8.47];
```

```
Cib1 -> 'ADH' + 'NAD' [0.00308];
```



```

Cia1 + 'NAD' -> Cfa1 [8.47];
Cib1 + 'EtOH' -> Cfb1 [0.0344309];
Cfa1 -> Cia1 + 'NAD' [0.00308];
Cfb1 -> Cib1 + 'EtOH' [0.00308];
metabolite Cio1 hide;
Cfa1 -> Cio1 + 'acetaldehyde' [0.0308];
Cfb1 -> Cio1 + 'acetaldehyde' [0.0308];
Cio1 -> 'ADH' + 'NADH' [1];

```

```

init (0 mM, 'EtOH');
init (500 uM, 'Lactate');

```

Biocham model code:

```

parameter (k1, 3.88656e-006).
parameter (k2, 10).
parameter (k3, 100).
parameter (k4, 4.63083e-006).
parameter (k5, 23.83).
parameter (k6, 238.3).
parameter (k7, 10000).
parameter (k8, 4.8661e-007).
parameter (k9, 30.8).
parameter (k10, 0.000119706).
parameter (k11, 30.8).
parameter (k12, 0.000119706).
parameter (k13, 4.8661e-007).
parameter (k14, 30.8).
parameter (k15, 30.8).
parameter (k16, 308).
parameter (k17, 308).
parameter (k18, 10000).
parameter (k19, 0.000590758).
parameter (k20, 76).
parameter (k21, 1.64099e-005).
parameter (k22, 76).
parameter (k23, 760).
parameter (k24, 5e-3).
parameter (k25, 1e-2).

```

```

MA(k24) for Lactateext => Lactate.
MA(k25) for EtOHext => EtOH.

```

```

MA(k1) for NADH + ABTSOX => Cf3.
MA(k2) for Cf3 => NADH + ABTSOX.
MA(k3) for Cf3 => NAD + ABTS.
MA(k4) for LO + Lactate => Cf2.
MA(k5) for Cf2 => LO + Lactate.
MA(k6) for Cf2 => Cio2 + H2O2.
MA(k7) for Cio2 => LO + Pyruvate.
MA(k8) for ADH + EtOH => Cia1.
MA(k9) for Cia1 => ADH + EtOH.
MA(k10) for ADH + NAD => Cib1.
MA(k11) for Cib1 => ADH + NAD.
MA(k12) for Cia1 + NAD => Cfa1.
MA(k13) for Cib1 + EtOH => Cfb1.
MA(k14) for Cfa1 => Cia1 + NAD.
MA(k15) for Cfb1 => Cib1 + EtOH.
MA(k16) for Cfa1 => Cio1 + acetaldehyde.
MA(k17) for Cfb1 => Cio1 + acetaldehyde.
MA(k18) for Cio1 => ADH + NADH.
MA(k19) for POD + ABTS => Cia5.
MA(k20) for Cia5 => POD + ABTS.
MA(k21) for Cia5 + H2O2 => Cib5.
MA(k22) for Cib5 => Cia5 + H2O2.
MA(k23) for Cib5 => POD + ABTSOX.

```

```

present(Lactateext, d).
present(EtOHext, e).
present(EtOH, 0).
present(acetaldehyde, 0).

```


3.5 GluNOx batch mode

HSIM model code:

```
title = "GluNOx_batch";
geometry = 1000:1000;

metabolite
    'glucose',
    'gluconolactone',
    'G_1DH',
    'NO3',
    'NO2',
    'NR',
    'NADH',
    'NAD',
    'N2O3',
    'DAFF',
    'NO2b',
    'O2',
    'NO',
    'DAF';

//      R 8: NO decay (Application of carbon fiber composite minielectrodes for measurement of kinetic constants of nitric oxide decay
in solution.)

metabolite Cf8 hide;
'NO' -> Cf8 [1.9e-007];

//      R 7:

metabolite Cf7 hide;
'NO2b' + 'NO' -> Cf7 [0.25e-005];
Cf7 -> 'NO2b' + 'NO' [6.25e-007];
Cf7 -> 'N2O3' [0.2];

//      R 6:

metabolite Cf6 hide;
'O2' + 'NO' -> Cf6 [0.25e-005];
Cf6 -> 'O2' + 'NO' [6.25e-007];
Cf6 -> 'NO2b' [0.0002];

//      R 5:

metabolite Cf5 hide;
'DAF' + 'N2O3' -> Cf5 [0.25e-005];
Cf5 -> 'DAF' + 'N2O3' [6.25e-007];
Cf5 -> 'DAFF' [0.2];

// metabolite Cf5 hide;
// 'DAF' + 'NO' -> Cf5 [0.25e-005];
// Cf5 -> 'DAF' + 'NO' [6.25e-007];
// Cf5 -> 'DAFF' [0.000628];

//      R 2:

metabolite Cfa2 hide, Cfb2 hide, Cia2 hide, Cib2 hide;
'NR' + 'NO3' -> Cia2 [0.385];
Cia2 -> 'NR' + 'NO3' [0.00021];
'NR' + 'NADH' -> Cib2 [0.144375];
Cib2 -> 'NR' + 'NADH' [0.000021];
Cia2 + 'NADH' -> Cfa2 [0.144375];
Cib2 + 'NO3' -> Cfb2 [0.385];
Cfa2 -> Cia2 + 'NADH' [0.000021];
Cfb2 -> Cib2 + 'NO3' [0.00021];
metabolite Cio2 hide;
Cfa2 -> Cio2 + 'NO2' [0.021];
Cfb2 -> Cio2 + 'NO2' [0.021];
Cio2 -> 'NR' + 'NAD' [1];
```

```
//      R 3:

metabolite Cfa3 hide, Cfb3 hide, Cib3 hide;
'NR' + 'NO2' -> Cib3 [0.0743243];
Cib3 -> 'NR' + 'NO2' [2e-005];
Cib3 + 'NADH' -> Cfa3 [0.1375];
Cib2 + 'NO2' -> Cfb3 [0.0743243];
Cfa3 -> Cib3 + 'NADH' [2e-005];
Cfb3 -> Cib2 + 'NO2' [2e-005];
metabolite Cio3 hide;
Cfa3 -> Cio3 + 'NO' [0.0002];
Cfb3 -> Cio3 + 'NO' [0.0002];
Cio3 -> 'NR' + 'NAD' [1];
```

```
//      R 1:

metabolite Cfa1 hide, Cia1 hide, Cib1 hide;
'G_1DH' + 'NAD' -> Cia1 [0.023913];
Cia1 -> 'G_1DH' + 'NAD' [0.004];
Cia1 + 'glucose' -> Cib1 [0.0006875];
Cib1 -> Cia1 + 'glucose' [0.002];
Cib1 -> Cfa1 + 'NADH' [0.02];
Cfa1 -> 'G_1DH' + 'gluconolactone' [0.04];
```

```
init (5000uM, 'glucose');
init (2uM, 'G_1DH');
init (5000uM, 'NO3');
init (20uM, 'NO2');
init (100uM, 'NAD');
init (0.253mM, 'O2');
init (10uM, 'DAF');
init (4.2uM, 'NR');
```

Biocham model code:

```
parameter (k1, 0.0019).
parameter (k2, 7.77313e-006).
parameter (k3, 0.2).
parameter (k4, 2).
parameter (k5, 0.00777313).
parameter (k6, 200).
parameter (k7, 2000).
parameter (k8, 0.00777313).
parameter (k9, 200).
parameter (k10, 2000).
parameter (k11, 1.05042e-006).
parameter (k12, 0.2).
parameter (k13, 1.94328e-006).
parameter (k14, 0.2).
parameter (k15, 1.94328e-006).
parameter (k16, 1.05042e-006).
parameter (k17, 0.2).
parameter (k18, 0.2).
parameter (k19, 2).
parameter (k20, 2).
parameter (k21, 10000).
```

```
MA(k1) for NO => volatNO.
MA(k2) for O2 + NO => Cf6.
MA(k3) for Cf6 => O2 + NO.
MA(k4) for Cf6 => O2 + NO2b.
MA(k5) for NO + NO2b => Cf5.
MA(k6) for Cf5 => NO + NO2b.
MA(k7) for Cf5 => N2O3.
MA(k8) for DAF + N2O3 => Cf4.
MA(k9) for Cf4 => DAF + N2O3.
MA(k10) for Cf4 => DAFF.
MA(k11) for NR + NO2 => Cia3.
```

MA(k12) for Cia3 => NR + NO2.
MA(k13) for NR + NADH => Cib3.
MA(k14) for Cib3 => NR + NADH.
MA(k15) for Cia3 + NADH => Cfa3.
MA(k16) for Cib3 + NO2 => Cfb3.
MA(k17) for Cfa3 => Cia3 + NADH.
MA(k18) for Cfb3 => Cib3 + NO2.
MA(k19) for Cfa3 => Cio3 + NO.
MA(k20) for Cfb3 => Cio3 + NO.
MA(k21) for Cio3 => NR + NAD.

parameter (k22, 5.44119e-005).
parameter (k23, 21).
parameter (k25, 0.000204045).
parameter (k26, 5.44119e-005).
parameter (k27, 21).
parameter (k28, 21).
parameter (k29, 210).
parameter (k30, 210).
parameter (k31, 10000).
parameter (k32, 3.37961e-007).
parameter (k33, 40).
parameter (k34, 9.71641e-009).
parameter (k35, 20).
parameter (k36, 200).
parameter (k37, 400).

MA(k22) for NR + NO3 => Cia2.
MA(k23) for Cia2 => NR + NO3.
MA(k25) for Cia2 + NADH => Cfa2.
MA(k26) for Cib3 + NO3 => Cfb2.
MA(k27) for Cfa2 => Cia2 + NADH.
MA(k28) for Cfb2 => Cib3 + NO3.
MA(k29) for Cfa2 => Cio2 + NO2.
MA(k30) for Cfb2 => Cio2 + NO2.
MA(k31) for Cio2 => NR + NAD.
MA(k32) for G_1DH + NAD => Cia1.
MA(k33) for Cia1 => G_1DH + NAD.
MA(k34) for Cia1 + glucose => Cib1.
MA(k35) for Cib1 => Cia1 + glucose.
MA(k36) for Cib1 => Cfa1 + NADH.
MA(k37) for Cfa1 => G_1DH + gluconolacrone.

present(glucose, d).
present(NO3, e).
present(NO2, f).
present(gluconolacrone, 0).
present(G_1DH, a).
present(NR, b).
present(NADH, 0).
present(NAD, g).
present(N2O3, 0).
present(DAFF, 0).
present(NO2b, 0).
present(O2, 79780608).
present(NO, 0).
present(DAF, 63067600).
present(volatNO, 0).
present(NOp, 0).
present(Cf6, 0).
present(Cf5, 0).
present(Cf4, 0).
present(Cfa3, 0).
present(Cfb3, 0).
present(Cia3, 0).
present(Cib3, 0).
present(Cio3, 0).
present(Cfa2, 0).
present(Cia2, 0).
present(Cfa2, 0).
present(Cfb2, 0).

```

present(Cia2, 0).
present(Cio2, 0).
present(Cfa1, 0).
present(Cia1, 0).

parameter(b, 26488392).
parameter(a, 36295404).
parameter(g, 630676000).
parameter(d, 438319820).
parameter(e, 315338000).
parameter(f, 1261356).

hide_molecules(?).
show_molecules({NADH, NAD, DAFF, NO}).

```

3.6. GluNOx protosensor mode

HSIM model code:

```

title = "GluNOx_protosensor";
geometry = 1500:1500;

metabolite
    'glucose',
    'gluconolactone',
    'G_1DH',
    'NO3',
    'NO2',
    'NR',
    'NADH',
    'NAD',
    'N2O3',
    'DAFF',
    'NO2b',
    'O2',
    'NO',
    'DAF';

compartment {
    geometry = 1000:1000+0+0+0; // 10000 nm de long, 10000 nm de diamtre
    init (115.092uM, 'G_1DH');
    init (2000uM, 'NAD');
    init (200uM, 'DAF');
    init (84uM, 'NR');
}

diffusion (glucose) = 5e-3; // Arrayed lipid bilayer chambers allow single-molecule analysis of membrane transporter activity
diffusion (NO3) = 5e-3; //
diffusion (O2) = 5e-3; //

//      R 8: NO decay (Application of carbon fiber composite minielectrodes for measurement of kinetic constants of nitric oxide decay
//      in solution.)

metabolite Cf8 hide;
'NO' -> Cf8 [1.9e-007];

//      R 7:

metabolite Cf7 hide;
'NO2b' + 'NO' -> Cf7 [0.25e-005];
Cf7 -> 'NO2b' + 'NO' [6.25e-007];
Cf7 -> 'N2O3' [0.2];

//      R 6:

metabolite Cf6 hide;
'O2' + 'NO' -> Cf6 [0.25e-005];
Cf6 -> 'O2' + 'NO' [6.25e-007];
Cf6 -> 'NO2b' [0.0002];

```

```

//      R 5:

metabolite Cf5 hide;
'DAF' + 'N2O3' -> Cf5 [0.25e-005];
Cf5 -> 'DAF' + 'N2O3' [6.25e-007];
Cf5 -> 'DAFF' [0.2];

// metabolite Cf5 hide;
// 'DAF' + 'NO' -> Cf5 [0.25e-005];
// Cf5 -> 'DAF' + 'NO' [6.25e-007];
// Cf5 -> 'DAFF' [0.000628];

//      R 2:

metabolite Cfa2 hide, Cfb2 hide, Cia2 hide, Cib2 hide;
'NR' + 'NO3' -> Cia2 [0.385];
Cia2 -> 'NR' + 'NO3' [0.00021];
'NR' + 'NADH' -> Cib2 [0.144375];
Cib2 -> 'NR' + 'NADH' [0.000021];
Cia2 + 'NADH' -> Cfa2 [0.144375];
Cib2 + 'NO3' -> Cfb2 [0.385];
Cfa2 -> Cia2 + 'NADH' [0.000021];
Cfb2 -> Cib2 + 'NO3' [0.00021];
metabolite Cio2 hide;
Cfa2 -> Cio2 + 'NO2' [0.021];
Cfb2 -> Cio2 + 'NO2' [0.021];
Cio2 -> 'NR' + 'NAD' [1];

//      R 3:

metabolite Cfa3 hide, Cfb3 hide, Cib3 hide;
'NR' + 'NO2' -> Cib3 [0.0743243];
Cib3 -> 'NR' + 'NO2' [2e-005];
Cib3 + 'NADH' -> Cfa3 [0.1375];
Cib2 + 'NO2' -> Cfb3 [0.0743243];
Cfa3 -> Cib3 + 'NADH' [2e-005];
Cfb3 -> Cib2 + 'NO2' [2e-005];
metabolite Cio3 hide;
Cfa3 -> Cio3 + 'NO' [0.0002];
Cfb3 -> Cio3 + 'NO' [0.0002];
Cio3 -> 'NR' + 'NAD' [1];

//      R 1:

metabolite Cfa1 hide, Cia1 hide, Cib1 hide;
'G_1DH' + 'NAD' -> Cia1 [0.023913];
Cia1 -> 'G_1DH' + 'NAD' [0.004];
Cia1 + 'glucose' -> Cib1 [0.0006875];
Cib1 -> Cia1 + 'glucose' [0.002];
Cib1 -> Cfa1 + 'NADH' [0.02];
Cfa1 -> 'G_1DH' + 'gluconolactone' [0.04];

init (5000uM, 'glucose');
init (5000uM, 'NO3');
init (0uM, 'NO2');
init (0.253mM, 'O2');

```

Biocham model code:

```

parameter (k1, 0.0019).
parameter (k2, 7.77313e-006).
parameter (k3, 0.2).
parameter (k4, 2).
parameter (k5, 0.00777313).
parameter (k6, 200).
parameter (k7, 2000).
parameter (k8, 0.00777313).
parameter (k9, 200).
parameter (k10, 2000).
parameter (k11, 1.05042e-006).
parameter (k12, 0.2).

```

parameter (k13, 1.94328e-006).
parameter (k14, 0.2).
parameter (k15, 1.94328e-006).
parameter (k16, 1.05042e-006).
parameter (k17, 0.2).
parameter (k18, 0.2).
parameter (k19, 2).
parameter (k20, 2).
parameter (k21, 10000).

MA(k1) for NO => volatNO.
MA(k2) for O2 + NO => Cf6.
MA(k3) for Cf6 => O2 + NO.
MA(k4) for Cf6 => O2 + NO2b.
MA(k5) for NO + NO2b => Cf5.
MA(k6) for Cf5 => NO + NO2b.
MA(k7) for Cf5 => N2O3.
MA(k8) for DAF + N2O3 => Cf4.
MA(k9) for Cf4 => DAF + N2O3.
MA(k10) for Cf4 => DAFF.
MA(k11) for NR + NO2 => Cia3.
MA(k12) for Cia3 => NR + NO2.
MA(k13) for NR + NADH => Cib3.
MA(k14) for Cib3 => NR + NADH.
MA(k15) for Cia3 + NADH => Cfa3.
MA(k16) for Cib3 + NO2 => Cfb3.
MA(k17) for Cfa3 => Cia3 + NADH.
MA(k18) for Cfb3 => Cib3 + NO2.
MA(k19) for Cfa3 => Cio3 + NO.
MA(k20) for Cfb3 => Cio3 + NO.
MA(k21) for Cio3 => NR + NAD.

parameter (k22, 5.44119e-005).
parameter (k23, 21).
parameter (k25, 0.000204045).
parameter (k26, 5.44119e-005).
parameter (k27, 21).
parameter (k28, 21).
parameter (k29, 210).
parameter (k30, 210).
parameter (k31, 10000).
parameter (k32, 3.37961e-007).
parameter (k33, 40).
parameter (k34, 9.71641e-009).
parameter (k35, 20).
parameter (k36, 200).
parameter (k37, 400).

MA(k22) for NR + NO3 => Cia2.
MA(k23) for Cia2 => NR + NO3.
MA(k25) for Cia2 + NADH => Cfa2.
MA(k26) for Cib3 + NO3 => Cfb2.
MA(k27) for Cfa2 => Cia2 + NADH.
MA(k28) for Cfb2 => Cib3 + NO3.
MA(k29) for Cfa2 => Cio2 + NO2.
MA(k30) for Cfb2 => Cio2 + NO2.
MA(k31) for Cio2 => NR + NAD.
MA(k32) for G_1DH + NAD => Cia1.
MA(k33) for Cia1 => G_1DH + NAD.
MA(k34) for Cia1 + glucose => Cib1.
MA(k35) for Cib1 => Cia1 + glucose.
MA(k36) for Cib1 => Cfa1 + NADH.
MA(k37) for Cfa1 => G_1DH + gluconolacrone.

parameter (k38, 5e-3).
parameter (k39, 5e-3).
parameter (k40, 5e-3).

MA(k38) for glucoseext => glucose.
MA(k39) for NO3ext => NO3.
MA(k40) for NO2ext => NO2.


