

Vishwesh V. Kulkarni · Guy-Bart Stan
Karthik Raman *Editors*

A Systems Theoretic Approach to Systems and Synthetic Biology II: Analysis and Design of Cellular Systems

 Springer

A Systems Theoretic Approach to Systems
and Synthetic Biology II: Analysis and Design
of Cellular Systems

Vishwesh V. Kulkarni
Guy-Bart Stan · Karthik Raman
Editors

A Systems Theoretic
Approach to Systems
and Synthetic Biology II:
Analysis and Design
of Cellular Systems

 Springer

Editors

Vishwesh V. Kulkarni
Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN
USA

Guy-Bart Stan
Department of Bioengineering
Imperial College
London
UK

Karthik Raman
Department of Biotechnology
Bhupat and Jyoti Mehta School
of Biosciences
Indian Institute of Technology Madras
Chennai
India

ISBN 978-94-017-9046-8 ISBN 978-94-017-9047-5 (eBook)
DOI 10.1007/978-94-017-9047-5
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2014942544

© Springer Science+Business Media Dordrecht 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law. The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

श्रेयो हि ज्ञानमभ्यासाज्ज्ञानाद्भ्यानं विशिष्यते ।
ध्यानात् कर्मफलत्यागस्त्यागाच्छान्तिरनन्तरम् ॥ १२-१२ ॥
– श्रीमद्भगवद्गीता

*Understanding is superior to mere practice
Union with the subject matter supersedes that
Dispassion towards all results is better still
And manifests peace immediately*

Bhagavat Gita (12:12)

*To my father Vasant, brother Vinay, sister
Ketki, and Prof. Peter Falb
To my mother in memory*

Vishwesh V. Kulkarni

*To my parents, Florentina and Stephan
To my wife Cristina and my daughter
Eva-Victoria*

Guy-Bart Stan

*To my parents and teachers
In memory of Sunder Mama
and Prof. E. V. Krishnamurthy*

Karthik Raman

Foreword

There is no design template more versatile than DNA. Nor are any designs more consequential than those whose blueprints DNA encodes. This exquisite substance has been shaped over billions of years by the creative combination of mutation and selection. Yet in the very long history of this template, it is only during our times that complex living organisms are beginning to understand and manipulate the very template whose sequences define them. But how should we go about this understanding? And how can we use this understanding to more effectively and responsibly alter the DNA template?

The complexity and diversity of living organisms are daunting. *Systems biology* aims at reverse engineering biological complexity for the purpose of understanding their design principles. By measuring and characterizing interactions of key biological molecules in response to stimuli and perturbations, systems biology aims to construct models that capture the complexity of endogenous biological networks. Through the systematic understanding of such models, it is hoped that one will achieve a holistic understanding of biological networks and the way they achieve biological function.

At the same time, the versatility of DNA and the dramatic decrease in the cost of DNA synthesis is making it possible to economically design and test new complex genetic circuits. This has given impetus to a new field: *Synthetic biology*. In our quest to understand biological complexity, we have examined endogenous biological subsystems and ascribed functions and design principles to their components. But a true understanding of these biological design principles is demonstrated only when one can build such systems *de novo* and demonstrate their function. When these circuits do not exhibit behavior consistent with our models, further investigations will lead to a deeper understanding of the underlying biology. Synthetic biology, therefore, serves as an important testbed for our understanding of biological principles. But the promise of synthetic biology extends beyond scientific understanding. Whether it be the detection and interference with the course of disease through the introduction of designer circuits, the cost-effective synthesis of new bio-substances, or the development of improved food products, synthetic biology provides a tremendous opportunity to alleviate suffering and improve the quality of our lives.

In both systems and synthetic biology, challenges abound. Quantitative modeling, analysis, and design of biological networks must contend with difficulties arising from the inescapable fact that at its most basic level, biology involves complex dynamic interactions among nonlinear stochastic components, taking place at multiple temporal and spatial timescales. The complexity of network interconnections of such components and the crosstalk between them adds another level of difficulty.

System theory has emerged as a field to deal with the challenges and complexities emerging from the interconnection of engineered systems, many of which are shared with biological systems. Notions from system theory such as nonlinearity, stochasticity, feedback, loading, modularity, robustness, identifiability, etc., are needed for a deeper understanding of biological complexity and for a more reliable design of biological circuits. These concepts are now being utilized to help us expand our understanding of endogenous biological circuits and to design novel ones. The articles in this book make significant strides in this direction.

While system theory will undoubtedly aid our understanding and design of biological systems, there is no doubt that the study of biological designs that have evolved over billions of years will also shape the future of system theory. For example, evolution and development are two central themes in biology that have little analogy with engineered man-made systems. Through the study of these and other biological themes, new systems notions and insights will undoubtedly emerge, enriching system theory in the process. One need only look at the history of *feedback*, a predominant concept in system theory, to imagine what is possible. While its human discovery can be traced back a little over one millennium, it is likely that feedback was invented by nature more than three billion years earlier. Since then, it has been wildly successful as a biological design principle, as evidenced by its prevalence at every level of biological organization. One wonders if an early systematic understanding of this concept in its biological context could have sped up the course of our own technological development.

As the physical sciences helped us understand the physical world around us over the last few centuries, so will quantitative biological science help us understand who we are, how we function, and how we can effectively and responsibly synthesize this most consequential of substances, the DNA. I believe that system theory will be central to this understanding.

Zürich, September 2013

Mustafa Khammash

Preface

Underlying every living cell are billions of molecules interacting in a beautifully concerted network of pathways such as metabolic, signalling, and regulatory pathways. The complexity of such biological systems has intrigued scientists from many disciplines and has given birth to the highly influential field of *systems biology* wherein a wide array of mathematical techniques, such as flux balance analysis, and technology platforms, such as next generation sequencing, is used to understand, elucidate, and predict the functions of complex biological systems. This field traces its roots to the general systems theory of Ludwig von Bertalanffy and effectively started in 1952 with a mathematical model of the neuronal action potential for which Alan Hodgkin and Andrew Huxley received the Nobel Prize in 1963. More recently, the field of *synthetic biology*, i.e., *de novo* engineering of biological systems, has emerged. Here, the phrase ‘biological system’ can assume a vast spectrum of meanings: DNA, protein, genome, cell, cell population, tissue, organ, ecosystem, and so on. Scientists from various fields are focusing on how to render this *de novo* engineering process more predictable, reliable, scalable, affordable, and easy. Systems biology and synthetic biology are essentially two facets of the same entity. As was the case with electronics research in the 1950s, a large part of synthetic biology research, such as the *BioFab* project, has focused on reusable macromolecular “parts” and their standardization so that composability can be guaranteed. Recent breakthroughs in DNA synthesis and sequencing combined with newly acquired means to synthesize plasmids and genomes have enabled major advances in science and engineering and marked the true beginning of the era of synthetic biology. Significant industrial investments are already underway. For example, in 2009, Exxon Mobil set up a collaboration worth \$600 million with Synthetic Genomics to develop next generation biofuels.