```
% ROBUSTNESS MEASURE
```

```
robustness([c,a,b],[0.5,0.5,0.5],F(G((N > [NADH]) & ([DAFF] > R))), [N,R], [3500,50000000],600).
```

```
% VISUALIZATION
```

```
landscape([g,a],[(0,1000000000),(0,1000000000)],F(G((N > [NADH]) & ([DAFF] > R))), [N,R], [3500,50000000], 10, 300, landG1DHADH).
```

```
% PARAMETER SEARCH
```

```
search_parameters_cmaes([g,a,b],[(0,1000000000),(0,1000000000),(0,1000000000)],F(G((N>[NADH])&([DAFF]>R))),[N,R],[3500,50000000],600).
```

4. Scripts

4.1. HSIM batch mode

This SHELL script was used to automatize HSIM simulation in batch mode on a linux server, in order to generate heat maps figures. It extracts input values for a HSIM model from a .txt file and performs HSIM computation with given parameters for all combination of the input. Finally, it generates a csv with detail on the computed model state for all given input points.

```
#!/bin/sh
# DEFAULT PARAMS
hsim_command="bhsim"
hsim_param_file="hsim_param_file.cfg"
hsim_param_values="hsim_param_values.txt" #File containing concentration inputs for the heat map to generate
IR=3600
M=3600
verbose=0
dryrun=0

die () {
    echo >&2 "$@"
    exit 1
}

usage () {
    echo "-----"
    echo "Usage:"
    echo " sh hsim.sh -c 'hsim hsim_command path' -f 'paramFile to parse' -p 'params values file' -i 'IR hsim param' -m 'm hsim
param' -v(verbose mode) -d(dry_run) ?(print usage)"
    echo "Format for params values file => ParamName:01nm,02nm,03NM... comma separated with \n for last"
    echo "-----"
    exit 1
}

while getopts ":c:d:f:i:m:b:p:vd" opt; do
    case "$opt" in
        c)    hsim_command="$OPTARG"
              ;;
        p)    hsim_param_values="$OPTARG"
              ;;
        f)    hsim_param_file="$OPTARG"
              ;;
        i)    IR="$OPTARG"
              ;;
        m)    M="$OPTARG"
              ;;
        v)    verbose=1
              ;;
        d)    dryrun=1
              ;;
        \?)  usage
              ;;
    esac
done

if [ ! -f "${hsim_param_file}" ]
then
    die "Missing param file (-f)"
    usage
fi
if [ ${dryrun} -eq 1 ]
then
    echo "#####"
    echo "# Dryrun mode, no command actually executed #"
    echo "#####"
fi
```

```

declare -a paramNames
declare -a paramValues
paramValues=()
paramNames=()
nbparams=`cat ${hsim_param_values} | wc -l`
if [ ${verbose} -eq 1 ]
then
    echo "#####"
    echo "# Verbose mode"
    echo "# - nbparams $nbparams"
    echo "# - params file $hsim_param_file : '${hsim_param_file}'"
    echo "# - params file $hsim_param_values '${hsim_param_values}':"
    cat ${hsim_param_values}
    echo "#####"
fi
index=0
while read line
do
    pName=`echo ${line} | cut -d"." -f1`
    pValues=`echo ${line} | cut -d"." -f2`
    paramNames[$index]=${pName}
    paramValues[$index]=${pValues}
    if [ ${verbose} -eq 1 ]
    then
        echo "#####"
        echo "# Verbose mode"
        echo "line: $line"
        echo "pName: $pName"
        echo "pValues: $pValues"
        echo "#####"
    fi
    (( index += 1 ))
done < ${hsim_param_values}
# create all combinations
myVar=`printf '%s/' ${paramValues[@]}`
count=1
if [ ! -d cfg ]
then
    mkdir cfg
else
    rm -f cfg/*
fi
if [ ! -d csv ]
then
    mkdir csv
else
    rm -f csv/*
fi
if [ ! -d final ]
then
    mkdir final
else
    rm -f final/*
fi

currentSet=()
for i in $(eval echo "$myVar")
do
    count2digit=`printf %02i $count`
    cp -f ${hsim_param_file} cfg/${count2digit}.cfg
    cfg_name="cfg/${count2digit}.cfg"
    IFS='/ ' read -a currentSet <<< "$i"
    for j in $(seq 1 ${nbparams});
    do
        pName=${paramNames[$j-1]}
        pValue=${currentSet[$j-1]}
        if [ ${verbose} -eq 1 ]
        then
            echo "current parameter $pName : $pValue"
        fi
    done
done
if [ ${dryrun} -eq 1 ]

```

```

        then
            echo "sed -i -e s/^init (\(.*\), \${pName}\);/init (\${pValue}, \${pName}\);/ \${cfg_name}"
        else
            sed -i -e "s/^init (\(.*\), \${pName}\);/init (\${pValue}, \${pName}\);/" \${cfg_name}
        fi
    done
    (( count += 1 ))
    if [ \${dryrun} -eq 1 ]
    then
        echo "\${hsim_command} \${cfg_name} -ir \${IR} -m \${M} -bbd csv/\${count2digit}+ &"
    else
        \${hsim_command} \${cfg_name} -ir \${IR} -m \${M} -bbd csv/\${count2digit}+ &
    fi
done
wait
count=1

for i in \$(eval echo "\${myVar}")
do
    count2digit=`printf %02i \${count}`
    echo "fileName csv/\${count2digit}-cell.csv"
    if [ -z "\${firstLineSet}" ]
    then
        firstLineSet=1
        line=$(head -n 1 csv/\${count2digit}-cell.csv)
        addedColumn=${paramNames[@]// /\t}
        if [ \${dryrun} -eq 1 ]
        then
            echo "printf '%s\n' \"\${addedColumn} \${line}\" > final/results.csv"
        else
            printf '%s\n' "\${addedColumn} \${line}" > final/results.csv
        fi
    fi
    sed 1d csv/\${count2digit}-cell.csv | while read line
    do
        addedColumn=${i//\// }
    if [ \${dryrun} -eq 1 ]
    then
        echo "printf '%s\n' \"\${addedColumn} \${line}\" >> final/results.csv"
    else
        printf '%s\n' "\${addedColumn} \${line}" >> final/results.csv
    fi
    done
    (( count += 1 ))
done

#rm -f csv/*
#rm -f cfg/*

```

4.2. ImageJ particle sizing

This script was used to automatize protocell size evaluation, taking microphotograph as inputs.

Protocell_sizing.ijm

```

setBatchMode(true);
function action(input, output, filename) {
    open(input + filename);
    run("8-bit");
    run("Enhance Contrast", "saturated=4 normalize");
    run("Threshold", "method=Default white");
    run("Watershed");
    run("Set Measurements...", "area mean center perimeter fit shape display redirect=None decimal=5");
    run("Set Scale...", "distance=852.26 known=50 pixel=1 unit=µm global");
}

```

```
run("Analyze Particles...", "size=.00-1 circularity=0.10-1.00 show=[Overlay Outlines] display exclude clear  
include");  
saveAs("PNG", output+filename);  
  
saveAs("Results", output+filename+"results.csv" );  
run("Clear Results");  
  