Recent advances in systems and synthetic biology clearly demonstrate the benefits of a rigorous and systematic approach rooted in the principles of systems and control theory—not only does it lead to exciting insights and discoveries but it also reduces the inordinately lengthy trial-and-error process of wet-lab experimentation, thereby facilitating significant savings in human and financial resources. So far, state-of-the-art systems and control-theory-inspired results in systems and synthetic biology have been scattered across various books and journals from various disciplines. Hence, we felt the need for an edited book that provides a

panoramic view and illustrates the potential of such systematic and rigorous mathematical methods in systems and synthetic biology.

Systems and control theory is a branch of engineering and applied sciences that rigorously deals with the complexities and uncertainties of interconnected systems with the objective of characterising fundamental systemic properties such as stability, robustness, communication capacity, and other performance metrics. Systems and control theory also strives to offer concepts and methods that facilitate the design of systems with rigorous guarantees on these fundamental properties. For more than 100 years, the insights and techniques provided by systems and control theory have enabled outstanding technological contributions in diverse fields such as aerospace, telecommunication, storage, automotive, power systems, and others. Notable examples include Lyapunov's theorems, Bellman's theory of dynamic programming, Kalman's filter, H^∞ control theory, Nyquist-Shannon sampling theorem, Pontryagin's minimum principle, and Bode's sensitivity integral. Can systems and control theory have, or evolve to have, a similar impact in biology? The chapters in this book demonstrate that, indeed, systems and control theoretic concepts and techniques can be useful in our quest to understand how biological systems function and/or how they can be (re-)designed from the bottom up to yield new biological systems that have rigorously characterized robustness and performance properties.

Several barriers must be overcome to contribute significantly in this exciting journey. One of these is the language barrier, e.g., what a systems theorist means by the word *sensitivity* is different from what a biologist means by it. Another one is the knowledge barrier as, traditionally, systems and control theorists and biologists are not well versed with each other's knowledge base (although that scenario is now fast changing for the better with the introduction of bioengineering courses in systems and control theory at the undergraduate and graduate levels). A third barrier is due to the sheer volume of *big data*: the European Bioinformatics Institute in Hixton, UK, which is one of the world's largest biological data repositories, currently stores 20 petabytes of data and backups about genes, proteins and small molecules, and this number is more than doubling every year. Finally, a fourth barrier comes from the effort required to produce timely contributions based on currently available models. As an example of this last barrier, the systems and control theory community could have played a greater role than it did in two of the most significant technological advances of the last 50 years: VLSI and Internet. In retrospect, besides the fact that the systems and control theorists caught on to the Internet too late, by which time infrastructures based on TCP/IP were already in place, the main difficulty posed by the Internet for the systems and control theory community was a lack of *good* models of the underlying networked system. This lack-of-good-models barrier is even more daunting in biology since some of the currently available *big data* are not guaranteed to be reproducible. As Prof. M. Vidyasagar illustrates and observes in the September 2012 issue of IEEE Lifesciences, one of the major challenges to the application of systems and control

theory concepts in biology comes from “the fact that many biological experiments are not fully repeatable, and thus the resulting datasets are not readily amenable to the application of methods that people like us [i.e., systems and control theorists] take for granted.”

The chapters in this book serve to propose ways to overcome such barriers and to illustrate that biologists as well as systems and control theorists can make deep and timely contributions in life sciences by collaborating with each other to solve important questions such as how to devise experiments to obtain models of biological systems, how to obtain predictive models using information extracted from experimental data, how to choose components for (re-)engineering biological networks, how to adequately interconnect biological systems, and so on. Furthermore, and as Prof. Mustafa Khammash observes in his foreword, this research will fundamentally enrich systems and control theory as well by forcing it to investigate currently open questions that are specific to living biological systems, e.g., Why do biological systems naturally evolve the way they do? Can the evolvability of biological systems be consciously exploited for (re-)design and optimization purposes?

This book is intended for (1) systems and control theorists interested in molecular and cellular biology, and (2) biologists interested in rigorous modelling, analysis, and control of biological systems. We believe that research at the intersection of these disciplines will foster exciting discoveries and will stimulate mutually beneficial developments in systems and control theory and systems and synthetic biology.

The book consists of 12 chapters contributed by leading researchers from the fields of systems and control theory, systems biology, synthetic biology, and computer science. Chapters 1–6 highlight some state-of-the-art methods used to address currently open questions in systems biology. Chapters 7–12 discuss frameworks and methods required to enable a bottom-up design of synthetic biology systems of increasing complexity. These chapters are organized into two main parts as follows.

- **Part I—Systems Biology:** Chapters 1–6 focus on specific problems in modelling biological systems. Examples of such problems include: characterization and synthesis of memory, understanding how homeostasis is maintained in the face of shocks and relatively gradual perturbations, understanding the functioning and robustness of biological clocks such as those at the core of circadian rhythms, and understanding how the cell cycles can be regulated, among others. A brief summary of each chapter is as follows.
 - Chapter 1: Today, several approaches used to identify biomarkers for a specific disease rely on genome-wide gene expression profiles without an explicit regard for how the genes are correlated. Wang and Chen present a network biomarker construction scheme that integrates microarray gene expression profiles and protein–protein interaction information so as to enable molecular investigation and diagnosis of lung cancer.

- Chapter 2: In this overview chapter, Lal and Seshashayee discuss next-generation sequencing techniques and illustrate how these have been used, at the scale of the whole bacterial genome, to investigate a variety of problems, from the analysis of gene expression and protein–DNA interactions to that of bacterial community function and evolution.
 - Chapter 3: Given a biological network of oscillators, such as circadian rhythm networks for instance, how do biological parameter variations affect the oscillation characteristics? Sacré and Sepulchre present a novel and scalable approach to characterize the parameter sensitivity of models of oscillators, and illustrate its use on a circadian rhythm network model.
 - Chapter 4: Osmosis facilitates the basic mechanism by which water is transported into and out of cells. Montefusco et al. demonstrate how a control theoretic analysis of the osmosis regulation system of *Saccharomyces cerevisiae* can be used to explain how cells maintain homeostasis in the face of osmotic perturbations.
 - Chapter 5: State synchronization is a recurring theme in neuronal networks and coupled networks of genetic clocks, among others. Hamadeh et al. explain how incremental dissipativity theory can be used to systematically analyse and/or synthesize feedback interconnections that ensure state synchronization in networks of identical oscillators and illustrate its use in the context of realising synchronization in a genetic repressilator network.
 - Chapter 6: Multistability is a key property of biological systems that characterizes salient phenotypes such as memory. Salerno et al. present a systematic approach to characterize bistability and explain its utility in characterizing the memory of the galactose regulatory system of *Saccharomyces cerevisiae*.
- **Part II—Synthetic Biology:** Chapters 7–12 focus on how biomacromolecules, platforms, and scalable architectures should be chosen and synthesized in order to build programmable *de novo* biological systems. For example, a standardization of the components used is a necessary step in the modular design of large scale systems and presents an opportunity to develop *in silico* design tools that optimize these systems with respect to a set of formal specifications. What are the types of constrained optimization problems encountered in this process and how can these be solved efficiently? Should DNA be used as the basic macromolecule in synthesising artificial biological networks or should it be used with other macromolecules to enable certain applications? This set of chapters aims at answering such questions. A brief summary of each chapter is as follows.
 - Chapter 7: Modern nucleic acid biochemistry extensively uses protein enzymes to manipulate nucleic acids. However, predictive modification of the behavior of protein enzymes remains a very difficult problem. Chandran et al. show how meta-biochemical systems offer the possible advantage of being far easier in terms of re-engineering and programming. They show how a biochemical system can be synthesized based entirely on strands of DNA as the

- only component molecule. These *meta-DNAs* have the same pairing mechanism as DNA but have a much larger alphabet of bases, thereby providing an increased power of base addressing.
- Chapter 8: An open challenge today is to specify synthetic biological systems using high level languages. Chen and Cai choose a rule-based modelling framework that was originally developed for systems biology and extend it to synthetic biology. They introduce a new model-specification language that facilitates the swift generation of mathematical models that encode the phenotypic behaviors of biological systems.
 - Chapter 9: Krishnan and Liu address how bistable and monostable switches give rise to irreversible transitions and decision making in cell cycles. They propose a modular framework to address such questions for binary signalling mechanisms, outlining some of the design principles of signalling networks, which can be exploited in synthetic biology.
 - Chapter 10: In standardising the components for a scalable design, an important constrained optimization problem concerns the selection of kinetic parameters and protein abundances. Koepl et al. explain how this inverse problem can be solved more elegantly, by linearising the forward operator that maps parameter sets to specifications, and then inverting it locally, rather than relying on a brute force random sampling approach.
 - Chapter 11: Marchisio and Stelling demonstrate how concepts and algorithms from electrical engineering can be exploited to set up a framework for the computation-based automated design of genetic Boolean gates and devices. They also explain how the Karnaugh algorithm used in the design of electrical circuits can be modified when it comes to the design of genetic circuits.
 - Chapter 12: Kim and Franco focus on how to synthesize and couple transcriptional circuits by exploiting the modular architecture of nucleic acid templates as well as the catalytic power of natural enzymes. They illustrate the programmability of dynamic behaviors for elementary circuits such as adapters, bistable switches, and oscillators. They also present insulating and amplifying devices as a solution for the scaling-up of biomolecular networks.

The burgeoning fields of systems biology and synthetic biology have thrown up a very large number of interesting research problems. As the pre-eminent computer scientist Donald Knuth put it, “biology easily has 500 years of exciting problems to work on.” The chapters in this book address but a small fraction of these interesting challenges. Nevertheless, we believe this book can serve as a good introduction on some of the currently open problems and on some of the state-of-the-art concepts and techniques available to propose solutions to such problems.

We are very grateful to all authors for their invaluable time and contributions and to Prof. Mustafa Khammash (ETH Zürich) for his stimulating foreword. We are also grateful to our institutions: University of Minnesota (Minneapolis, USA), Imperial College (London, UK), and Indian Institute of Technology Madras (Chennai, India) for their support and for providing a stimulating work

environment. Finally, we thank and acknowledge the financial support of our respective funding agencies: the National Science Foundation, the UK Engineering and Physical Sciences Research Council, and the Ministry of Human Resource and Development of the Government of India.

Minneapolis, MN, USA, September 2013
London, UK
Chennai, India

Vishwesh V. Kulkarni
Guy-Bart Stan
Karthik Raman

Contents

Part I Systems Biology

- 1 Network Biomarker Construction for Molecular Investigation and Diagnosis of Lung Cancer via Microarray Data** 3
Yu-Chao Wang and Bor-Sen Chen
- 2 The Impact of Next-Generation Sequencing Technology on Bacterial Genomics** 31
Avantika Lal and Aswin Sai Narain Seshasayee
- 3 Sensitivity Analysis of Circadian Entrainment in the Space of Phase Response Curves.** 59
Pierre Sacré and Rodolphe Sepulchre
- 4 Modelling and Analysis of Feedback Control Mechanisms Underlying Osmoregulation in Yeast.** 83
Francesco Montefusco, Ozgur E. Akman, Orkun S. Soyer and Declan G. Bates
- 5 Analysis of Synchronizing Biochemical Networks via Incremental Dissipativity.** 117
Abdullah Hamadeh, Jorge Gonçalves and Guy-Bart Stan
- 6 Robustness Model Validation of Bistability in Biomolecular Systems** 141
Luca Salerno, Carlo Cosentino, Alessio Merola, Declan G. Bates and Francesco Amato

Part II Synthetic Biology

- 7 Meta-DNA: A DNA-Based Approach to Synthetic Biology** 171
Harish Chandran, Nikhil Gopalkrishnan, Bernard Yurke and John Reif

8	Towards Modeling Automation for Synthetic Biology	201
	Chen Liao and Yizhi Cai	
9	An Investigation of Signal Transduction and Irreversible Decision Making Through Monostable and Bistable Switches. . . .	219
	J. Krishnan and C. Liu	
10	From Specification to Parameters: A Linearization Approach . . .	245
	Heinz Koepl, Marc Hafner and James Lu	
11	Simplified Computational Design of Digital Synthetic Gene Circuits	257
	Mario Andrea Marchisio and Jörg Stelling	
12	Synthetic Biochemical Devices for Programmable Dynamic Behavior	273
	Jongmin Kim and Elisa Franco	
Index	297

Part I
Systems Biology

Chapter 1

Network Biomarker Construction for Molecular Investigation and Diagnosis of Lung Cancer via Microarray Data

Yu-Chao Wang and Bor-Sen Chen

Abstract Lung cancer is the leading cause of cancer deaths worldwide. Many studies have investigated the carcinogenic process and identified the biomarkers for signature classification. However, those biomarkers are mainly identified based only on analysis of genome-wide expression profiles, that is, the identification method cannot elucidate how the different genes in the biomarker gene set are related to each other. Therefore, from the systems perspective, we developed a network biomarker construction scheme, which integrated microarray gene expression profiles and protein-protein interaction information, for molecular investigation and diagnosis of lung cancer. The network biomarker consisted of two protein association networks constructed for cancer samples and non-cancer samples. Based on the network biomarker, a total of 40 significant proteins were identified with carcinogenesis relevance values (CRVs) to gain insights into the lung carcinogenesis mechanism. In addition, the network biomarker was also acted as the diagnostic tool, demonstrated to be effective to diagnose the smokers with lung cancer. Taken together, the network biomarker not only successfully sheds light on the mechanisms in lung carcinogenic process but also provides potential therapeutic targets to combat against cancer.