close();  
}  
input = "/tmp/input/";  
output = "/tmp/output/";  
list = getFileList(input);  
for (i = 0; i < list.length; i++)  
    action(input, output, list[i]);  
setBatchMode(false);
```

6. Manuscripts

DIAGNOSTICS

Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates

Alexis Courbet,¹ Drew Endy,² Eric Renard,³ Franck Molina,^{1*} Jérôme Bonnet^{4*}

Whole-cell biosensors have several advantages for the detection of biological substances and have proven to be useful analytical tools. However, several hurdles have limited whole-cell biosensor application in the clinic, primarily their unreliable operation in complex media and low signal-to-noise ratio. We report that bacterial biosensors with genetically encoded digital amplifying genetic switches can detect clinically relevant biomarkers in human urine and serum. These biosensors perform signal digitization and amplification, multiplexed signal processing with the use of Boolean logic gates, and data storage. In addition, we provide a framework with which to quantify whole-cell biosensor robustness in clinical samples together with a method for easily reprogramming the sensor module for distinct medical detection agendas. Last, we demonstrate that biosensors can be used to detect pathological glycosuria in urine from diabetic patients. These next-generation whole-cell biosensors with improved computing and amplification capacity could meet clinical requirements and should enable new approaches for medical diagnosis.

INTRODUCTION

In vitro diagnostic tests (IVDs) are growing in importance in the global health arena because of their noninvasive nature and resulting ease of use and scale (1, 2). However, conventional detection methods for IVDs are often expensive and complex, and thus difficult to implement in resource-limited settings (3). In response to these challenges, bioengineers have developed attractive methodologies that rely on synthetic nanoprobe (4–6) or microfluidics (7, 8). Yet, there remains a need for easy-to-use, portable biosensor devices that can be used by nonspecialists to make clinical measurements at home or in remote locations (4, 9, 10).

Among biosensing devices, whole-cell biosensors mainly based on bacteria have proven to be applicable for the detection and quantification of a wide range of analytes (11, 12). Living cells have many attractive properties when it comes to diagnostics development. Cells detect biomolecules with high sensitivity and specificity and are capable of integrated and complex signal processing. Cells also provide a self-manufacturing platform via autonomous replication (12, 13), and the production of laboratory prototypes can be scaled using existing industrial frameworks (14). Spores from whole-cell biosensors can remain functional for extended periods of time, increasing the shelf life of a diagnostic product in harsh storage conditions (15). Last, whole-cell biosensors are highly versatile and can be used as stand-alone devices or interfaced with other technologies such as electronics, microfluidics, or micropatterning (16–18). All of these advantages have prompted the development of whole-cell biosensors that measure a variety of clinical parameters (19–24). However, whole-cell biosensor systems have not yet been applied for the monitoring of medically relevant parameters in a clinical context.

Many challenges have limited whole-cell biosensor translation to the clinic: (i) unreliable operation and low signal-to-noise ratio in complex and heterogeneous clinical samples; (ii) the inability to engi-

neer ligand-tailored sensors; (iii) limited signal-processing capability, which precludes the integration of several biomarker signals for accurate diagnosis; (iv) lack of consistent frameworks for the assessment of robustness in challenging clinical conditions; (v) response times that are not compatible for diagnosis that require fast delivery of results; and (vi) compliance to clinical formats (fig. S1).

The emerging field of synthetic-biology research aims at streamlining the rational engineering of biological systems (25). In the field of health care, synthetic biology has delivered breakthroughs in drug biosynthesis (26–29) and new hope for compelling translational medicine applications (30–32). As proof of concept, researchers have embedded medical algorithms within living cells for diagnosis, disease classification, and treatment (33–37). However, the use of synthetic biology tools and concepts to improve IVD technologies has been limited mainly to bacteriological tests using engineered bacteriophages (38) or to a recently developed mammalian cell-based allergy profiler (39).

Synthetic biology focuses on parts and systems standardization, the engineering of modular components, and systematic strategies for the engineering of biological systems and new biological functions with reliable and predictable behaviors. Molecular modules such as sensors, reporters, or switches could ultimately be assembled at a systems level to perform specific tasks. Genetic devices that support in vivo computation were developed recently and enable living cells to perform sophisticated signal-processing operations such as Boolean logic, edge detection, or cellular profiling (40, 41). Therefore, synthetic biology could presumably support the design of cell-based biosensors that meet medical specifications and help to translate whole-cell biosensors to clinical applications.

Here, we investigate the use of recently developed digital amplifying genetic switches and logic gates (42) to bring the performance of whole-cell biosensor closer to clinical requirements. These genetic devices enabled bacteria to perform, in human clinical samples, reliable detection of clinically relevant biomarkers, multiplexing logic, and data storage. We also provide a framework for quantifying whole-cell biosensor robustness in clinical samples together with a method for easy reprogramming of the sensor module for distinct medical detection agendas. Hence, our platform architecture is highly modular and could be repurposed for various applications. We anticipate that such

¹Sys2Diag FRE3690-CNRS/ALCEDIAG, Cap Delta, 34090 Montpellier, France. ²Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. ³Department of Endocrinology, Diabetes, Nutrition, Montpellier University Hospital; INSERM 1411 Clinical Investigation Center; Institute of Functional Genomics, CNRS UMR 5203, INSERM U661, University of Montpellier, 34090 Montpellier, France. ⁴Centre de Biochimie Structurale, INSERM U1054, CNRS UMR5048, University of Montpellier, 29 Rue de Navacelles, 34090 Montpellier, France.

*Corresponding author. E-mail: jerome.bonnet@inserm.fr (J.B.); franck.molina@sys2diag.cnrs.fr (F.M.)

engineered bacterial biosensors (“bactosensors”) that are capable of in vivo computation could be tailored according to medical knowledge and used as expert biosensing devices for medical diagnosis (Fig. 1 and fig. S1).

RESULTS

Behavior and robustness of bacterial chassis in human clinical samples

Our first goal was to determine the operational characteristics of bacterial chassis of interest in terms of growth, viability, and functionality of gene networks in human body fluids of clinical relevance: urine and blood serum. We chose to evaluate the robustness of Gram-negative and Gram-positive bacterial models—*Escherichia coli* and *Bacillus subtilis*, respectively—that have been used in previous whole-cell biosensors designs. To this end, we collected urine and serum from healthy volunteers, pooled the samples to average molecule concentrations to account for possible variations among individuals, and prepared dilutions with a defined culture medium (see Materials and Methods). We then inoculated various clinical sample dilutions with cells from stationary cultures of *E. coli* DH5 α Z1 or *B. subtilis* 168, grew these cultures for 18 hours at 25°, 30°, or 37°C, and measured their optical densities. For both cell types, we observed cell growth across the complete range of sample dilutions and at all three temperatures

(fig. S2). However, growth was strongly inhibited at 100% urine or serum concentrations, probably because of the lack of nutrient provided by the diluted culture medium. Growth of both cell types decreased with increasing urine or serum concentration, but cell death was insignificant (<2% for all samples; fig. S3). These results demonstrate that both Gram-positive and Gram-negative bacteria can survive and proliferate in human clinical samples for several hours. Because of the larger number of tools available for the reliable control of gene expression (43, 44), we chose *E. coli* for further engineering of a prototype bactosensor.

We next assessed the capacity of *E. coli* cells growing in clinical samples to respond in a reliable way to exogenously added molecular signals using the model transcriptional promoters pTET [which responded to anhydrotetracycline (aTc)] and pBAD [which responded to arabinose (ara)], both driving expression of a reporter gene that encodes the green fluorescent protein (GFP). Both promoters were functional at all concentrations of clinical samples (Fig. 2, A and B), but cells that were induced in 100% serum failed to produce GFP, which indicates that serum has an inhibitory effect on bacterial gene expression. Increasing sample concentrations produced variations in auto-fluorescence, which could be corrected using a reference promoter (45). Such reference promoters could be used as internal standards to increase measurement reliability in a clinical setting (fig. S4 and note S1). These results demonstrate that synthetic gene networks can remain functional in clinical samples.

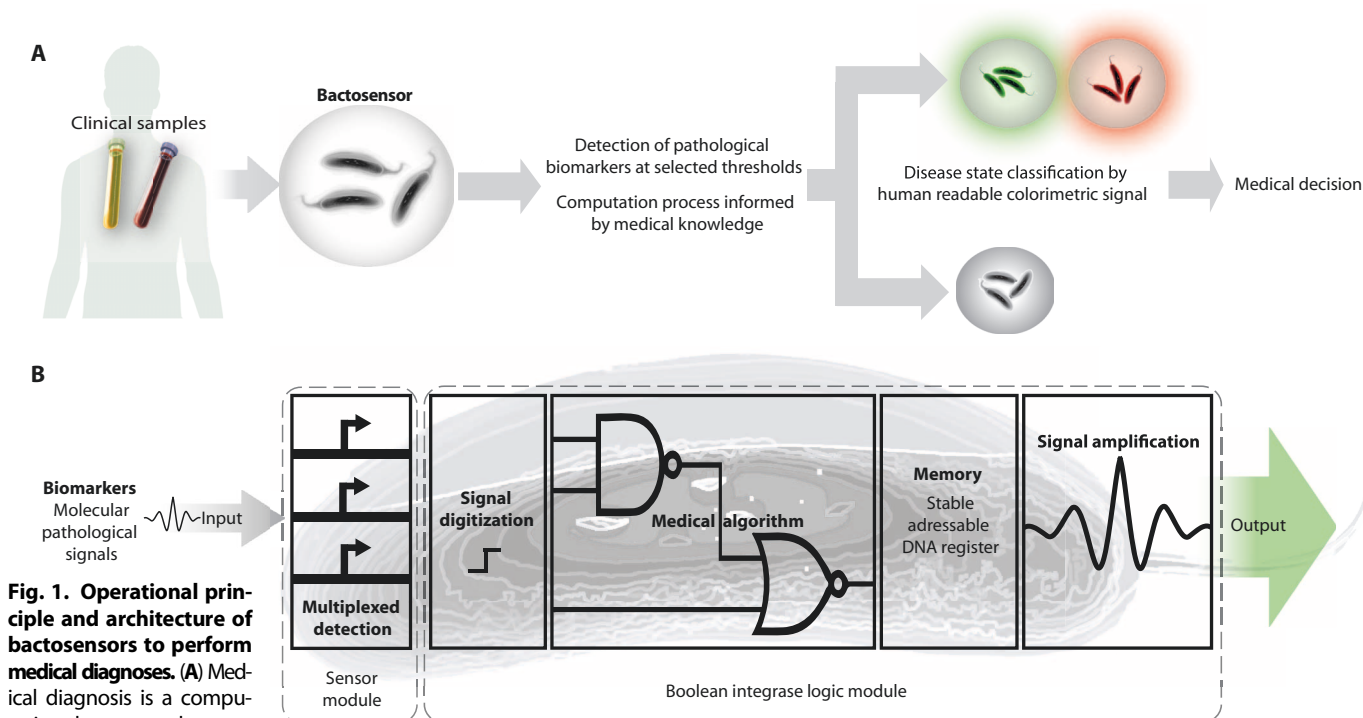


Fig. 1. Operational principle and architecture of bactosensors to perform medical diagnoses. (A) Medical diagnosis is a computational process that can be formalized using Boolean logic in vivo and embedded into a bactosensor. The bactosensor is capable of detecting a pattern of specific biomarkers in human clinical samples at selected thresholds and integrates these signals using an in vivo computational process. If a pathological pattern of biomarker is detected, the bactosensor generates a colorimetric output. (B) Schematic architecture of a bactosensor. A sensor module enables multiplexed detection

of pathological biomarkers. These control signals drive a Boolean integrase logic gate module, which is the biological support for a user-defined digital medical algorithm. Boolean integrase logic gates also enable signal digitization and amplification along with storage of the diagnosis test’s outcome in a stable DNA register that can be interrogated a posteriori.

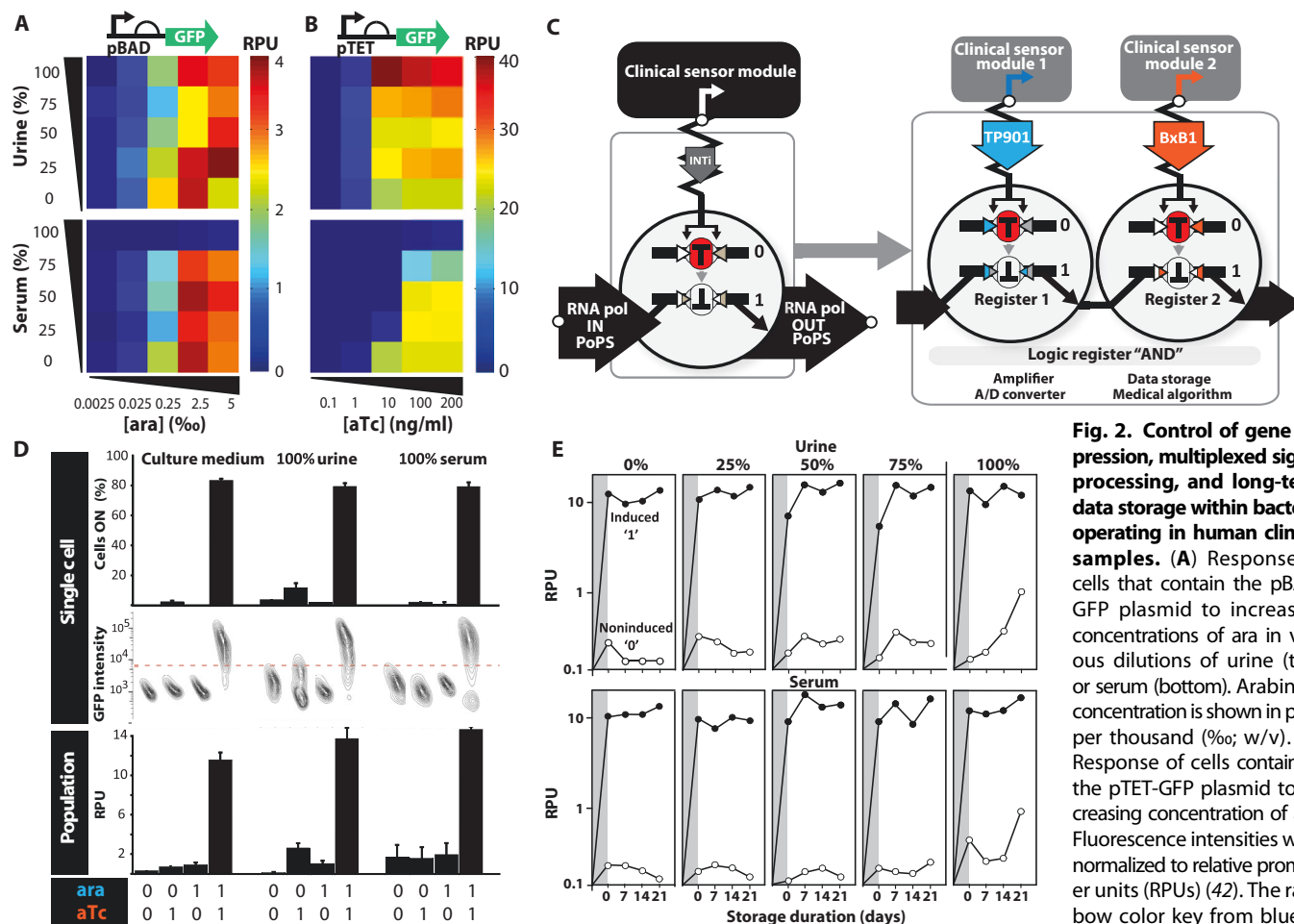


Fig. 2. Control of gene expression, multiplexed signal processing, and long-term data storage within bacteria operating in human clinical samples. (A) Response of cells that contain the pBAD-GFP plasmid to increasing concentrations of ara in various dilutions of urine (top) or serum (bottom). Arabinose concentration is shown in parts per thousand (‰; w/v). **(B)** Response of cells containing the pTET-GFP plasmid to increasing concentration of aTc. Fluorescence intensities were normalized to relative promoter units (RPUs) (42). The rainbow color key from blue to red depicts increasing signal

intensities measured in RPUs. **(C)** Architecture and functional composition of transcriptor-based digital amplifying genetic switches. The clinical sensor promoter drives integrase expression, which inverts a transcriptor module that controls the flow of RNA polymerase (RNA pol) along the DNA. Two transcriptors that respond to different signals can be composed in a series to produce an AND gate. A/D, analog to digital. **(D)** Operation of an AND gate at 25°C, at various dilutions (0, 100%) of human urine and serum in response to ara (0.5% w/v) and aTc (200 ng/ml). The 0 or 1 values symbolize absence or presence, respectively, of a particular inducer. Population (bottom, RPUs) and single-cell (top) fluorescence intensity measurements are shown. The middle row shows raw flow cytometry data (x axis: side scatter). Errors bars indicate SDs from three independent experiments, each performed in triplicate. **(E)** Stability of functional memory in various dilutions of urine (top row) or serum (bottom row) in living cells. The AND gate was switched with 0.5% (w/v) ara and aTc (200 ng/ml). Cells were then kept at 4°C for 7, 14, or 21 days and then grown overnight in fresh medium. For each medium concentration, GFP fluorescence in RPUs is represented for non-induced cells (open circles) and induced cells (filled circles). The gray-shaded regions depict the duration of the period in which cells were exposed to the inducing signal.

Multiplexing logic and memory in human clinical samples

Multiplexed biomarker assays are known to improve the performance and robustness of diagnostic tests (3). Signal processing allows an assay to integrate the detection of multiple inputs and to perform complex analytical tasks such as diagnostics algorithms informed by medical knowledge. Performing such integrated multiplexed detection and analysis within living cells thus requires some form of engineered biomolecular computation. We recently designed three-terminal devices, termed transcriptors, that use serine integrase enzymes to control the flow of RNA polymerase along DNA via unidirectional inversion of an asymmetric transcriptional terminator (42, 46).

Transcriptors are digital amplifying switches that operate as analog-to-digital converters, are capable of signal amplification, can perform data storage and record transient signals, and can be composed to pro-

duce a variety of Boolean integrase logic gates (Fig. 2C). We thus wanted to assess whether Boolean integrase logic gates could enable whole-cell biosensors, operating in clinical samples, to execute complex signal-processing algorithms. To improve the performance of synthetic circuits for the clinics, we incorporated in our design recently developed standardized regulatory genetic elements (see note S2). We first evaluated the functionality and robustness of an AND gate that responded to ara and aTc in clinical samples and found that the logic gate operated reliably at room temperature in 100% urine and serum (Fig. 2D). We obtained similar results using NAND and NOR gates (fig. S5). Moreover, after gate switching, cells stored at 4°C could be regrown and the fluorescent output measured after up to 3 weeks of storage time (Fig. 2E). Moreover, the signal stored within the DNA register could be recovered from bacterial cells that had been dead for

8 months using the polymerase chain reaction (PCR) or Sanger DNA sequencing (fig. S6). Together, these results show that living cells with embedded Boolean integrase logic gates can perform programmable, multiplexed signal processing in clinical samples. The ability to perform stable data storage over extended periods of times provides new opportunities for delayed readout in clinical environments.

Analytical evaluation of biosensors for the detection of biological parameters in clinical samples

We next aimed to detect signals of clinical interest in urine and serum. We first chose nitrogen oxides (NO_x), a biomarker for various pathologies involving inflammation (47). Using a GFP reporter, we measured the response of pYeaR, a nitrite/nitrate-sensitive transcriptional promoter (48), to increasing concentrations of NO_x at various urine and serum dilutions (Fig. 3A). The pYeaR activation threshold decreased with increasing concentrations of urine or serum and was activated in 100% urine without the addition of NO_x, probably due to the presence of endogenous NO_x. Moreover, pYeaR was totally inhibited in 100% serum. These results highlight the sensitivity of whole-cell biosensors to context perturbations that need to be overcome for successful medical applications.

Thresholding, digitization, and amplification of biologically relevant signals in clinical samples using digital amplifying genetic switches

We then tested whether transcription-based digital amplifying switches could improve the detection of signals of clinical interest. We thus cloned pYeaR upstream of the *Bxb1* or *TP901-1* integrase genes, which encode the enzymes that control transcription switching. Because even weak, leaky promoter activity can drive integrase expression and non-specific switching, we used a directed evolution approach that combined randomization of regions that regulate integrase gene expression [that is, the ribosomal binding site (RBS), ATG translation start codon, and nucleic acid sequences that encode a C-terminal SsrA tag for cytoplasmic degradation] coupled with bacterial library screening to obtain switches that activated in response to NO_x (Fig. 3B). From this library, we selected and characterized three switches that contained variations in the RBS, start codon, integrase type, and SsrA proteolysis tag (46, 49). Switches were activated at different NO_x thresholds that spanned several orders of magnitude (Fig. 3C). These data suggest that digital amplifying genetic switches could be tailored to detect a specific biomarker over defined pathological thresholds that meet clinical requirements (fig. S7).