Keywords Network biomarker · Lung cancer · Protein association network · Microarray data · Analysis of variance (ANOVA) · Protein–protein interaction (PPI) · Akaike’s information criterion

B.-S. Chen (✉)

Laboratory of Control and Systems Biology, Department of Electrical Engineering,
National Tsing Hua University, Hsinchu 30013, Taiwan
e-mail: bschen@ee.nthu.edu.tw

Y.-C. Wang

Institute of Biomedical Informatics,
National Yang-Ming University, Taipei 11221, Taiwan
e-mail: yuchao@ym.edu.tw

1.1 Introduction

Cancer, the complex disease of uncontrolled cell growth, is the leading cause of human death worldwide and the deaths from cancer are projected to continue rising [19, 49]. Among all types of cancer, the mostly diagnosed and the most common cause of cancer deaths are lung cancer and the mortality rate within 5 years is as high as 80–85 % [49, 59]. Lung cancer can be divided into two main types, small cell lung carcinoma (SCLC) and non-small cell lung carcinoma (NSCLC). NSCLCs are further divided into three main subtypes: squamous cell carcinoma, adenocarcinoma, and large cell carcinoma [6]. Previous study has shown that all these major histological types of lung cancer are associated with cigarette smoking [33]. Therefore, many researchers devoted themselves to investigate the molecular alterations resulted from cigarette smoking and the mechanism that links cigarette smoking and lung cancer. Spira et al. used DNA microarray to compare the gene expressions of large-airway epithelial cells from nonsmokers and smokers and to define how cigarette smoking alters the transcriptome [58]. Hecht indicated that many tobacco smoke carcinogens such as polycyclic aromatic hydrocarbons and nicotine-derived nitrosamine ketone are likely to play major roles in lung cancer induction [23]. Recently, Takahashi et al. showed that induction of IKK β - and JNK1-dependent inflammation is likely to be an important contributor to the tumor-promoting activity of tobacco smoke [61].

In addition to the investigation of carcinogenesis, many studies identified the cancer biomarkers through analysis of genome-wide expression profiles [2, 21]. The biomarkers are used as diagnostic evaluation to determine the patient with or without the cancer or used as prognostic indicator to evaluate the patient's prognosis. In lung cancer, Spira et al. used gene expression profile from patient samples to identify an 80-gene biomarker that distinguishes smokers with and without lung cancer [59]. The 80-gene biomarker could be beneficial for the decrease of the high mortality rate since the poor prognosis of lung cancer is closely related to the fact that there is no effective screening tool to diagnose the disease at an early stage [26, 59]. However, the biomarker identification method based only on gene expression profiles cannot elucidate how the different genes in the biomarker gene set are related to each other, i.e., the biomarkers are not identified from the systems perspective. Further, the gene lists obtained for the same clinical types of patients by different groups differ widely and have only very few genes in common [17].

Due to these kinds of limitation and the widely accepted opinion that cancer is a disease of pathways [22, 68], protein-protein interaction (PPI) and pathway information are integrated for biomarker identification. Chuang et al. developed a protein-network-based approach that identifies biomarkers not as individual genes but as subnetworks extracted from protein interaction databases. They showed that the subnetwork classification could achieve higher accuracy in the signature discrimination and are informative of the network structure [11]. Many other network-based approaches were also developed for prioritizing disease genes and protein interaction subnetworks that are discriminative of disease signature [9, 45, 46, 65]. In addition, the dynamic structure of the human protein interaction network was examined to

predict breast cancer prognosis, suggesting that the network modularity might be a defining feature of tumor phenotype [62].

Network analysis has shown that under different cellular states or in response to diverse stimuli, transcription factors alter their interactions to regulate different genes, thereby rewiring the network [41]. The same situation happens for protein interaction networks [62, 76]. Motivated by the dynamic structure of human protein interaction network and the observation that interacting proteins tend to result in similar disease phenotypes when dysregulated [47], we develop a computational framework to construct the network biomarker for molecular investigation and diagnosis of lung cancer via microarray data. The network biomarker consists of two protein association networks for cancer and non-cancer smokers. Based on the concept of network comparison [72], 40 significant proteins that may play important roles in lung carcinogenesis are identified. With the help of the network biomarker, the smokers suspect with cancer can be classified into smokers with cancer or without cancer, making the network biomarker a useful tool for molecular diagnosis. Hopefully, the proposed method can help understand the lung carcinogenesis and provide potential drug targets for humans to combat against lung cancer.