We then mapped the transfer function of one of the switches at various sample dilutions (Fig. 3D) and observed signal digitization and marked improvement of the signal-to-noise ratio compared to the pYeaR-GFP construct. The inhibitory effect of 100% serum on NO_x detection was overcome, although signal interference was still observed in 100% urine. Using an amplification reporter system, we quantified pYeaR switch-mediated signal amplification across a range of signals and sample concentrations (Fig. 3E) and measured maximum gain values between 10 and 15 dB (decibels). These results show that digital amplifying switches increase the robustness of whole-cell biosensing systems and could thus enable the development of clinically compliant biosensors.

Detection of a metabolized biological signal—glucose—in clinical samples

Glucose is a biomarker of clinical interest whose blood levels can be used for the monitoring of diabetes (high blood glucose or glycemia)

and whose presence in urine (glycosuria) marks the onset of or presence of uncontrolled diabetes. Point-of-care technologies that enable clinicians to detect glycosuria or continuously monitor glycemia remotely can greatly improve and simplify care of diabetic patients (50). The fact that glucose is one of the primary carbon sources metabolized by bacterial cells makes it a challenging molecular signal to monitor. To perform glucose detection, we chose the pCpxP promoter as a driver of target gene expression, which is activated in the presence of glucose, pyruvate, or acetate (51). pCpxP showed a high basal level of expression in bacterial growth medium and human serum and a low signal-to-noise ratio (Fig. 4A, maximum fold change ~1.5 in medium, ~2.2 in urine, and ~1.7 in serum). Moreover, pCpxP was inhibited by increasing urine concentrations and glucose concentration greater than 10⁻² M. For the latter case, we confirmed, by kinetic assays, the time-dependent inhibition of pCpxP putatively as a result of a glucose-induced drop in the pH of the medium (52) (fig. S8).

We next built a pCpxP switch and again observed a marked improvement in the signal-to-noise ratio (Fig. 4A, maximum fold changes ~12.4 in medium, 12.6 in urine, and ~20.6 in serum) and a near-digital switching (that is, the system responded in a nearly all-or-none fashion). Response of the pCpxP switch to glucose was detectable up to 100% urines, indicating that a low signal produced by the promoter in these conditions was detected, amplified, and stored by the switch. The transient pCpxP activity at high glucose concentrations was also detected and stored by the pCpxP switch (Fig. 4A and fig. S8). Therefore, the detection of multiple clinically relevant signals can be systematically improved by digital amplifying switches. Finally, as a proof of concept, we performed dual detection and multiplexed signal processing of clinical biomarkers by building two-input logic gates controlled by NO_x and glucose and performing various computation processes (fig. S9).

Quantifying biosensor robustness in clinical samples

We next quantified the improvement in signal digitization conferred by the digital amplifying switches. For both pYeaR and pCpxP, we measured the digitization error rates (DERs; the combined probability of scoring a false-positive or a false-negative) (42) of promoter-only constructs and promoter-switch constructs in the presence of minimal and maximal inducer concentrations (fig. S10). We found that sensors that incorporated digital amplifying switches generally displayed a reduced DER, demonstrating the improvement in signal digitization provided by the switches.

We then aimed at establishing a quantitative framework with which to evaluate the robustness of the biosensor response against clinical sample-induced perturbations (fig. S11). At each inducer concentration and for each clinical sample dilution, we quantified the change in signal relative to cells grown in culture medium. The relative changes in signal values were averaged to obtain a global robustness score (RS), which was inversely proportional to the robustness of the biosensor against sample-induced perturbations. For pCpxP, use of the switch reduced RS values from 0.6 to 0.3 in urine and from 0.27 to 0.20 in serum. For pYeaR, RS values decreased from 2.1 to 0.44 in urine and from 1.15 to 0.69 in serum. Using digital amplifying switches thus systematically improved the robustness of the biosensor response against sample perturbations. Part of this improvement in robustness also resulted from the use of standard parts for the translational control of the switch output (43) (fig. S12 and note S2).

Together, these results demonstrate that digital amplifying switches can markedly improve the reliability of the detection of clinically relevant signals in clinical samples.

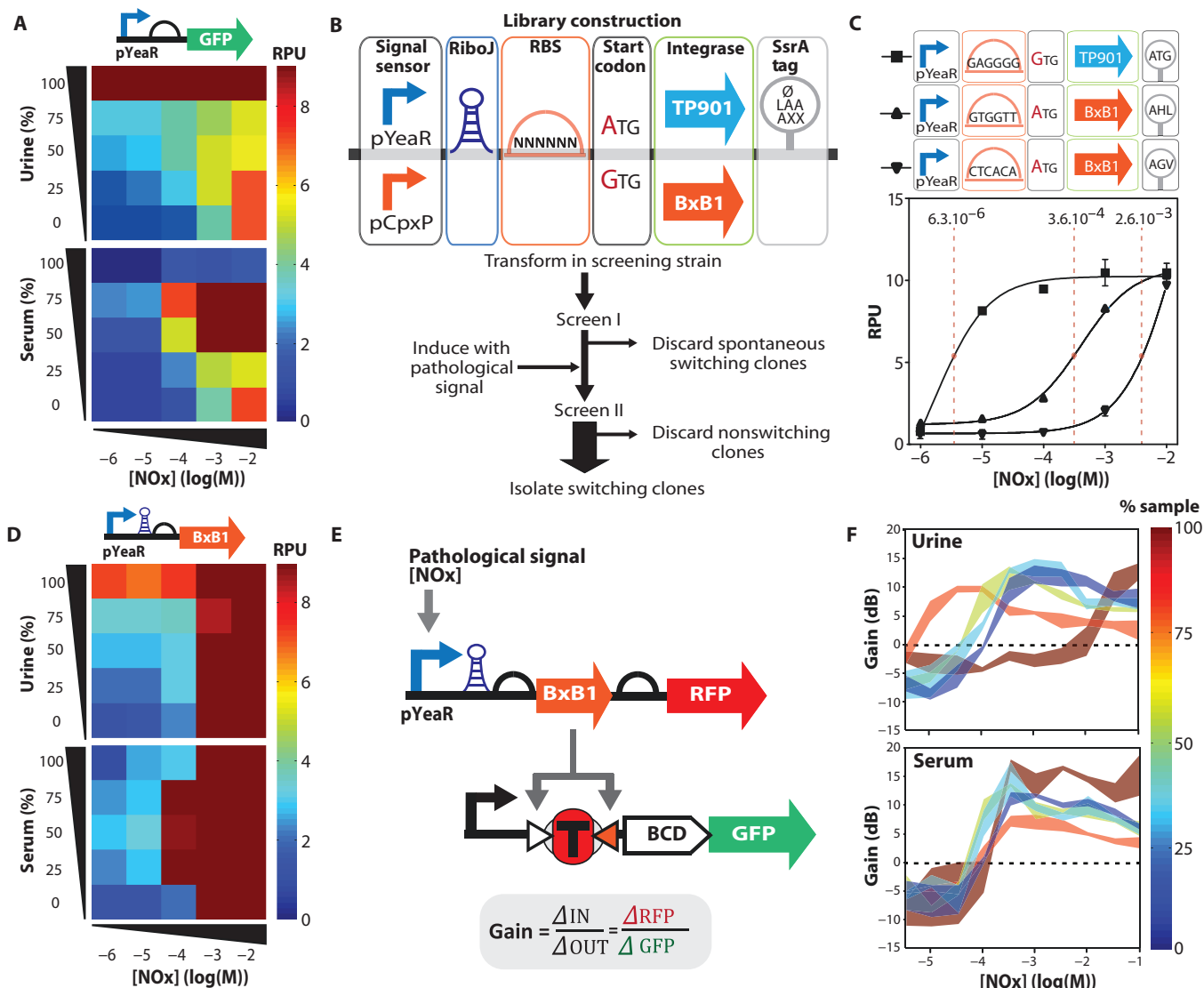


Fig. 3. Thresholding, digitization, and amplification of biologically relevant signals in clinical samples using amplifying genetic switches.

(A) Characterization of the NO_x-responsive promoter pYeaR driving expression of GFP in various dilutions of serum and urine. The rainbow color key from blue to red depicts increasing signal intensities measured in RPU. (B) Workflow for connecting biological signal-responder promoters to amplifying digital switches. An integrase expression cassette library driven by promoters of interest was built by introducing combinatorial diversity in RBSs, start codons, and C-terminal SsrA degradation tags. Libraries were transformed in a screening strain, spontaneously switching clones were eliminated, and the remaining cells were induced with the biological signal of interest. Switching clones were identified on a plate reader or using a fluorescence-activated cell sorter and isolated (see Materials and Methods for details). (C) Multiple switching threshold for biomarker detection. Clones were isolated from the various pYeaR li-

braries and characterized. Midpoint switching values are indicated. Variation in sequences among the isolated clones is depicted in the upper panel, along with the correspondence between graphs symbols and a particular switch sequence. (D) Digitization and amplification of the NO_x signal in urine and serum using amplifying digital switches. Cells cotransformed with pYeaR switch and exclusive OR (XOR)-GFP gates (42) were induced with NO_x, and bulk fluorescence was measured on a plate reader. (E) The plasmid to measure amplification of NO_x input consists of bicistronic BxB1-RFP cassette driven by pYeaR. This construct was cotransformed with a XOR-GFP gate to enable the precise simultaneous measurement of fold change in input control signal (RFP) and output signal (GFP) after induction with NO_x. (F) In vivo molecular pathological signal amplification in clinical samples. Gain in decibels was calculated as the 10log of the RFP/GFP ratio. The line thickness represents the SD over three independent experiments.

Bactosensor detection of glycosuria in clinical samples from diabetic patients

To assess the relevance of digital amplifying switches for disease detection in a clinical assay, we sought to develop a proof-of-concept sensor that detects endogenous levels of a pathological biomarker

in clinical samples from patients. As a preliminary validation and proof of concept, we aimed to detect glycosuria using the pCpxP switch.

To do so, we used a prototype clinical format and encapsulated viable bactosensors in polyvinyl acetate (PVA)/alginate hydrogel beads

Downloaded from stm.sciencemag.org on May 28, 2015

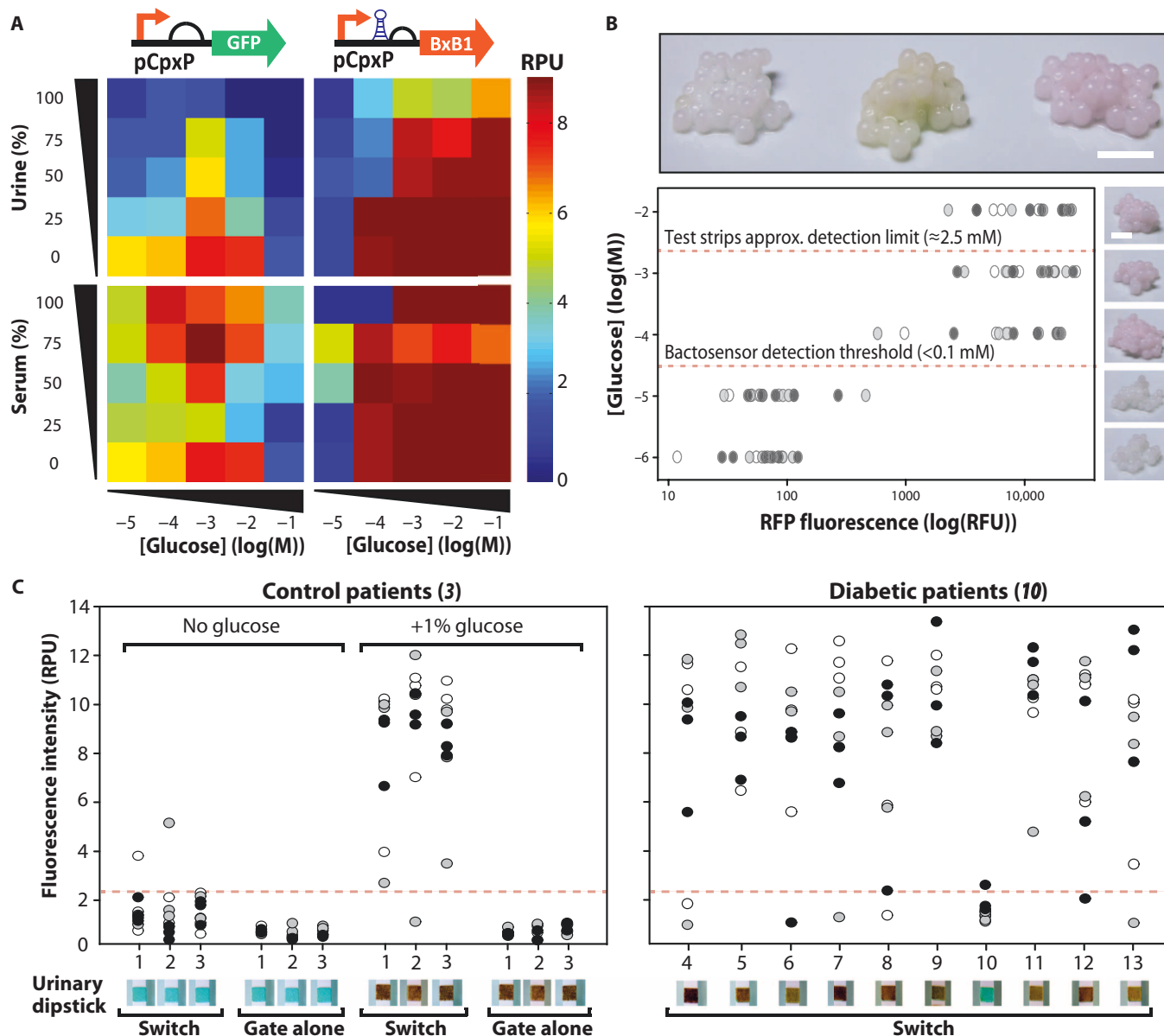


Fig. 4. Bactosensor-mediated detection of pathological glucose levels in clinical samples from diabetic patients. (A) Comparison of the response of pCpxP-GFP (left) and pCpxP-Bxb1 (right) constructs to increasing concentrations of glucose in various concentrations of urine (upper panel) or serum (lower panel). (B) Upper panel: Insulation of bactosensors in stable hydrogel beads. Left: Uninduced beads. Middle: NO_x-induced beads with GFP reporter. Right: NO_x-induced beads with RFP reporter. Scale bar, 0.5 cm. Lower panel: Operability of amplifying digital switches encapsulated in PVA/alginate beads. Beads that contained cells cotransformed with the XOR-RFP gate and pCpxP-TP901 were incubated in culture medium supplemented with various concentrations of glucose, and RFP fluorescence was measured after 24 hours. RFU, relative fluorescence unit. Detection thresholds for urinary dipsticks and for bactosensor are indicated. Pictures of the beads at various inducer concentrations are shown. Scale bar, 0.5 cm. (C) Bactosensor-

mediated detection of abnormal glycosuria levels in clinical samples from diabetic patients. PVA/alginate beads encapsulating cells transformed with both the pCpxP-TP901 controller and the XOR-GFP gate, XOR-GFP gate alone, or the reference construct J23100-GFP were incubated in urine samples. Left panel: three glucose-negative samples from independent individuals and three positive controls [same samples supplemented with 1% (w/v) glucose]. Right panel: Urine samples from 10 nonstabilized individual diabetic patients. GFP fluorescence was measured after 24 hours. Response to glucose was compared with standard urinary dipsticks (lower panels). The lower panels show the glucose reactive band of the urinary dipstick. In the absence of glucose, the band is blue. When glucose is present, the band turns brown. Data from three different experiments performed on different days are depicted by black, gray, and white circles. Each circle corresponds to one replicate (three replicates for each experiment).

(Fig. 4B and note S3) (53–55). First, using the red fluorescent protein (RFP) as a reporter, we tested the response of beads that encapsulated the pCpxP or pYeaR switch to increasing inducer concentrations in

culture medium and observed digital switching detectable with the naked eye (Fig. 4B and fig. S13). The pCpxP switch was activated at a threshold concentration under 0.1 mM glucose, outperforming the

detection limit of urinary dipsticks, the gold standard point-of-care test for glycosuria, by an order of magnitude.

We then tested the beads in urine samples. The pCpxP switch beads produced a robust and specific response in nonpathological urine samples exogenously supplemented with glucose (Fig. 4C, left panel). Finally, we tested the pCpxP switch beads in individual urine samples from 13 patients diagnosed with diabetes but not yet stabilized (Fig. 4C, right panel). The assay reliably detected glycosuria in samples from diabetic patients, with a sensitivity of 88.9% and a specificity of 96.3% (fig. S14). We observed some variability in beads response, which we attributed to our bead fabrication process (see note S3). Improvements in the encapsulation process should increase the reliability of our assay and reduce bead-to-bead variability. Nevertheless, our system was capable of reliably detecting the presence of endogenous glucose in urine from 12 different diabetic patients, suggesting that bactorsensors are relatively robust when faced with interindividual variations in sample composition. Together, our data demonstrate that digital amplifying genetic switches can enable whole-cell biosensor to operate in a clinical assay and detect endogenous biomarkers of disease in patient samples.

DISCUSSION

The past decade has witnessed the development of innovative biodiagnostic technologies and biosensor approaches, promising a new era of fast, versatile, easy-to-use, and reliable point-of-care diagnostic devices (4). However, within the biosensing device family, whole-cell biosensors—despite their potential—have not been able to be translated into real-world clinical applications. Here, we bring whole-cell biosensors closer to medical application by addressing some of the limitations that have hindered their translation to the clinics.

As a prototype, our system demonstrates that digital amplifying switches and logic gates can overcome several typical problems faced in the clinical application of whole-cell biosensors: low signal-to-noise ratios, partial inhibition of the sensor by the samples, and logic processing of multiple biomarkers signals. Digitalizing along with amplifying and multiplexing signal detection improved sensitivity, mediated sharp response profiles, and offered an all-or-none response based on a pathological biomarker threshold (for example, detection of fold change). Moreover, digital switching provided constant outputs and dynamic ranges irrelevant of the control signal and greatly facilitated clinical assay standardization and high-throughput measurements. In addition, transient signals (glucose) that are undetectable in endpoint assays using conventional whole-cell biosensors were detected and stored using our system. This long-term data storage property enables diagnostics tests to be performed and results stored for several months under harsh conditions. Last, we have provided a quantitative framework to evaluate the function and robustness of whole-cell biosensors in clinical samples.

Several hurdles still need to be solved to fully translate whole-cell biosensors into clinical applications. First, methods to engineer new sensing modules tailored to detect ligands of interest are lacking. Current research efforts focus on mining databases for transcriptional regulators that respond to various biological signals and engineering tailored ligand-responsive RNA switches or transmembrane receptors (56–59). Recent encouraging successes suggest that we will witness significant progress in this field in the near future. Multiple sensors that specifically respond to clinically relevant biomarkers could then be

connected to Boolean integrase logic gates to perform multiplexed biomarker detection and analysis in clinical samples (fig. S15).

Second, response times of bactorsensors can be too long for clinical assays. For example, here we measured signal after 18-hour incubations and were able to detect an interpretable output response after 4 to 5 hours (fig. S8). Given the current response time of our assay, it is thus unlikely to compete with test strips for the detection of glycosuria. Further work should thus be devoted to obtaining the shorter response times needed for diagnostic tests, for example, by interfacing our sensors with electronic devices or engineering circuits that rely on post-translational signaling (such as protein phosphorylation). Nevertheless, whereas long measurement times are not compatible with timely diagnosis for certain applications (toxicological emergencies), our system would still be relevant for addressing certain medical questions that require less urgent results such as large-scale population screening, monitoring of chronic disease evolution, or companion diagnostics.

From a broader perspective, our work could be a stepping stone toward future applications that use living cells to perform *in vivo* diagnostics coupled with *in situ* synthesis and delivery of therapeutic molecules (60). Last, although our study addresses robustness and standardization issues essential for commercialization approval, regulatory and safety concerns regarding the use of engineered living organisms in the clinics remain, and societal and ethical questions must be addressed before such agents can be effectively used in the clinic (61). All switches, logic gates, and uses demonstrated or disclosed herein have been contributed to the public domain via the BioBrick public agreement (62).

MATERIALS AND METHODS

Study design

Our goal in this study was to investigate the use of recently developed digital amplifying genetic switches and logic gates (42) to bring the performance of whole-cell biosensor closer to clinical requirements. In particular, we wanted to assess whether digital amplifying switches could overcome typical problems faced in the clinical application of whole-cell biosensors, such as low signal-to-noise ratios, partial inhibition of the sensor by the samples, and logic processing of multiple biomarker signals. Using glycosuria as a model system, we aimed at demonstrating the detection of an endogenous clinically relevant biomarker in a clinical setup using samples from diabetic patients.

We developed a technological platform that was used to build several biosensors capable of detecting various biomolecules. To this aim, we use synthetic biology principles (including standardization and modularity) and provided a method to couple new detection sources to our system. The gates are fully modular (that is, the logic can be easily altered by changing the target DNA sequence for the recombinases) (see figs. S5 and S9).