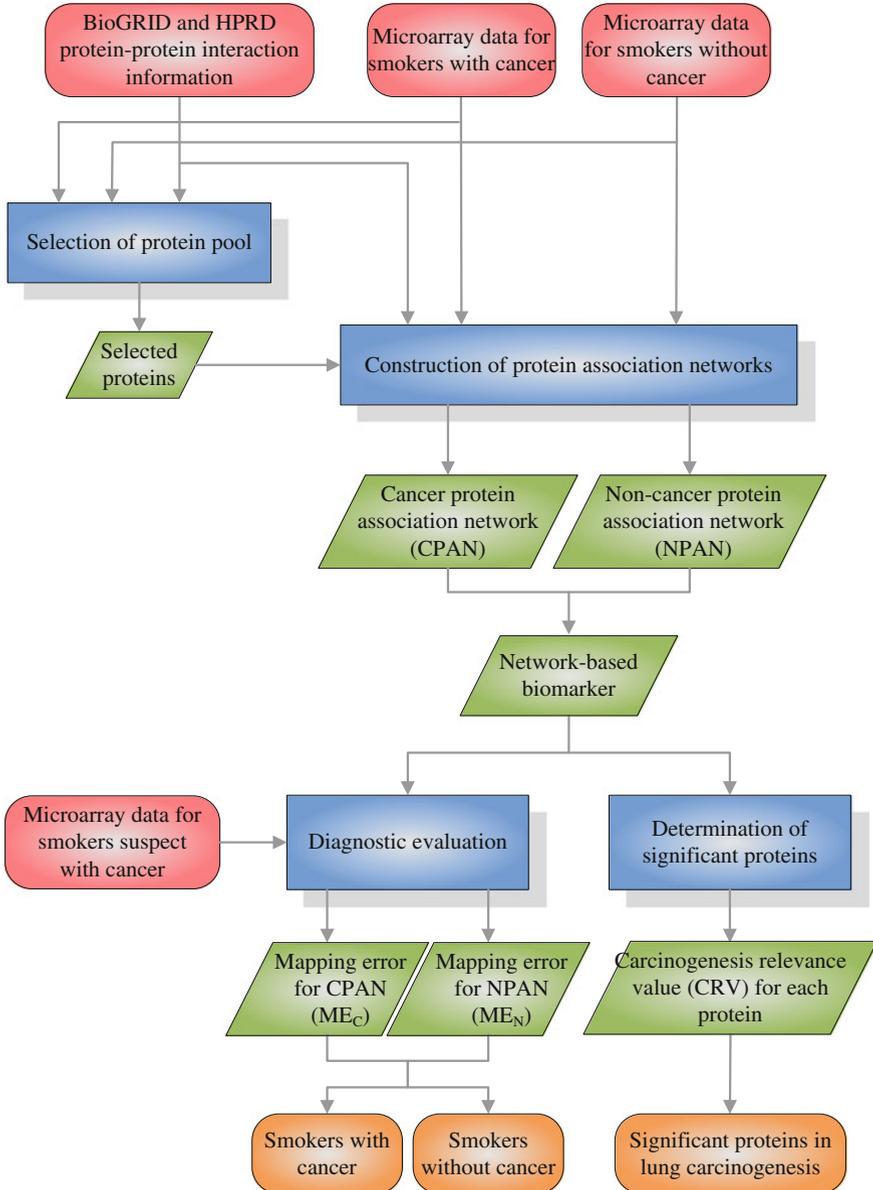
1.2 Methods

1.2.1 Overview of the Network Biomarker Approach for Lung Cancer Investigation

The overall flowchart of the proposed network biomarker approach is shown in Fig. 1.1. Our goal is to investigate the lung cancer by the construction of network biomarker which is composed of protein association networks for smokers with and without cancer. Microarray gene expression profiles of patient samples and protein-protein interaction information were integrated for protein selection and network construction. Two protein association networks with quantitative protein association abilities for cancer and non-cancer smokers were constructed respectively. Based on the comparison of two protein association networks within the network biomarker, a score named carcinogenesis relevance value (CRV) was computed to correlate proteins with significance of lung carcinogenesis. A higher score suggests that the particular protein plays a more critical role in lung carcinogenesis. According to the CRV for each protein and the statistical assessment, a set of significant proteins was selected. Furthermore, given the microarray data for the smokers suspect with cancer, mapping errors can be computed for diagnostic evaluation that the smokers with or without cancer.

1.2.2 Data Selection and Preprocessing

Here, two kinds of data, microarray gene expression profile and protein-protein interaction information, were integrated. The microarray data was



◀ **Fig 1.1** The flowchart of constructing the network biomarker for lung cancer investigation and diagnosis. The figure indicates the flowchart of the proposed method. The *red rounded rectangles* represent the data needed. The *blue rectangles* denote the processing steps of the approach. The *green parallelograms* are the processed results for each step and the *orange rounded rectangles* are the overall results for the whole method. In summary, two kinds of data, microarray data and PPI information, are needed for the proposed method. These data are used for protein pool selection. Then the selected proteins and the input data are used for protein association network construction, resulting in cancer protein association network (CPAN) and non-cancer protein association network (NPAN). The two constructed protein association networks constitute the network biomarker, which can be used for either determination of significant proteins or diagnostic evaluation. With the help of the network biomarker, carcinogenesis relevance value (CRV) is computed for each protein and significant proteins in lung carcinogenesis are determined based on the CRVs. These significant proteins provide targets for further characterization. On the other hand, given the microarray data for smokers suspect with cancer, mapping errors for CPAN and NPAN can be computed, respectively, which help diagnose the smokers with cancer or without cancer.

downloaded from GEO database <http://www.ncbi.nlm.nih.gov/geo/> (accession number GSE4115). Spira et al. performed gene expression profiling in histologically normal large-airway epithelial cells obtained at bronchoscopy from current and former smokers. Each individual was followed after bronchoscopy until a final diagnosis of lung cancer or not lung cancer was made [59]. Data was collected from a total of 187 subjects and was divided into primary and prospective data sets (79 smokers with lung cancer and 73 smokers without lung cancer in the primary data set; 18 smokers with lung cancer and 17 smokers without lung cancer in the prospective data set). The primary data set was used for network biomarker construction and the prospective data set was used for diagnostic evaluation. Protein-protein interaction (PPI) data was extracted from BioGRID <http://thebiogrid.org/> and HPRD <http://www.hprd.org/databases>. The Biological General Repository for Interaction Datasets (BioGRID) database was developed to house and distribute collections of protein and genetic interactions from major model organism species. BioGRID currently contains over 340,000 interactions as derived from both high-throughput studies and conventional focused studies [60]. The Human Protein Reference Database (HPRD) is a database that integrates a wealth of information relevant to human proteome, including protein-protein interactions, post-translational modifications, disease associations, and tissue expression [50]. Prior to further processing, the gene expression value g_{ij} is normalized to z-transformed scores z_{ij} so that for each gene i the normalized expression value has mean $\mu_i = 0$ and standard deviation $\sigma_i = 1$ over sample j .

1.2.3 Selection of Protein Pool and Construction of Network Biomarker

To integrate the gene expression and PPI information data and construct the network biomarker consisting of protein association networks, the expression value of each gene was first overlaid on its corresponding protein. The gene expression for

each protein was then used to select differentially expressed protein using one-way analysis of variance (ANOVA) where the null hypothesis is that the average expression levels for the protein are the same for smokers with and without cancer [48]. The proteins with Bonferroni adjusted p -values less than 0.05 were selected in the protein pool. Since we aimed at investigating the lung cancer using the network biomarker, the differentially expressed proteins without interaction information were excluded from the protein pool. In addition to the proteins that differentially expressed, the proteins which are highly connected with the proteins in the protein pool based on the PPI information were also included into the pool. In other words, the protein pool consists of both differentially expressed proteins and the proteins that are highly connected with them. On the basis of the protein pool and the PPI information, the rough PPI network can be easily constructed by linking the proteins that have interactions among them. One thing should be noted is that since the data for cancer and non-cancer samples are limited, the number of proteins selected for rough PPI network construction is also restricted. That is, in order to avoid overfitting in network construction, the maximum degree of the proteins in the rough PPI network should be less than the cancer/non-cancer sample number, thereby restricting the size of the rough PPI network.