To evaluate the robustness of our system and its functionality in clinical samples, we used serum and urine pools from healthy individuals as well as urine samples from healthy individuals and diabetic patients. Regarding collection of clinical samples, nonpathological (control) and glycosuric (diabetic) urine samples were obtained from the Department of Endocrinology of the Lapeyronie Hospital, Montpellier, France, under the supervision of E. Renard. Individual informed consents were obtained from the patients and control individuals. Glycosuric urine samples were collected from 10 newly discovered, nonstabilized

diabetic patients. Human serum pools from numerous blood donors were obtained from the Etablissement Français du Sang, Montpellier, France. Serum was heat-inactivated by incubation in a 56°C bath for 30 min. Serum and urine samples were stored at -80°C before use.

Molecular biology

Constructs used in this study were cloned using standard molecular biology procedures or one-step isothermal assembly (63). All enzymes were purchased from New England BioLabs (NEB). PCRs were performed using the Q5 PCR mastermix (NEB, 1-min extension time per kilobase). Most primers were generated using J5 [j5 DNA Assembly Design Automation Software (64); <http://j5.jbei.org>]. Primers were purchased from Eurofins Genomics and IDT (Carlsbad, USA). Detailed information and plasmid maps, primers, and Gblocks sequences can be found in the Supplementary Materials.

Boolean integrase logic

GFP Boolean integrase logic gates and pBAD/pTET plasmid constructs used in this study come from previous work (42). XOR, AND, NOR, and NAND gates were then modified by Gibson assembly to replace GFP with mKate2.

Library design, construction, and screening

To obtain functional synthetic networks, we needed to finely tune gates operation so that translation levels of integrases match relevant clinical dynamic ranges for our application. To introduce diversity within integrase expression cassettes, we built combinatorial libraries of pCpxP and pYeaR promoters driving expression of integrase TP901 and BxB1, by using primers (JB 587, JB588, with G1005) and randomizing (i) RBS (4096 variants), (ii) RBS and initiation codon (8192 variants), and (iii) RBS, initiation codon, and SsrA tag (AXX) (1,179,648 variants). Effective randomization at specific positions was achieved using degenerated primers, and amplified fragments were cloned in a medium copy plasmid (J64100, chloramphenicol resistance). The library was then electroporated in DH10B electrocompetent *E. coli* (Life Technologies) and plated on chloramphenicol plates. After overnight growth, ~8000 colonies per library were counted. The libraries were grown overnight at 30°C in 10 ml of LB with chloramphenicol and mini-prepped. The libraries were then transformed into a chemically competent screening strain containing an episomal XOR-BCD-RFP logic gate. To isolate NO_x-responsive switching clones, the pYeaR library cells were plated, and 600 clones were picked and induced overnight in 400 µl of LB with chloramphenicol with 10 mM NO_x. Clones were then screened using a plate reader by measuring RFP fluorescence levels. Different clones switching after induction were kept for further investigation, yielding controller 1 and controller 2. To obtain controller 3, the TP901 fragment library was cloned in pYeaR_J64100, the library was cotransformed with XOR-RFP gate and induced with 10 mM NO_x, and 400 clones were screened using a plate reader. To isolate glucose-responsive switching clones, the pCpxP-BxB1 library was cotransformed with XOR-RFP gate and sorted using a FACSAria (BD Biosciences): On a first sort step, constitutively switching cells were discarded and the remaining clones were kept and then induced in LB medium containing 0.5% (w/v) glucose for 6 hours. After induction, cells were washed and grown in fresh LB medium overnight at 30°C. On a second sort step, switching cells (~1000 clones) were kept, and nonswitching cells were discarded. One pCpxP-BxB1 controller clone was finally kept for use.

Beads assay

To test the operability of bactorsensors in PVA/alginate beads, beads were inoculated in 300 µl of culture medium with or without inducer, or urine from patient diluted at a ratio of 1:4 in culture medium for a total volume of 300 µl in a 96-well plate. After 24 hours of incubation, fluorescence was read using a synergy H1 plate reader (more details on encapsulation of bacteria in beads can be found in the Supplementary Materials). We concomitantly tested these urines from nonstabilized diabetic patients using the Siemens Multistix 8 SG reagent strip according to the manufacturer's protocol.

Cell culture and data collection

We used *E. coli* DH5αZ1 and *B. subtilis* 168 1A1 for all measurements. Cells were cultivated with shaking at 400 rpm and grown for 18 hours at 25°, 30°, or 37°C, in either Azure Media (Teknova) or LB phosphate buffer adjusted to pH 7. Antibiotics used were carbenicillin (25 µg/ml), kanamycin (30 µg/ml), and chloramphenicol (25 µg/ml) (Sigma). D-Glucose, nitrate, L-ara, and aTc (Sigma) were used at final concentrations of 0.5% (w/v), 0.1 M, 0.5% (w/v), and 200 ng/ml, respectively. Cells were streaked from a glycerol stock, and then one clone was inoculated in 5 ml of LB with carbenicillin and/or chloramphenicol and grown overnight at 30°C. The cells were then diluted at a ratio of 1:200 and grown for 6 hours at 30°C until an optical density (OD) of ~0.5. The cells were then back-diluted at a ratio of 1:100 into 1 ml of azure medium (Teknova) and diluted with urine or serum, induced with 0.5% (w/v) ara, aTc (200 ng/ml), 0.5% (w/v) glucose, or 10 mM NO₃⁻, and grown for 18 hours at 25°C in 96 DeepWell plates. The next morning, the cells were put on ice, and we measured RFP/GFP fluorescence levels (588ex/633em, 485ex/528em, respectively) and OD₆₀₀ using a synergy H1 plate reader (BioTek) and a Beckman Coulter FC 500 flow cytometer recording 50,000 events per samples. Events were gated on forward and side scatter to exclude debris, dead cells, and doublets. The overnight growth, back dilution, and measurement procedure were performed three times on separate days in triplicates. Measurements for each data point were normalized using a reference promoter (BBa_J23101) driving expression of sfGFP (low-copy plasmid pSC101 origin with chloramphenicol resistance).

For functional and genetic memory experiments, cells were cotransformed with pTET/pBAD dual controller plasmid and AND-BCD logic gate plasmid, and induced overnight at 25°C with 0.5% (w/v) ara and aTc (200 ng/ml), in 300 µl of urine, serum, or Azure medium in p96 plates. The plates were kept for 8 months at 4°C. Plasmid DNA was then recovered by scrapping and dissolving the dry cellular residues of cells in phosphate-buffered saline, and extracted using QIAamp (Qiagen) kit. We used specifically designed primers (attL/attR Bxb1 or TP901) to PCR-amplify the recombined targets. For Sanger sequencing experiments, the gate plasmid DNA was amplified using primers G1004 and G1005, and the PCR product was sent for sequencing.

Data analysis and statistics

Experimental values are reported as means ± SD. All experiments were performed at least three times on different days and in triplicate. Data, statistics, graphs, and tables were processed and generated using MATLAB (MathWorks), SigmaPlot (Systat Software Inc.), and the R with ggplot2 package. Flow cytometry was performed using an FC 500 (Beckman Coulter Inc.), and data were analyzed using FlowJo and Flowing Software (Turku Centre for Biotechnology). We used RPU to integrate into clinical measurements an in vivo internal standard for

bactosensor operation and signal generation (45). For signal amplification experiments, amplification was quantified by the gain defined as the 10log ratio between the fractional change in the output signal GFP and the fractional change in the input signal RFP. For receiver operating characteristic analysis, a set of 27 measurements performed in nonpathological urine were compared to 27 measurements performed in urine containing 1% glucose. See the Supplementary Materials for details on calculations.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/7/289/289ra83/DC1

Materials and Methods

Notes

Fig. S1. Conceptual workflow for the systematic development of medical bactosensors.

Fig. S2. Bacterial chassis growth characteristics in urine and serum.

Fig. S3. Bacterial chassis viability in urine and serum from single-cell measurements.

Fig. S4. Influence of clinical media (urine and serum) on GFP fluorescence output generation and measurement.

Fig. S5. Population and single-cell measurements of multiplexing Boolean integrase logic gates operation in urine and serum.

Fig. S6. Stability of genetic memory in clinical samples by addressing DNA register with Sanger sequencing.

Fig. S7. Workflow for engineering switches that respond to biological signals over user-defined thresholds.

Fig. S8. Kinetic measurements and transfer functions of promoters and switches.

Fig. S9. Multiplexing detection of glucose and NO_x with AND-BCD-RFP, NAND-BCD-RFP, and NOR-BCD-RFP Boolean integrase logic gates.

Fig. S10. Single-cell measurements of fold changes and DERs for promoters and switches.

Fig. S11. Evaluation of the robustness of inducible systems against clinical media-induced perturbation (urine and serum).

Fig. S12. Comparison of the operational characteristic of AND logic gates with and without the bicistronic device (BCD) for their use in clinical samples.

Fig. S13. Insulation of bactosensors in hydrogel beads and pathological signal detection.

Fig. S14. ROC analysis for bactosensor-mediated detection of glucose in urine.

Fig. S15. Potential modalities of bactosensor-based diagnosis and composability of integrase-based logic for the development of decision-making tests.

Data file S1. Numerical values of OD₆₀₀ and fluorescence from plate reader measurements and calculated RPU's used to develop plots and figures.

Note S1. Experimental design: Use of reference promoters.

Note S2. Genetic design: Use of standardized genetic elements.

Note S3. Bacterial encapsulation toward a clinical format.

References (65–71)

REFERENCES AND NOTES

- D. C. Hay Burgess, J. Wasserman, C. A. Dahl, Global health diagnostics. *Nature* **444**, 1–2 (2006).
- E. Fu, P. Yager, P. N. Floriano, N. Christodoulides, J. T. McDevitt, Perspective on diagnostics for global health. *IEEE Pulse* **2**, 40–50 (2011).
- D. A. Giljohann, C. A. Mirkin, Drivers of biodiagnostic development. *Nature* **462**, 461–464 (2009).
- A. P. F. Turner, Biosensors: Sense and sensibility. *Chem. Soc. Rev.* **42**, 3184–3196 (2013).
- R. De la Rica, M. M. Stevens, Plasmonic ELISA for the ultrasensitive detection of disease biomarkers with the naked eye. *Nat. Nanotechnol.* **7**, 821–824 (2012).
- X. Chi, D. Huang, Z. Zhao, Z. Zhou, Z. Yin, J. Gao, Nanoprobes for in vitro diagnostics of cancer and infectious diseases. *Biomaterials* **33**, 189–206 (2012).
- A. W. Martinez, S. T. Phillips, G. M. Whitesides, E. Carrilho, Diagnostics for the developing world: Microfluidic paper-based analytical devices. *Anal. Chem.* **82**, 3–10 (2010).
- P. Yager, T. Edwards, E. Fu, K. Helton, K. Nelson, M. R. Tam, B. H. Weigl, Microfluidic diagnostic technologies for global public health. *Nature* **442**, 412–418 (2006).
- M. U. Ahmed, I. Saaem, P. C. Wu, A. S. Brown, Personalized diagnostics and biosensors: A review of the biology and technology needed for personalized medicine. *Crit. Rev. Biotechnol.* **34**, 180–196 (2014).
- V. Gubala, L. F. Harris, A. J. Ricco, M. X. Tan, D. E. Williams, Point of care diagnostics: Status and future. *Anal. Chem.* **84**, 487–515 (2012).
- L. Su, W. Jia, C. Hou, Y. Lei, Microbial biosensors: A review. *Biosens. Bioelectron.* **26**, 1788–1799 (2011).
- J. R. Van der Meer, S. Belkin, Where microbiology meets microengineering: Design and applications of reporter bacteria. *Nat. Rev. Microbiol.* **8**, 511–522 (2010).
- N. Raut, G. O'Connor, P. Pasini, S. Daunert, Engineered cells as biosensing systems in biomedical analysis. *Anal. Bioanal. Chem.* **402**, 3147–3159 (2012).
- American Society for Microbiology, *Manual of Industrial Microbiology and Biotechnology* (ASM Press, Washington, DC, 2010).
- A. Date, P. Pasini, A. Sangal, S. Daunert, Packaging sensing cells in spores for long-term preservation of sensors: A tool for biomedical and environmental analysis. *Anal. Chem.* **82**, 6098–6103 (2010).
- F. Cortés-Salazar, S. Beggah, J. R. van der Meer, H. H. Girault, Electrochemical As(III) whole-cell based biochip sensor. *Biosens. Bioelectron.* **47**, 237–242 (2013).
- A. Prindle, P. Samayoa, I. Razinkov, T. Danino, L. S. Tsimring, J. Hasty, A sensing array of radically coupled genetic 'biopixels'. *Nature* **481**, 39–44 (2011).
- S. Melamed, T. Elad, S. Belkin, Microbial sensor cell arrays. *Curr. Opin. Biotechnol.* **23**, 2–8 (2012).
- J. Horswell, S. Dickson, Use of biosensors to screen urine samples for potentially toxic chemicals. *J. Anal. Toxicol.* **27**, 372–376 (2003).
- C. Lewis, S. Beggah, C. Pook, C. Guitart, C. Redshaw, J. R. van der Meer, J. W. Readman, T. Galloway, Novel use of a whole cell *E. coli* bioreporter as a urinary exposure biomarker. *Environ. Sci. Technol.* **43**, 423–428 (2009).
- A. Roda, P. Pasini, N. Mirasoli, M. Guardigli, C. Russo, M. Musiani, M. Baraldini, Sensitive determination of urinary mercury (II) by a bioluminescent transgenic bacteria-based biosensor. *Anal. Lett.* **34**, 29–41 (2001).
- K. Turner, S. Xu, P. Pasini, S. Deo, L. Bachas, S. Daunert, Hydroxylated polychlorinated biphenyl detection based on a genetically engineered bioluminescent whole-cell sensing system. *Anal. Chem.* **79**, 5740–5745 (2007).
- E. Michelini, M. Magliulo, P. Leskinen, M. Virta, M. Karp, A. Roda, Recombinant cell-based bioluminescence assay for androgen bioactivity determination in clinical samples. *Clin. Chem.* **51**, 1995–1998 (2005).
- H. M. Alloush, E. Anderson, A. D. Martin, M. W. Ruddock, J. E. Angell, P. J. Hill, P. Mehta, M. A. Smith, J. G. Smith, V. C. Salisbury, A bioluminescent microbial biosensor for in vitro pretreatment assessment of cytarabine efficacy in leukemia. *Clin. Chem.* **56**, 1862–1870 (2010).
- D. Endy, Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
- C. J. Paddon, J. D. Keasling, Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
- M. N. Thaker, W. Wang, P. Spanogiannopoulos, N. Waglechner, A. M. King, R. Medina, G. D. Wright, Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
- P. K. Ajikumar, W. H. Xiao, K. E. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, Isoprenoid pathway optimization for Taxol precursor over-production in *Escherichia coli*. *Science* **330**, 70–74 (2010).
- K. Thodey, S. Galanie, C. D. Smolke, A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* **10**, 837–844 (2014).
- W. C. Ruder, T. Lu, J. J. Collins, Synthetic biology moving into the clinic. *Science* **333**, 1248–1252 (2011).
- W. Weber, M. Fussenegger, Emerging biomedical applications of synthetic biology. *Nat. Rev. Genet.* **13**, 21–35 (2011).
- A. G. Planson, P. Carbonell, I. Grigoras, J. L. Faulon, A retrosynthetic biology approach to therapeutics: From conception to delivery. *Curr. Opin. Biotechnol.* **6**, 948–956 (2012).
- Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* **333**, 1307–1311 (2011).
- S. Culler, K. G. Hoff, C. D. Smolke, Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins. *Science* **330**, 1251–1255 (2010).
- C. Kemmer, M. Gitzinger, M. Daoud-El Baba, V. Djonov, J. Stelling, M. Fussenegger, Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nat. Biotechnol.* **28**, 355–360 (2010).
- F. Duan, K. L. Curtis, J. C. March, Secretion of insulinotropic proteins by commensal bacteria: Rewiring the gut to treat diabetes. *Appl. Environ. Microbiol.* **23**, 7437–7438 (2008).
- Y. Y. Chen, M. C. Jensen, C. D. Smolke, Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8531–8536 (2010).
- T. K. Lu, J. Bowers, M. S. Koeris, Advancing bacteriophage-based microbial diagnostics with synthetic biology. *Trends Biotechnol.* **31**, 325–327 (2013).
- D. Ausländer, B. Eggerschwiler, C. Kemmer, B. Geering, S. Ausländer, M. Fussenegger, A designer cell-based histamine-specific human allergy profiler. *Nat. Commun.* **5**, 4408 (2014).
- Y. Benenson, Biomolecular computing systems: Principles, progress and potential. *Nat. Rev. Genet.* **7**, 455–468 (2012).
- J. A. Brophy, C. A. Voigt, Principles of genetic circuit design. *Nat. Methods* **5**, 508–520 (2014).

42. J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, Amplifying genetic logic gates. *Science* **340**, 599–603 (2013).
43. V. K. Mutalik, J. C. Guimaraes, G. Cambray, C. Lam, M. J. Christoffersen, Q. A. Mai, A. B. Tran, M. Paull, J. D. Keasling, A. P. Arkin, D. Endy, Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
44. C. Lou, B. Stanton, Y. J. Chen, B. Munsky, C. A. Voigt, Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