From the process above, we have selected a protein pool and constructed a rough PPI network among them. The rough PPI network comprises all possible protein interactions under all kinds of experimental conditions. Consequently, the network should be further pruned using microarray data to indicate the effective protein associations for samples with and without lung cancer. Here, a simple linear regression model was applied to prune the rough PPI network to obtain the protein association networks independently for samples with and without cancer, according to their respective data sets. For a target protein i in the rough PPI network, the protein was described by the following protein association model [73].

$$y_i[n] = \sum_{k=1}^{N_i} \alpha_{ik} y_{ik}[n] + \varepsilon_i[n] \quad (1.1)$$

where $y_i[n]$ represents the gene expression level of the target protein i for the sample n , α_{ik} denotes the association ability between the target protein i and its k th interactive protein, which quantifies the expression relation between the interactive proteins and can be identified using the data we have, $y_{ik}[n]$ indicates the gene expression level of the k th protein that interacts with the target protein i for the sample n , N_i is the number of proteins interacting with the target protein i and can be obtained from the rough PPI network, $\varepsilon_i[n]$ denotes the stochastic noises due to other factors or model uncertainty. The biological meaning of Eq. (1.1) is that the expression level of the target protein i is associated with the expression levels of the proteins interacted with it. For each protein in the protein pool, a protein association model was constructed.

After the protein association model of the rough PPI network was constructed, the association parameters in Eq. (1.1) were identified using maximum likelihood estimation method [8, 32] by microarray data (see Appendix 1 for details). Since

there are two data sets of microarray data (smokers with and without cancer), the association parameters were separately identified for cancer data set and non-cancer data set, resulting in $\alpha_{ik,C}$ and $\alpha_{ik,N}$. In this case, for each protein in each phenotype, i.e., with cancer and without cancer, a mathematical description was constructed to characterize the expression association, respectively. Once the association parameters for all proteins in the rough PPI network were identified, the significant protein associations were determined based on the estimated association abilities α_{ik}' s. Akaike Information Criterion (AIC) [1, 32] and student's t-test [48] were employed for both model order selection and significance determination of protein associations (see Appendix 2 for details). In this way, the rough PPI network was pruned and the protein association networks for smokers with and without cancer were constructed, respectively.

On the basis of the identified protein association abilities, two matrices were established to represent the cancer protein association network (CPAN) and the non-cancer protein association network (NPAN).

$$C = \begin{bmatrix} \alpha_{11,C} & \alpha_{12,C} & \cdots & \alpha_{1K,C} \\ \alpha_{21,C} & \alpha_{22,C} & \cdots & \alpha_{2K,C} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{K1,C} & \alpha_{K2,C} & \cdots & \alpha_{KK,C} \end{bmatrix} \quad (1.2)$$

$$N = \begin{bmatrix} \alpha_{11,N} & \alpha_{12,N} & \cdots & \alpha_{1K,N} \\ \alpha_{21,N} & \alpha_{22,N} & \cdots & \alpha_{2K,N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{K1,N} & \alpha_{K2,N} & \cdots & \alpha_{KK,N} \end{bmatrix}$$

where $\alpha_{ij,C}$ and $\alpha_{ij,N}$ indicate the quantitative protein association ability between protein i and protein j for CPAN and NPAN, respectively, and K is the number of proteins in the protein association network. For any protein i and protein j in the protein association network, the association ability α_{ij} quantifies the expression relation between the interactive proteins. If the estimated protein association ability α_{ij} equals to zero, it means that there is no association between protein i and protein j . In addition, we said that protein i is associated with protein j means that the expression level changes of protein i account for the expression level changes of protein j and vice versa. As a consequence, when the estimated protein association ability α_{ij} does not equal to α_{ji} , the one which has larger absolute value would be selected as the association ability between protein i and protein j , i.e., $\alpha_{ij} = \alpha_{ji}$. The resulting cancer and non-cancer protein association networks (CPAN and NPAN) constituted the network biomarker, which was used for determining the significant proteins playing important roles in lung carcinogenesis and for diagnostic evaluation.

1.2.4 Determination of Significant Proteins in Lung Carcinogenesis via the Network Biomarker

According to equations (1.1) and (1.2), the protein association models for CPAN and NPAN can be represented as the following equations.

$$\begin{aligned} Y_C &= CY_C + E_C \\ Y_N &= NY_N + E_N \end{aligned} \quad (1.3)$$

where $Y_C = [y_{1,C}[n] \ y_{2,C}[n] \ \dots \ y_{K,C}[n]]^T$, $Y_N = [y_{1,N}[n] \ y_{2,N}[n] \ \dots \ y_{K,N}[n]]^T$ denotes the vectors of expression levels; E_C and E_N indicate the noise vectors in cancer case and non-cancer case, respectively. A matrix indicating the difference between two protein association networks is defined as $C - N$ [73].

$$\begin{aligned} D &= \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1K} \\ d_{21} & d_{22} & \cdots & d_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ d_{K1} & d_{K2} & \cdots & d_{KK} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_{11,C} - \alpha_{11,N} & \alpha_{12,C} - \alpha_{12,N} & \cdots & \alpha_{1K,C} - \alpha_{1K,N} \\ \alpha_{21,C} - \alpha_{21,N} & \alpha_{22,C} - \alpha_{22,N} & \cdots & \alpha_{2K,C} - \alpha_{2K,N} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{K1,C} - \alpha_{K1,N} & \alpha_{K2,C} - \alpha_{K2,N} & \cdots & \alpha_{KK,C} - \alpha_{KK,N} \end{bmatrix} \end{aligned} \quad (1.4)$$

where d_{ij} denotes the protein association ability difference between CPAN and NPAN among protein i and protein j . Using the matrix D to show the difference of network structure between CPAN and NPAN, a score named carcinogenesis relevance value (CRV) was then presented to quantify the correlation of each protein with significance of lung carcinogenesis. To identify the significant proteins for lung carcinogenesis, two important issues were taken into consideration. First, the magnitude of the association abilities α_{ij} 's denotes the significance of one protein to the other one. A higher absolute value of α_{ij} implies that the two proteins are more tightly associated. Second, if a protein plays more crucial roles in lung carcinogenesis, the difference of association numbers linked to the protein for CPAN and NPAN would be larger. For instance, if one protein associates with a lot of proteins in CPAN but associates with no protein in NPAN, it would be more likely involved in lung carcinogenesis. As a result, the CRV was determined based on the difference of protein association abilities as the following equation.