45. J. R. Kelly, A. J. Rubin, J. H. Davis, C. M. Ajo-Franklin, J. Cumbers, M. J. Czar, K. de Mora, A. L. Glielberman, D. D. Monie, D. Endy, Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3**, 4 (2009).
46. J. Bonnet, P. Subsoontorn, D. Endy, Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8884–8889 (2012).
47. J. N. Sharma, A. Al-Omran, S. S. Parvathy, Role of nitric oxide in inflammatory diseases. *Inflammopharmacology* **15**, 252–259 (2007).
48. H. Y. Lin, P. J. Bledsoe, V. Stewart, Activation of *yeaR-yoaG* operon transcription by the nitrate-responsive regulator NarL is independent of oxygen-responsive regulator Fnr in *Escherichia coli* K-12. *J. Bacteriol.* **189**, 7539–7548 (2007).
49. C. M. Farrell, A. D. Grossman, R. T. Sauer, Cytoplasmic degradation of *ssrA*-tagged proteins. *Mol. Microbiol.* **57**, 1750–1761 (2005).
50. American Diabetes Association, Standards of medical care in diabetes—2014. *Diabetes Care* **37**, S14–S80 (2014).
51. B. P. Lima, H. Antelmann, K. Gronau, B. K. Chi, D. Becher, S. R. Brinsmade, A. J. Wolfe, Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol. Microbiol.* **81**, 1190–1204 (2011).
52. A. J. Wolfe, N. Parikh, B. P. Lima, B. Zemaitaitis, Signal integration by the two component signal transduction response regulator CpxR. *J. Bacteriol.* **190**, 2314–2322 (2008).
53. P. S. J. Cheetham, K. W. Blunt, C. Bocke, Physical studies on cell immobilization using calcium alginate gels. *Biotechnol. Bioeng.* **21**, 2155–2168 (1979).
54. Team: Paris Bettencourt—2012.igem.org. (2012); http://2012.igem.org/Team:Paris_Bettencourt.
55. T. Takei, K. Ikeda, H. Ijima, K. Kawakami, Fabrication of poly(vinyl alcohol) hydrogel beads crosslinked using sodium sulfate for microorganism immobilization. *Process Biochem.* **46**, 566–571 (2011).
56. H. Salis, A. Tamsir, C. A. Voigt, Engineering bacterial signals and sensors. *Contrib. Microbiol.* **16**, 194–225 (2009).
57. M. N. Win, J. C. Liang, C. D. Smolke, Frameworks for programming biological function through RNA parts and devices. *Chem. Biol.* **16**, 298–310 (2009).
58. W. A. Lim, Designing customized cell signalling circuits. *Nat. Rev. Mol. Cell Biol.* **11**, 393–403 (2010).
59. M. Sadelain, R. Brentjens, I. Rivière, The promise and potential pitfalls of chimeric antigen receptors. *Curr. Opin. Immunol.* **21**, 215–223 (2009).
60. J. W. Kotula, S. J. Kerns, L. A. Shaket, L. Siraj, J. J. Collins, J. C. Way, P. A. Silver, Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4838–4843 (2014).
61. P. B. Archana Chugh, Synthetic biology based biosensors and the emerging governance issues. *Curr. Synth. Syst. Biol.* **1**, 108 (2013).
62. BioBrick public agreement, <https://biobricks.org/bpa>
63. D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
64. N. J. Hillson, R. D. Rosengarten, J. D. Keasling, j5 DNA assembly design automation software. *ACS Synth. Biol.* **1**, 14–21 (2012).
65. R. Kwok, Five hard truths for synthetic biology. *Nature* **463**, 288–290 (2010).
66. L. J. Bugaj, D. V. Schaffer, Bringing next-generation therapeutics to the clinic through synthetic biology. *Curr. Opin. Chem. Biol.* **16**, 355–361 (2012).
67. A. P. Arkin, A wise consistency: Engineering biology for conformity, reliability, predictability. *Curr. Opin. Chem. Biol.* **17**, 893–901 (2013).
68. A. Struss, P. Pasini, C. M. Ensor, N. Raut, S. Daunert, Paper strip whole cell biosensors: A portable test for the semiquantitative detection of bacterial quorum signaling molecules. *Anal. Chem.* **82**, 4457–4463 (2010).
69. E. Michelini, A. Roda, Staying alive: New perspectives on cell immobilization for biosensing purposes. *Anal. Bioanal. Chem.* **402**, 1785–1797 (2012).
70. J. Lehtinen, J. Nuutila, E. M. Lilius, Green fluorescent protein–propidium iodide (GFP-PI) based assay for flow cytometric measurement of bacterial viability. *Cytometry A* **60**, 165–172 (2004).
71. K. Otto, T. J. Silhavy, Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2287–2292 (2002).

Acknowledgments: We thank P. Amar, N. Salvétat, and members of the Endy Lab for fruitful discussions; C. Duperray and N. Vié from the Montpellier Rio Imaging cytometry platform and C. Crumpton and M. Bigos from the Stanford Shared FACS Facility for the cytometry experiments; the Centre Regional d'Imagerie Cellulaire platform and C. Cazevielle for scanning electron microscopy experiments; the Paris Bettencourt iGEM team 2012 for igniting our interest in the encapsulation of bacteria into alginate beads; and the patients for participating in this study. **Funding:** This work was supported by Stanford University, the France-Stanford Center for Interdisciplinary Studies, the Institut National de la Santé et de la Recherche Médicale (INSERM), and the Centre National pour la Recherche Scientifique (CNRS). A.C. is a recipient of a fellowship from the French Ministry of Health and a pharmacy resident at the Montpellier Centre Hospitalier Régional Universitaire (CHRU). J.B. is supported by grants from the INSERM Avenir-ATIP program and the Bettencourt Schuller Foundation. **Author contributions:** A.C., J.B., D.E., and F.M. designed the study. A.C. and J.B. performed library construction and screening. A.C. performed all other experiments and the statistical analysis. E.R. collected clinical samples. A.C. and J.B. analyzed and displayed the data and wrote the manuscript. All authors participated in the interpretation of the results and in the editing of the manuscript. **Competing interests:** The authors declare that they have no competing financial interests. **Data and materials availability:** Sequences for constructs generated in this study are deposited in GenBank (accession numbers: KM234313-23 and KM347896). Plasmids are available from Addgene.

Submitted 24 November 2014

Accepted 21 April 2015

Published 27 May 2015

10.1126/scitranslmed.aaa3601

Citation: A. Courbet, D. Endy, E. Renard, F. Molina, J. Bonnet, Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Sci. Transl. Med.* **7**, 289ra83 (2015).

Editor's Summary

A little help from our (little) friends

It's only logical: Translation of diagnostics to home health care or remote settings requires simple methods for measuring markers in complex clinical samples. And living cells—with their ability to detect biomolecules, process the signal, and respond—are logical choices as biosensing devices. The recent buzz on the human microbiota has expanded our view of bacteria beyond infectious enemies to metabolic buddies. Now, Courbet *et al.* refine that view further by engineering bacteria to serve as whole-cell diagnostic biosensors in human biological samples.

Although whole-cell biosensors have been shown to serve as analytical tools, their quirky operation and low signal-to-noise ratio in complex clinical samples have limited their use as diagnostic devices in the clinic. The authors engineered bacterial biosensors capable of signal digitization and amplification, multiplexed signal processing (with the use of Boolean logic gates), and months-long data storage. As a proof of concept, the "bactosensors" detected pathological levels of glucose in urine from diabetic patients, providing a framework for the design of sensor modules that detect diverse biomarkers for diagnostics.

A complete electronic version of this article and other services, including high-resolution figures, can be found at:

<http://stm.sciencemag.org/content/7/289/289ra83.full.html>

Supplementary Material can be found in the online version of this article at:

<http://stm.sciencemag.org/content/suppl/2015/05/22/7.289.289ra83.DC1.html>

Related Resources for this article can be found online at:

<http://stm.sciencemag.org/content/scitransmed/5/179/179ps7.full.html>

<http://stm.sciencemag.org/content/scitransmed/6/267/267ra174.full.html>

<http://stm.sciencemag.org/content/scitransmed/7/283/283rv3.full.html>

<http://stm.sciencemag.org/content/scitransmed/6/253/253rv2.full.html>

<http://stm.sciencemag.org/content/scitransmed/7/273/273re1.full.html>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Computing with Synthetic Protocells

Alexis Courbet² · Franck Molina² · Patrick Amar¹

Received: 5 February 2015 / Accepted: 5 May 2015
© Springer Science+Business Media Dordrecht 2015

Abstract In this article we present a new kind of computing device that uses biochemical reactions networks as building blocks to implement logic gates. The architecture of a computing machine relies on these generic and composable building blocks, *computation units*, that can be used in multiple instances to perform complex boolean functions. Standard logical operations are implemented by biochemical networks, encapsulated and insulated within synthetic vesicles called protocells. These protocells are capable of exchanging energy and information with each other through transmembrane electron transfer. In the paradigm of computation we propose, *protoputing*, a machine can solve only one problem and therefore has to be built specifically. Thus, the programming phase in the standard computing paradigm is represented in our approach by the set of assembly instructions (specific attachments) that directs the wiring of the protocells that constitute the machine itself. To demonstrate the computing power of protocellular machines, we apply it to solve a NP-complete problem, known to be very demanding in computing power, the 3-SAT problem. We show how to program the assembly of a machine that can verify the satisfiability of a given boolean formula. Then we show how to use the massive parallelism of these machines to verify in less than 20 min *all* the valuations of the input variables and output a fluorescent signal when the formula is satisfiable or no signal at all otherwise.

Keywords Synthetic biology · Biocomputing · 3-SAT · Protocell · Protoputing

✉ Patrick Amar
pa@lri.fr

¹ LRI, Université Paris Sud - UMR CNRS 8623, Bât. 650, 91405 Orsay Cedex, France

² Sys2diag, FRE CNRS 3690, 1682 rue de la Valsière, 34184 Montpellier, France

1 Introduction

What is Computation? One definition could be “the goal-oriented process that transforms a representation of input information into a representation of output information”. The process itself can be *iterative* (or in another form *recursive*), in this case it is called an *algorithm*, but other forms of processing can be used, such as: *neural networks* or *first order logic*.

A computation process, whatever it is, has to be run by a *computer*, which can be a human being using pen and paper, or a *machine* specifically built for that purpose. The most popular form of computer is an electronic device that use a digital representation of data, and manipulate this representation according to a set of *instructions* that implements the algorithm that transforms them into *results*. The set of instructions is then called a *computer programme*.

Electronic computers use numbers, integer and floating point, to represent data. These numbers are commonly coded in base 2, which can also be directly used to encode boolean values and therefore easily implement *conditional* calculations. Electronic computers are mainly built from basic blocks, *logic gates*, that are interconnected to make the arithmetic and logic units, memory registers and micro-controllers that form the *Central Processing Unit* which in turn, along with the *Main Storage Unit*, and the *I/O Controllers* constitute the computer itself.

Therefore, one can build a digital computer using any technology that can mimic the logic gates and their interconnections. We will demonstrate in this article how to implement single logic gates using synthetic minimal biological systems embedded in a vesicle (*protocell*) and how to connect them together to get a device (*protocellular machine*) that computes a complex logical function. The computing model that underlies our biochemical implementation of a computer is similar to the one of an electronic computer, their computing capabilities are the same.

The fundamental characteristic of electronic computers is their ability to run a potentially infinite number of algorithms doing a wide variety of computations on data, because they are *programmable*: the same computer can run sequentially (or *pseudo-concurrently*) as many different programmes as those that can reside in its main memory storage, along with the associated data.

Here, we will show how to build a reduced kind of computer that can only solve one problem, but a problem belonging to a class known to be hard to solve: a NP-complete problem.

The *computational complexity theory* explores the feasibility of computational problems, in terms of computing time (or memory space) needed to solve a problem of a given size. In the *von Neumann* based architectures (standard electronic computers) the number of computing elementary steps (instructions) is often used to approximate the computing time, since each instruction takes approximately the same amount of time to be performed.

There are two main classes of computational problems, those that can be solved by a deterministic machine in a number of steps which can be expressed as a polynomial of the problem size (class P), and those that can be solved in polynomial time, but on a *non-deterministic* machine (class NP). Typically decision problems

where (1) a solution can be verified in polynomial time and (2) there is no other known algorithm except generate and verify all the potential solutions, are NP problems. Solving these problems on a von Neumann computer require an exponential number of steps with respect to the problem size.

A NP problem is said to be *NP-complete* if any other NP problem can be transformed into this problem in polynomial time (Karp 1972). In consequence NP-complete problems are more difficult to solve than any other NP problems because if one NP-complete problem is *quickly* solved (in polynomial time) then all the NP problems will be quickly solved. Of course all these complexity classes collapse if $P = NP$ (which is one of the great open conjectures in computer science).

We have chosen the 3-SAT problem, a variant of the *boolean satisfiability problem (SAT)*, as an example of NP-complete problem (Cook 1971) a protocellular computer can solve elegantly. This is mainly because the very small size of protocells and their 3D packing allow us to build a machine made of billions of logic gates specifically connected to solve a given 3-SAT problem. Another characteristic of our protocellular machines is that they are *disposable* in the sense that once the computation is done for a given set of input values, the machine is no more usable. But the counterpart is that the energy needed for the computation is very low (Sarpeshkar 2010).

Finally, the biochemical nature of the protocellular machines make them very easy to interface with living organisms. For example, they can be used for medical diagnosis to implement biosensing coupled with medical decision algorithm.

2 Methods

2.1 Protocell Logic Gates Definitions

The bottom-up design of biological systems is made possible by the synthetic biology approach that applies engineering principles to biology in order to design standardised biological parts, devices, systems in a systematic and rational manner. Hierarchical abstraction of biological functions enables the assembly at the system level of new biological systems with user-defined functionalities (Purnick and Weiss 2009; Canton et al. 2008; Endy 2005). The behaviour of synthetic systems is predictable and designs can be automatised *in silico* before attempting to implement them with biological components (Marchisio and Stelling 2009). In addition, the remarkable capacity of biological building blocks to compute in highly sophisticated ways has led scientists to design and engineer biomolecular computers (Benenson 2012). Thus far, most biocomputing has been investigated from the *top down* perspective, that is, by modifying existing organisms (Khalil and Collins 2010). The strategy we propose here, prototyping, is interested in implementing protocells from the bottom-up perspective to perform computation, where very little attention has been given Rasmussen et al. (2009), Luisi and Stanó (2011) and Smaldon et al. (2010).

Starting from an abstract operation that is to be computed, one can rationally and systematically choose biochemical species for the implementation (metabolites,

enzymes, nucleic acids...) (Fig. 1a). Standardised and robust biomolecular components and reactions can be engineered, tested and optimised to implement different types of biological functions or computations (Koepl 2011): simple boolean operations, memory devices, amplifiers, analog to digital converter, oscillators etc. Figure 1b. In addition, this process can be automatised using CAD tools recently developed for that purpose (Koepl 2011; Rialle et al. 2010; Chandran et al. 2011).