$$CRV_i = \sum_{j=1}^K |d_{ij}| \quad (1.5)$$

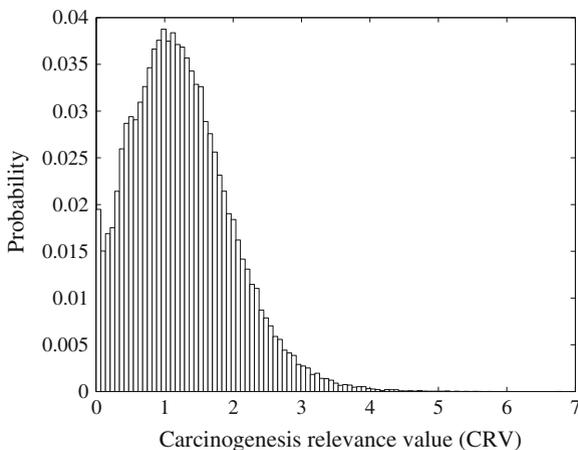


Fig. 1.2 Distribution of carcinogenesis relevance values (CRVs) of random networks. The null distribution of CRVs is generated by 100,000 randomly permuted network structures

For the i th protein in the network biomarker, the implication of Eq.(1.5) is that the CRV quantifies the extent of protein associations that differentiate CPAN from NPAN in the network biomarker.

For each protein, in addition to the CRV assigned, an empirical p -value was also computed to determine the significance of the CRV. To determine the p -value for an observed CRV, a null distribution of CRVs (Fig. 1.2) was generated by repeatedly permuting the network structure of the rough PPI network and computing the CRV for each random network structure. The permutation of the network structure was performed by keeping the network size, i.e., the proteins with which a particular protein interacted were permuted without changing the total number of protein interactions. The process was repeated 100,000 times and the p -value of the corresponding CRV was estimated as the fraction of random network structures whose CRV is at least as large as the CRV of the real network structure. The CRVs with p -value ≤ 0.05 were determined as significant CRVs and the corresponding proteins were identified as significant proteins in lung carcinogenesis.

1.2.5 Diagnostic Evaluation by the Network Biomarker

An important feature of the proposed network biomarker approach is that it can not only be used for investigation of significant proteins for lung cancer, but also for diagnosis of smokers suspect with lung cancer. Given the new microarray expression data for the smoker, we can classify the sample into smoker with or without cancer based on CPAN and NPAN within the network biomarker. The idea comes from the similarity comparison of new sample data between CPAN and NPAN.

Specifically, if a sample data is more similar to the network structure of CPAN than of NPAN, it would be regarded as the smoker with lung cancer, and vice versa. Since only one sample cannot be used for network construction, the new sample data was mapped to the CPAN and NPAN identified above and the mapping error would be employed as the criteria of classification. Suppose that we had a new sample data $Z = [z_1 \ z_2 \ \dots \ z_K]^T$ from a smoker, based on Eqs.(1.1) and (1.3), the mapping errors for CPAN and NPAN are respectively defined as

$$\begin{aligned} ME_C &= \|Z - C \cdot Z\|_2 \\ ME_N &= \|Z - N \cdot Z\|_2 \end{aligned} \quad (1.6)$$

where $\|P\|_2 = \left(\sum_{i=1}^K p_i^2\right)^{1/2}$ when $P = [p_1 \ p_2 \ \dots \ p_K]^T$. The mapping errors can be considered as the similarity measurement of the new sample Z to the systems CPAN and NPAN. The smaller the mapping error is, the more matching the sample data is to the protein association network. Consequently, if $ME_C < ME_N$, the new sample Z is more similar to the cancer system and is classified into the smokers with cancer category, and vice versa. The criteria of mapping errors have simultaneously taken account of the protein association network structures with quantitative association abilities and the expression levels of the proteins. Further, since the modeling error is regarded as the criterion of classification, it is a classification more dependent on network structure than data only and therefore could be also suitable for classification with independent data. We believe that the kind of classification approach can provide new perspective for diagnostic evaluation.

1.3 Results

1.3.1 Construction of Network Biomarker and Determination of Significant Proteins in Lung Carcinogenesis

We applied the proposed network biomarker approach for molecular investigation and diagnosis of lung cancer. The primary data set (79 smokers with lung cancer and 73 smokers without lung cancer) of GSE4115 downloaded from GEO database <http://www.ncbi.nlm.nih.gov/geo/> was used for construction of network biomarker. Based on the classical statistical method ANOVA, 199 proteins which have PPI information were identified as the differentially expressed proteins and were selected in the protein pool. In addition, the proteins that linked to three differentially expressed proteins in the protein pool according to PPI information were also included in the pool. In this case, the protein pool consisted of 339 proteins. Then, the proteins that have PPI information among them were linked together, resulting in the rough PPI network. The expression profiles for smokers with and without cancer and the