For example, an AND biochemical logic gate taking reduced metabolites as inputs (NADPH and FADH₂) can be implemented using a network of 3 different enzymes and 4 different metabolites connected by 3 biocatalytic reactions, and transferring electrons to NADH as an output. In the same way, we can implement a set of standardised *computation units* that recapitulate all boolean logic gates (see Fig. 3 for examples of implementations of AND, NOT and NOR gates). Electron transfer can also be coupled to various output biological functions to produce human readable signals (Fig. 2) or enable the selection of machines with specific behaviour for further analysis. We propose that specific reduction of species can trigger as an output, either luminescence or fluorescence (Candeias et al. 1998) or the transport of a ligand (or its receptor).

Our approach improves modularity of biomolecular computing systems by the fact that biochemical networks implementing boolean logic are encapsulated within synthetic vesicles, or *protocells*, distinguished by their high degree of organisation and control over biological processes provided by the membrane boundary (Elani et al. 2014). Such architecture of insulated computing units allows us to use many instances of the same type of protocell anywhere in the circuit when the same logic gate is needed. Moreover, this enables the connection of multiple layers of

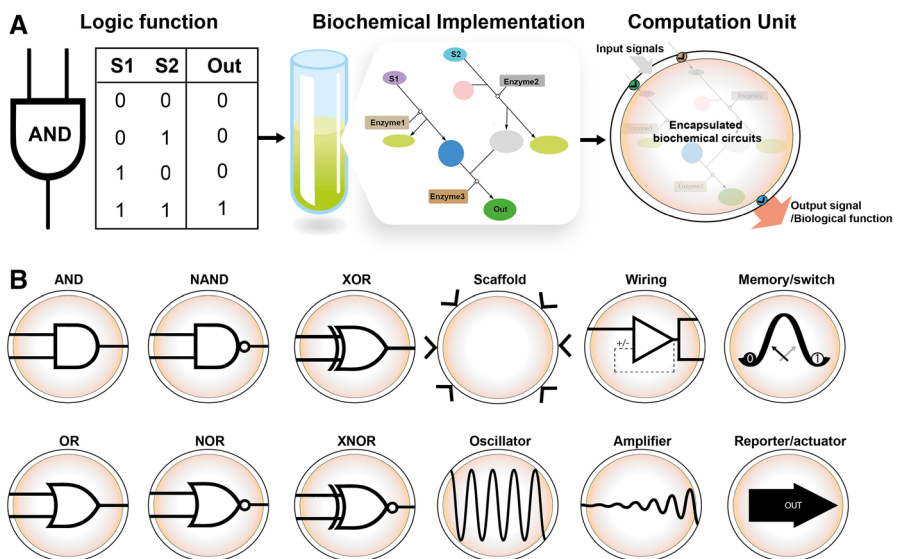


Fig. 1 a Rational design of a computation unit implementing a given logical function. (b) Different types of computation units. An AND gate outputs *true* only if the two inputs are *true*; An OR gate outputs *true* if at least one of the inputs is *true*; A XOR gate outputs *true* only when one of the inputs is *true*; The NAND, NOR and XNOR gates outputs the opposite value of the AND, OR and XOR gates respectively

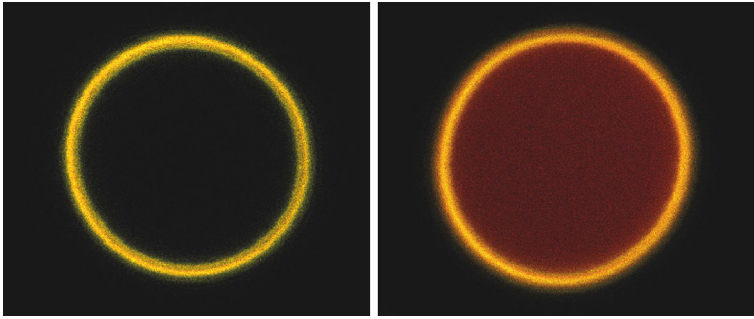


Fig. 2 Example of experimental fluorescence signal triggered in micrometric protocells. Ezymatic electron transfer from carbohydrate to the redox sensor probe (in that case resazurin is reduced into the *red fluorescent* product resorufin). Phospholipidic protocells encapsulating biochemical species were generated using microfluidic devices, and imaged using a confocal microscope. *Left* no induction; *right* induced with glucose. (Color figure online)

protocells to achieve complex information processing capabilities. In such architecture, input information arrives from upstream connections with previous protocells, to output connections to following computation units.

As each logic gate is encapsulated in an impermeable vesicle, the reactions that compute the output value will go from the non-equilibrium initial state to an equilibrium state. Therefore, once a logic gate has finished to compute the output, it is no more able to do another computation. So this first model of protocellular machine is in essence a kind of disposable computer.

Encapsulation of biochemical networks can be achieved using natural bilayer membranes (e.g. phospholipid bilayers, liposomes) (Noireaux and Libchaber 2004), or engineered membranes (e.g. copolymers, polymersomes) (Kamat et al. 2011), with respect to stoichiometry of internal species and incorporation of membrane proteins for connections (Chaize et al. 2004; Huang et al. 2014; Peters et al. 2014). This process is also known to stabilise enzymes, prevent cross-talk, denaturation or proteolysis and improve enzymatic properties (Yoshimoto 2011; Sunami et al. 2010). In addition, streamlined workflows, for example relying on microfluidics, are already available for the high-throughput generation of protocells that encapsulate various substrates (Richmond et al. 2011; Thiele et al. 2010; Duncanson et al. 2012; Matosevic and Paegel 2011; Teh et al. 2011). This strategy, extensively used in our lab, allowed us to test the implementation of various protocellular logic gates. Such vesicle have proven to be sufficiently stable (i.e. not prone to fuse together or physical disruption) to enable the construction of such multi vesicular assemblies (Stanish and Singh 2001; Teh et al. 2011). Tunable sizes ranging from 50 nm to 50 μm can be obtained, although in our approach, size should be kept as small as possible to obtain the highest density of computing operators.

2.2 Circuit Wiring

To obtain a full circuit implementing a given boolean function, we then need to concatenate and wire basic logic gates. The design of a function-specific

protocellular machine exploits the composability of *computation units*. Amongst a specific set of protocells, multiple instances of the same logic gates can be *wired* together to implement a user-defined function.

One way to achieve successive reactions in each layer of a protocellular machine, from input to output protocells, is to drive them using electrochemical potential (e.g. oxido-reduction reactions). By analogy with electronic computers, electrons are energy carriers and the redox potential is the *current* of the system, which could be measured with an electronic device. The major difference is that inside a protocell, wires are replaced by free molecules (e.g. NADH, NADPH, FADH₂), and effective wiring is achieved using chemical selectivity of enzymes. Molecules are either electron donors or acceptors, obeying biological enzymatic rules resulting in current and energy for computation. In such systems, the *in* → *out* direction is driven by the thermodynamics of the redox reaction. In our example, a protocell giving the *true* value would have a reductive state with high concentration of NADH, which can then transfer its electron to reduce the input of the next protocell. Conversely, a protocell giving the *false* value does not output any electron. In addition, electron transfer occurs only between physically connected protocells, through tight junctions putting into close contact electron transfer complexes, which carry out the connections between protocells and therefore between logic gates (Fig. 3).

We will build a protocellular machine from a set of protocell logic gates assembled in a tree-like layout (see Sect. 3). When set to *true*, the inputs of the machine initiate electron transfer through the chain of protocells that constitutes each branch of the tree, down to the root protocell.

In these input protocells, electron production is started by the specific oxidation of molecular species by oxidase enzymes. Electrons are then transferred down the protocell chain via transmembrane electron transport complexes that enable electron coupling (reduction) of specific molecular species. In that sense, input protocells can be seen as the *generators* that power the machine. Moreover *fuel* protocells, with a switch like behaviour, could be used to amplify and reshape the signal and therefore counteract its decay.

In order to implement specific electron transfer modules, we propose to exploit the modularity and thermodynamic reversibility of natural oxidative phosphorylation and photosynthesis complexes, which catalyse the electron transfer across natural membranes with specificity to NADH (Complex I), FADH₂ (complex II), and NADPH (NADPH quinone oxido reductase) (Osyczka et al. 2004). This includes quinone (or chemically related) and cytochrome c shuttle, which are delocalised mobile electron carriers that could be used as inter-protocell transfer molecules. In our design, we propose that a first quinone carrier (or related), could transfer electrons from a specific output signal (substrate specificity given by the first complex: I, II...) to a close complex III, which would then via a mobile cytochrome c transfer these electrons forward to the complex III belonging to the next protocell. This mechanism constitutes efficient reversible energy coupling, which has been shown to work via electron-tunneling across the proteins (Osyczka et al. 2004). Furthermore, recent studies have highlighted the possibility to re-engineer natural prokaryotic complexes for efficient and substrate specific synthetic

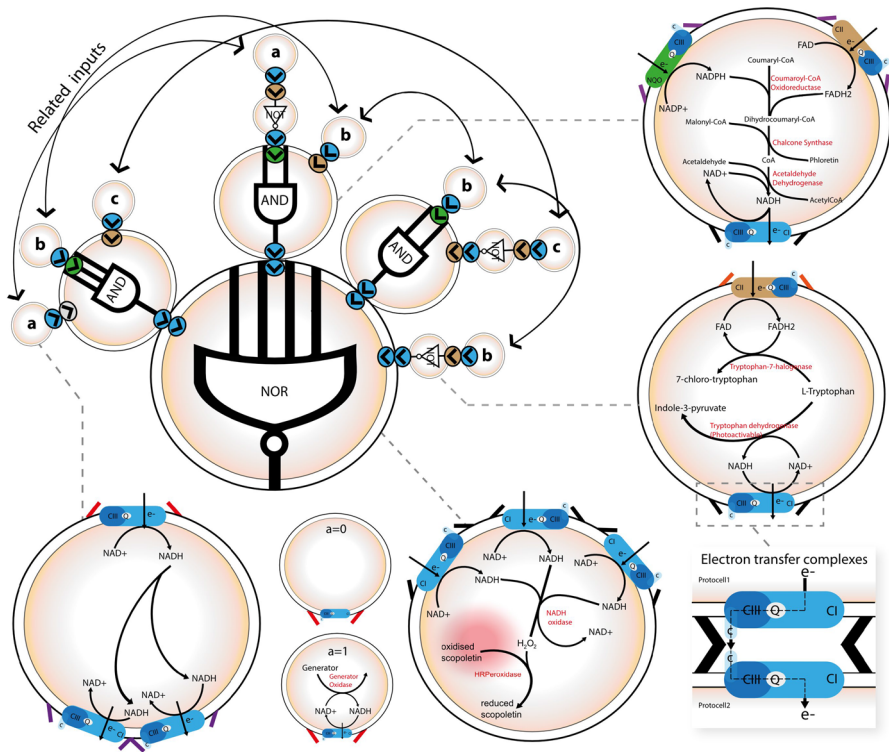


Fig. 3 Detail of a possible implementation of each type of protocell gate. Each type of logic gates has been simulated in silico with HSIM, and some of them are under test in the lab). The detail of the electron transfer mechanism is shown in the bottom right cartoon. For example, the fluorescent NOR gate uses a cascade of two enzymatic reactions (NADH oxidase, Horseradish Peroxidase) to consume the fluorescent oxidised scopoletin when NADH is present in the protocell, that is when at least one input is set to *true*, so is transferring electrons to make NADH from the initial pool of NAD^+

electron transporters (Katzen et al. 2002; Page et al. 1999; Wakeham and Jones 2005).

The architecture of a machine is controlled by the functional *wiring* of input and output of specific protocells. This can be achieved by using programmable junction modules, that can be selected to implement any protocellular machine in a *plug-and-play* way (Fig. 3). Biological function for these programmable attachments could be supported by couples of ligand/receptors with high binding affinity, such as aptameric nucleic acids (Hermann and Patel 2000; Smuc et al. 2013) or peptidic binders (Falciani et al. 2005), that could be straightforwardly produced in large combinatorial synthetic libraries using SELEX (Stoltenburg et al. 2007), or ribosome display respectively (Hanes et al. 2000; Binz et al. 2005).

Starting from a pre-built stock of computation units, the user can define a set of attachment instructions that corresponds to the boolean function to implement. Irreversible constructs can be achieved using cross-linking chemicals, so that no unbinding would occur (Song et al. 2012; Xiang et al. 2014). We assume that the

kinetics associated with such an assembly process would be of the order of minutes. Some attachments can also be set as *random*, to enable stochastic wiring of different types of protocells to specific positions. This could be used for example to solve problems involving the navigation through a large parameter space where protocellular machines could be used to compute a fitness function. Additionally selection methods could be implemented to isolate protocellular machines that exhibit specific behaviours. Positive selection can be done for example using FACS, conversely negative selection via a self-destruction mechanism.

3 The Case Study

3.1 Boolean Satisfiability Problem

The NP-complete problem we aim to solve is the 3-SAT problem. This problem can be simply defined as:

given any boolean formula in Conjunctive Normal Form (CNF), with at most 3 literals per clause, is there a valuation of the variables that satisfy the formula?

In other words, it asks whether the variables of a given boolean formula can be consistently replaced by the values *true* or *false* in such a way that the formula evaluates to *true*. If it is the case, the formula is called satisfiable. The literals are either a variable (v) or the negation of a variable ($\neg v$); They are connected with the *or* operator (\vee) to form a clause; The clauses are connected with the *and* operator (\wedge) to obtain the formula in CNF. For example:

$$F(a, b, c) = (a \vee \neg b \vee c) \wedge (b \vee \neg c) \wedge (a \vee b) \quad (1)$$

is true when $a = \text{true}$, $b = \text{true}$ and $c = \text{false}$, so the formula $F(a, b, c)$ is satisfiable. Conversely, the formula:

$$G(a, b, c) = (a \vee b \vee c) \wedge (\neg a \vee b) \wedge (b \vee \neg c) \wedge \neg b \quad (2)$$

is not satisfiable because all the eight possible valuations for a, b, c lead to $G = \text{false}$.

To find if a formula is satisfiable, we will build as many protocellular machines as there are combinations of valuations of the input variables. To do this, we will exploit the combinatorial power of ligand-receptor binding to link constant protocells (with *false* or *true* values) to the inputs of the protocellular machine to cover all the value space. A protocellular machine is dedicated to a specific formula, and therefore is not programmable in the sense an electronic computer is. The protocellular machines are self assembled according to the formula they have to check, so in our approach, the *programme* is the process that directs the assembly of the machines. We will ascertain that there is at least one instance of a protocellular machine per possible valuation of the variables.

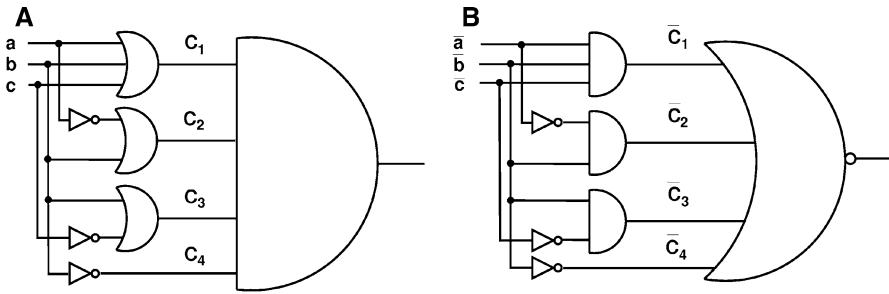


Fig. 4 **a** Direct implementation of the G formula in standard Conjunctive Normal Form. **b** Using the De Morgan laws, the same boolean function is rewritten using a NOR gate instead of the final AND gate, easier to build with a large number of inputs, and multiple 2- and 3-inputs AND gates fed with the complement of the original inputs

An instance of the machine can be made using 2- and 3-inputs OR gates connected to a big AND gate with as much inputs as there are clauses in the formula. Each input of a clause is connected to a protocell representing a variable v sending *true* or *false* when a specific start signal is given, or to an inverter protocell sending the negation of v when the start signal is given. The output of the AND gate is connected to a protocell that fluoresces when the input value is *true*. For example, the protocellular machine corresponding to the G formula would be made of a 4-input AND gate, two 2-input OR gates, one 3-input OR gate and three inverters connected as in the equation above (Fig. 4a).

As we have at least one (and probably more) instance of the machine for each possible valuation of the variables, if at least one of the protocellular machine fluoresces, the formula is satisfiable. Conversely, if there is no fluorescence at all then the formula is not satisfiable.

We can simplify the construction of the machines using the De Morgan laws to replace the big AND gate by a NOR gate, which is easier to build and also more efficient than an AND gate when there is a lot of inputs. Since the output of this NOR gate is the output of the whole machine, the final inverter can be made using an inhibitor of the fluorophores stored inside the protocell implementing the gate. We also need to feed the inputs of the AND gates with the complement of the variables, which could lead us to use a lot of inverters; But they can be avoided because these inputs are the inputs of the whole machine, and since we need to test all the valuations of the variables, these inputs will be fed with *constant* values. Therefore we can program the assembly of a machine with the constants already inverted (Fig. 4b) and we will need no more inverters than negated variables specified in the original formula.

3.2 The Assembly of the Machines

To obtain one instance of a computing protocellular machine, we need to direct the self assembly of as many copies of AND gate protocells as there are clauses in the formula (except when a clause has only one literal), the output of each AND gate

being connected to an input of a fluorescent NOR protocell. The inputs of each AND gate are also to be connected to the output of an inverter or to the output of a wiring protocell (representing the input variables of the formula). Then, to test a valuation of the variables of the formula, the input of each wiring protocell will be connected to special inputless protocells that output the constant value *true* or *false*. Once the machine and its inputs are assembled, when a *start signal* is given, after a few minutes, the NOR gate of this machine will fluoresce if the formula is *true* for this valuation of the variables, and therefore the formula is satisfiable.

We must ensure that *correlated* inputs of two (ore more) AND gates are fed with correlated values. In the previous formula (rewritten using a NOR of ANDs, with the complemented variables as shown in Fig. 4b)

$$G(a, b, c) = \overline{(\bar{a} \wedge \bar{b} \wedge \bar{c}) \vee (-\bar{a} \wedge \bar{b}) \vee (\bar{b} \wedge -\bar{c}) \vee -\bar{b}} \tag{3}$$

the first input of the first clause, \bar{a} , is always the opposite of the first input of the second clause ($-\bar{a}$), and the second input of the two first clauses, \bar{b} , have always the same value, etc. To achieve that we will use *inverter* protocells, and *wiring* protocells that can transfer their input to two or more outputs.

In this example, since there are 3 variables, we must assemble 8 protocellular machines to test each of the 8 possible valuations. Each line of the table in Table 1 shows the input values (0 for *false*, 1 for *true*) of one of the 8 different protocellular machines, the complemented value of each clause, and the value of the formula (3), which is always *false* (this formula is not satisfiable).

In order to have a efficient assembly mechanism, we split the process in two steps. The first one does not depend on a specific formula, but on the maximal numbers of variables (V_{max}) and of clauses (C_{max}) a formula can have. To be able to test any given formula within the limits of size we stated, we build a reservoir containing at most for one protocellular machine instance:

- one C_{max} -input NOR gate
- C_{max} 2- and 3-inputs AND gates.
- V_{max} *inverter* protocells
- $2 \cdot V_{max}$ types of inputless *constant protocells*, outputting the constant *false* or *true* to represent the two possible values of each variable.

Table 1 Complemented value of each clause for the eight possible valuations of the variables, and the corresponding value of the formula

\bar{a}	\bar{b}	\bar{c}	\bar{c}_1	\bar{c}_2	\bar{c}_3	\bar{c}_4	$G(a, b, c)$
0	0	0	0	0	0	1	0
0	0	1	0	0	0	1	0
0	1	0	0	1	1	0	0
0	1	1	0	1	0	0	0
1	0	0	0	0	0	1	0
1	0	1	0	0	0	1	0
1	1	0	0	0	1	0	0
1	1	1	1	0	0	0	0

- a formula dependent number of *wiring* protocells that duplicate their input to two (or more) outputs in order to cast each constant protocell output to the appropriate AND input or inverter.

Of course we can have a larger number of copies of these building blocks if we want to test more than one instance of the formula.

We can remark that depending on the formula we want to test, all the C_{\max} inputs of the NOR gate are not used and will stay not connected to any output, which is equivalent to a *false* value and so these inputs will not interfere with the computation since we are certain that nothing can be bound to them.

To verify the satisfiability of a formula made of $N < V_{\max}$ variables and $C < C_{\max}$ clauses, we need to build 2^N protocellular machine instances, (at least) one per possible valuation of the input variables. The building of these protocellular machines constitutes the second step. Although this step is specific to a given formula, its principle is generic enough to be applied to any formula. This resembles the *compilation* phase of a programme written in a high level programming language on a standard computer.

To assemble a machine we will *program* the binding of each input of one NOR gate to the output of a 2-inputs or a 3-inputs AND gate, or to one output of a wiring protocell, or to the output of an inverter. We will also need to program the binding of one wiring protocell per variable to some inverter, AND or NOR input, according to the formula. Then, to test a given valuation of the variables, we will need to bind the constant protocells corresponding to each variable of the formula to the inputs of this machine.

These programmed bindings are made possible because all the protocells in the reservoir have been built with specific tags on their inputs and outputs. These tags can be peptides/nucleic acids with a unique sequence to address them. The process of binding itself will be done by putting in the environment specific molecular attachment instructions that recognise and bind the tag on the output and the tag on the corresponding input. This will enable the binding of specific protocells together (Fig. 5).

Each input of the NOR gate is labeled with a tag implementing the number of the corresponding clause (0 to $C_{\max} - 1$). Similarly the output of each of the AND gate is labeled with the same number. Therefore, to connect an AND gate to the corresponding input of the NOR gate for one protocellular machine, we have to synthesise a molecular attachment that match at one end the tag labelling the output of the AND gate and at the other end, the tag labelling the input of the NOR gate.

The same mechanism is used for the input variables of the formula. The input of a wiring protocell that corresponds to a variable of the formula is labeled with a tag representing the variable number (0 to $V_{\max} - 1$). The constant protocells used for each variable, whether their output is *false* or *true*, are labeled with a tag matching the corresponding wiring protocell of the machine. Since there is a high number of constant protocells in the medium, the *false* and *true* version for each variable will be randomly bound to the corresponding input of the machines, and after some time, all the possible valuations will be covered.

It is important to notice that we must use constant protocells that output the boolean value *false*, even if a non-connected input is equivalent, because when we

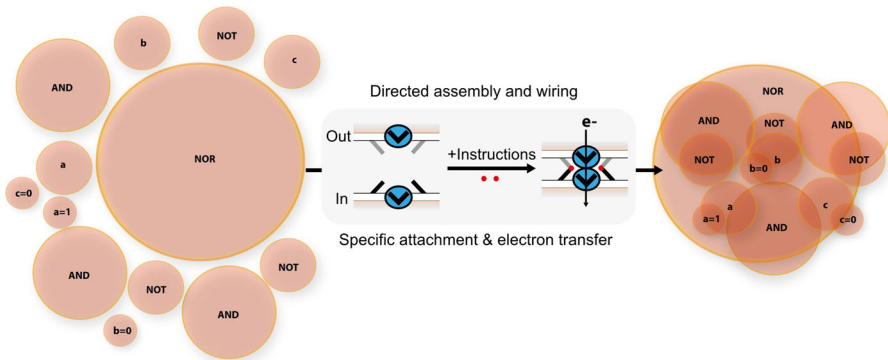


Fig. 5 Directed assembly and wiring via specific attachments of one instance of a protocellular machine for the formula $G(a, b, c)$. The inputs a , b and c are implemented with *wiring* protocells (one input, one, two or more outputs) that distributes the values of the variables to inverters or to the NAND gates according to the formula (3), see Fig. 4b. The NOR gate is a large protocell underneath the AND gates, where the outputs of the AND gates are bound. The small protocells $a, b, c = 0, 1$ are constant protocells for the input variables a, b and c (left). These input protocells will be randomly bound to constant *false* or *true* protocells to cover all the valuations of the variables. On the *right side*, the protocellular machine assembled tests the valuation $a = 1, b = 0, c = 0$

want to test a valuation where some variable is *false*, we must be certain that no *true* constant protocells can be bound to this input.

3.3 The Computation Process

The computation process may begin when we are certain that at least one copy of a protocellular machine is bound to each possible combination of input values.

This process is started by remotely triggering the whole population of *true* constant protocells and inverter protocells using, for example, light switchable enzymes (a tryptophan dehydrogenase engineered to bear a photoswitch moiety) (Strickland et al. 2012; Rakhit et al. 2014; Riggsbee and Deiters 2010).

Since all the machines run concurrently to compute the value of the formula, the total computing time is the time needed either by the first one that output *true* (that become fluorescent) or when we can be certain that the slowest machine that outputs *false* has finished (in this case they all do). If there is a small number of protocellular machines that fluoresces, we could enhance the signal/noise ratio by scattering the solution into several parts such that the concentration of the fluorescent machines would appear higher, and so helps its detection. Another way to easily detect the first (and possibly only) protocellular machine that outputs *true* would be that this machine triggers the fluorescence of those in its neighbourhood, and so increase the global fluorescence. Independently of the formula we want to test, the maximal number of reactions needed from one input to the output is very small: one inverter, a small number of wiring protocells, one AND, and one NOR.

Considering the kinetics of enzymatic processes for these simple reactions, we could assume that the calculation time of a single protocell (i.e. the time required for effective electron transfers through the protocell) would be in the order of a few minutes. The computation time for one protocellular machine would then be proportional to the number of layers of this machine. The total computing time would not exceed 20 min, whatever the number of protocellular machines is needed to solve the problem. This is of course mainly because the computing process is massively parallel and to a lesser extent because each *processor* is dedicated to the specific problem we want to solve.

Since the size of a complete protocellular machine is of the order of magnitude of a micron-cube, even less, we can have more than 10^{12} machines in a few ml of solution. As 10^3 is approximately equal to 2^{10} , we could theoretically have about $2^{10(12/3)} = 2^{40}$ machines in a few ml. Therefore using this technique, we could potentially solve any 3-SAT problem involving up to 40 variables in a few minutes. If we suppose that an electronic computer needs 1 μ s to generate and test one valuation of the variables, the average computing time would be of the order of $10^{12} \cdot 10^{-6} = 10^6$ s, which is more than 11 days and a half.

Moreover, if we suppose we use a low power electronic computer, for example 20 watts, the energy consumed at the end of the 11.5 days would be $10^6 \cdot 20 = 2 \cdot 10^7$ J (≈ 5.5 kWh), compared to a few joules for the protocellular machines.

4 Conclusion

The case studied here is an example of what we could do with protocellular machines, and how to make them. Of course, making the huge number of instances of protocellular machines needed to verify the satisfiability of a large formula is a bit speculative at the present day, but the mechanisms used to engineer their building blocks and to direct their assembly are already under test in the lab. Many implementations of logic gates (much more than those shown in Fig. 5) have been tested *in silico* using the HSIM (Amar et al. 2008) simulation system and proven to be functioning (Bouffard et al. 2015).

The computing time we claim, approximately one thousand times faster than a traditional electronic computer for a specific class and size of problem, is also a bit provocative, but the fact remains that this is an example of how to use the really massive parallelism of protocellular machines in order to solve dedicated problems. Moreover, to our knowledge, this is the first case where a synthetic biochemical computer could realistically compete with the speed of electronic computers, while being far less demanding in terms of energy.

Nevertheless, in our opinion, the most exciting perspective of protocellular machines is that they are electronically and biologically interfaceable. Thus they could be incorporated in living organisms, or into hybrid electronic/biological systems. Our approach allows us to design any given boolean function that can be

connected and triggered by any biological and/or electrical input, and generate chosen outputs in a similar way.

References

- Amar P, Legent G, Thellier M, Ripoll C, Bernot G, Nystrom T, Saier M Jr, Norris V (2008) A stochastic automaton shows how enzyme assemblies may contribute to metabolic efficiency. *BMC Syst Biol* 2:27
- Benenson Y (2012) Biomolecular computing systems: principles, progress and potential. *Nat Rev Genet* 13(7):455–468
- Binz HK, Amstutz P, Plückthun A (2005) Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol* 23(10):1257–1268
- Bouffard M, Molina F, Amar P (2015) Extracting logic gates from a metabolic network. In: P. Amar, F. Képès, V. Norris (eds.) Proceedings of the conference “advances in systems and synthetic biology”. EDP Sciences, Strasbourg, France, pp 63–76
- Candeias LP, MacFarlane DPS, McWhinnie SLW, Maidwell NL, Roeschlaub CA, Sammes PG, Rachel W (1998) The catalysed nadh reduction of resazurin to resorufin. *J Chem Soc Perkin Trans* 2:2333–2334
- Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* 26(7):787–793
- Chaize B, Colletier J-P, Winterhalter M, Fournier D (2004) Encapsulation of enzymes in liposomes: high encapsulation efficiency and control of substrate permeability. *Artif Cells Blood Substit Immobil Biotechnol* 32(1):67–75
- Chandran D, Bergmann FT, Sauro HM, Densmore D (2011) Computer-aided design for synthetic biology. In: Koepl H, Setti G, di Bernardo M, Densmore D (eds) Design and analysis of biomolecular circuits. Springer, New York, pp 203–224
- Cook SA (1971) The complexity of theorem-proving procedures. In: Proceedings of the third annual ACM symposium on theory of computing, STOC’71, New York, NY, USA, ACM, pp 151–158
- Duncanson WJ, Lin T, Abate AR, Seiffert S, Shah RK, Weitz DA (2012) Microfluidic synthesis of advanced microparticles for encapsulation and controlled release. *Lab Chip* 12(12):2135
- Elani Y, Law RV, Ces O (2014) Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat Commun* 5:5305
- Endy D (2005) Foundations for engineering biology. *Nature* 438(7067):449–453
- Falciani C, Lozzi L, Pini A, Bracci L (2005) Bioactive peptides from libraries. *Chem Biol* 12(4):417–426
- Hanes J, Schaffitzel C, Knappik A, Plückthun A (2000) Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat Biotechnol* 18(12):1287–1292
- Hermann T, Patel DJ (2000) Adaptive recognition by nucleic acid aptamers. *Science* 287(5454):820–825
- Huang X, Patil AJ, Li M, Mann S (2014) Design and construction of higher-order structure and function in proteinosome-based protocells. *J Am Chem Soc* 136(25):9225–9234
- Kamat NP, Katz JS, Hammer DA (2011) Engineering polymersome protocells. *J Phys Chem Lett* 2(13):1612–1623
- Karp RM (1972) Reducibility among combinatorial problems. In: Miller RE, Thatcher JW (eds) Complexity of computer computations. Plenum Press, New York, pp 85–103
- Katzen F, Deshmukh M, Daldal F, Beckwith J (2002) Evolutionary domain fusion expanded the substrate specificity of the transmembrane electron transporter DsbD. *EMBO J* 21(15):3960–3969
- Khalil AS, Collins JJ (2010) Synthetic biology: applications come of age. *Nat Rev Genet* 11(5):367–379
- Koepl H (2011) Design and analysis of bio-molecular circuits. Springer, Berlin
- Luisi PL, Stano P (2011) The minimal cell the biophysics of cell compartment and the origin of cell functionality. Springer, Berlin
- Marchisio MA, Stelling J (2009) Computational design tools for synthetic biology. *Curr Opin Biotechnol* 20(4):479–485
- Matosevic S, Paegel BM (2011) Stepwise synthesis of giant unilamellar vesicles on a microfluidic assembly line. *J Am Chem Soc* 133(9):2798–2800

- Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. *Proc Nat Acad Sci USA* 101(51):17669–17674
- Oszycza A, Moser CC, Daldal F, Leslie Dutton P (2004) Reversible redox energy coupling in electron transfer chains. *Nature* 427(6975):607–612
- Page CC, Moser CC, Chen X, Dutton PL (1999) Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* 402(6757):47–52
- Peters RJRW, Marguet M, Marais S, Fraaije MW, van Hest JCM, Lecommandoux S (2014) Cascade reactions in multicompartimentalized polymersomes. *Angew Chem Int Ed* 53(1):146–150
- Purnick PEM, Weiss R (2009) The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* 10(6):410–422
- Rakhit R, Navarro R, Wandless TJ (2014) Chemical biology strategies for posttranslational control of protein function. *Chem Biol* 21(9):1238–1252
- Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (2009) Protocells: bridging nonliving and living matter. MIT Press, Boston
- Rialle S, Felicori L, Dias-Lopes C, Peres S, Atia SE, Thierry AR, Amar P, Molina F (2010) BioNetCAD: design, simulation and experimental validation of synthetic biochemical networks. *Bioinformatics* 26(18):2298–2304
- Richmond DL, Schmid EM, Martens S, Stachowiak JC, Liska N, Fletcher DA (2011) Forming giant vesicles with controlled membrane composition, asymmetry, and contents. *Proc Nat Acad Sci* 108(23):9431–9436
- Riggsbee CW, Deiters A (2010) Recent advances in the photochemical control of protein function. *Trends Biotechnol* 28(9):468–475
- Sarpeshkar R (2010) Ultra low power bioelectronics: fundamentals, biomedical applications, and bio-inspired systems. Cambridge University Press, Cambridge
- Smaldon J, Romero-Campero FJ, Trillo FF, Gheorghe M, Alexander C, Krasnogor N (2010) A computational study of liposome logic: towards cellular computing from the bottom up. *Syst Synth Biol* 4(3):157–179
- Smuc T, Ahn I-Y, Ulrich H (2013) Nucleic acid aptamers as high affinity ligands in biotechnology and biosensors. *J Pharm Biomed Anal* 81–82:210–217
- Song S, Kole S, Bernier M (2012) A chemical cross-linking method for the analysis of binding partners of heat shock protein-90 in intact cells. *BioTechniques*. doi:[10.2144/000113856](https://doi.org/10.2144/000113856)
- Stanish I, Singh A (2001) Highly stable vesicles composed of a new chain-terminus acetylenic photopolymeric phospholipid. *Chem Phys Lipids* 112(2):99–108
- Stoltenburg R, Reinemann C, Strehlitz B (2007) SELEX-a (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol Eng* 24(4):381–403
- Strickland D, Lin Y, Wagner E, Hope CM, Zayner J, Antoniou C, Sosnick TR, Weiss EL, Glotzer M (2012) TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat Methods* 9(4):379–384
- Sunami T, Hosoda K, Suzuki H, Matsuura T, Yomo T (2010) Cellular compartment model for exploring the effect of the lipidic membrane on the kinetics of encapsulated biochemical reactions. *Langmuir* 26(11):8544–8551
- Teh S-Y, Khnouf R, Fan H, Lee AP (2011) Stable, biocompatible lipid vesicle generation by solvent extraction-based droplet microfluidics. *Biomicrofluidics* 5(4):044113
- Thiele J, Abate AR, Shum HC, Bachtler S, Förster S, Weitz DA (2010) Fabrication of polymersomes using double-emulsion templates in glass-coated stamped microfluidic devices. *Small* 6(16):1723–1727
- Wakeham MC, Jones MR (2005) Rewiring photosynthesis: engineering wrong-way electron transfer in the purple bacterial reaction centre. *Biochem Soc Trans* 33:851–857
- Xiang Z, Lacey VK, Ren H, Jing X, Burban DJ, Jennings PA, Wang L (2014) Proximity-enabled protein crosslinking through genetically encoding haloalkane unnatural amino acids. *Angew Chem Int Ed* 53(8):2190–2193
- Yoshimoto M (2011) Stabilization of enzymes through encapsulation in liposomes. In: Minteer SD (ed) *Enzyme stabilization and immobilization*. Humana Press, New York, pp 9–18