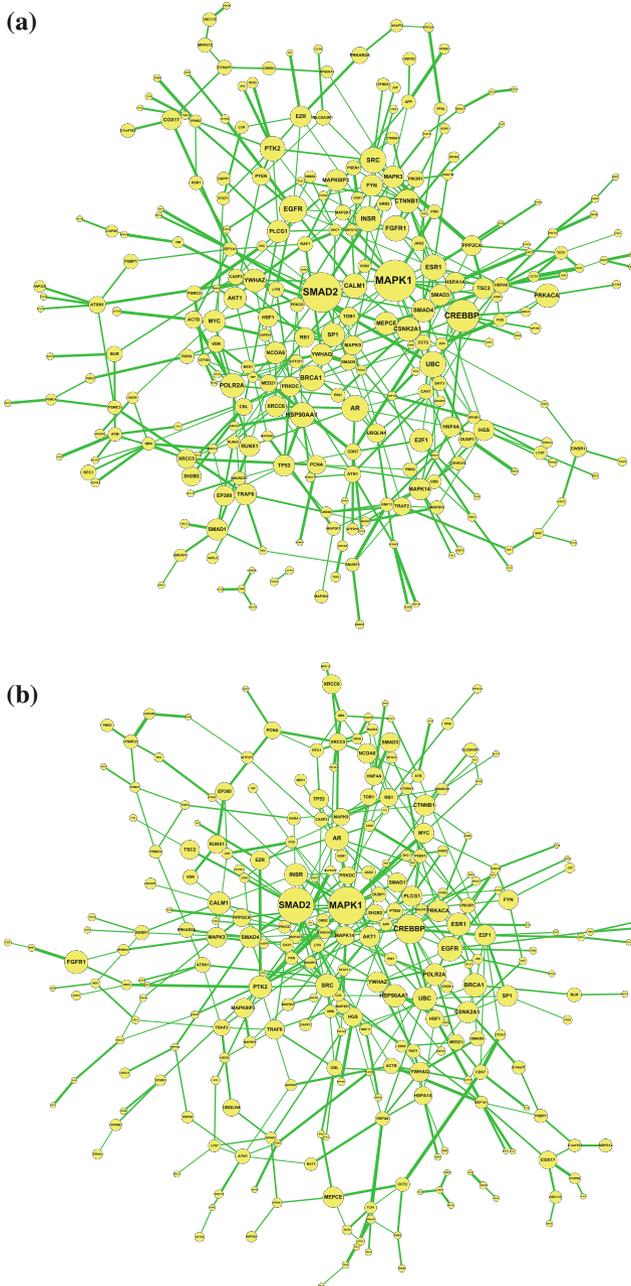


Fig. 1.3 The constructed network biomarker. **a** Cancer protein association network (CPAN). **b** Non-cancer protein association network (NPAN). The node size is proportional to the CRV for each protein and the edge width represents the magnitude of the association ability among two proteins. The figures are created using cytoscape [56]

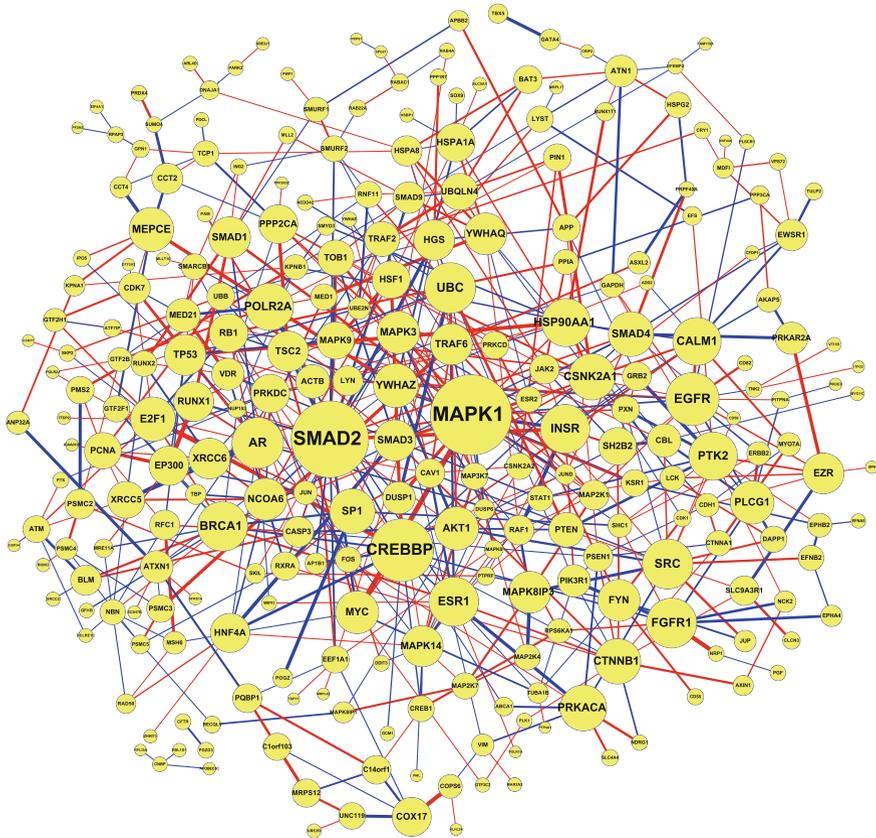


Fig. 1.4 The difference between CPAN and NPAN. The node size is proportional to the CRV for each protein and the edge width represents the magnitude of the association ability among two proteins. Red and blue edges indicate the positive and negative values of d_{ij} 's in Eq. (1.4), respectively. The figure is created using cytoscape [56]

protein association model Eq. (1.1) were further employed to prune the rough PPI network. The CPAN and NPAN, which consisted of 399 and 393 protein associations respectively, would constitute the network biomarker of lung cancer (Fig. 1.3). The difference between CPAN and NPAN was further shown in Fig. 1.4. According to the CPAN and NPAN with quantitative association abilities, the CRVs for each protein were computed and the significance of these CRVs was determined. Consequently, 40 proteins were identified to play significant roles in lung carcinogenesis and were shown in Table 1.1.

Table 1.1 The identified significant proteins in lung carcinogenesis

Protein symbol ^a	CRV	<i>p</i> -value	Functional annotation ^b			Literature evidence ^c
			Cell growth	Cell survival	Cell migration	
MAPK1	8.3418	<1e-5	+	+	+	[29, 67]
SMAD2	7.7901	<1e-5	+	+	+	[5]
CREBBP	5.7870	0.00002	+			[34]
EGFR	4.3635	0.00086	+	+	+	[16, 25, 38]
AR	4.0966	0.00159	+	+	+	[64]
UBC	4.0331	0.00180				
SRC	3.9446	0.00218	+	+	+	[7, 44]
FGFR1	3.9227	0.00237	+		+	[4]
BRCA1	3.9049	0.00243	+	+		[74]
ESR1	3.8409	0.00295	+	+	+	[24]
INSR	3.7946	0.00329	+		+	[13]
PTK2	3.6758	0.00432	+	+	+	[43, 44]
HSP90AA1	3.6732	0.00436	+	+	+	[20]
CALM1	3.6363	0.00482		+		
POLR2A	3.5701	0.00547				
CSNK2A1	3.4128	0.00761	+	+		[69]
PRKACA	3.3688	0.00856		+		
CTNNB1	3.2935	0.00994	+	+	+	[3]
SP1	3.2397	0.01133	+	+		[15]
SMAD4	3.1947	0.01266	+	+	+	[5]
E2F1	3.1382	0.01407	+	+		[30]
YWHAZ	3.1212	0.01467	+			[39]
MEPCE	3.0968	0.01545				
AKT1	3.0193	0.01857	+	+	+	[75]
PLCG1	2.9654	0.02069			+	[54]
MYC	2.8987	0.02385	+	+		[77]
MAPK3	2.8545	0.02654	+	+	+	[29, 67]
NCOA6	2.8132	0.02892	+	+		
FYN	2.7833	0.03089	+		+	[10]
MAPK8IP3	2.7746	0.03141			+	
YWHAQ	2.7582	0.03242	+			[70]
TRAF6	2.7150	0.03535		+		[31]
SMAD1	2.6940	0.03697	+	+	+	[37]
SMAD3	2.6815	0.03815	+	+	+	[5]
MAPK14	2.6727	0.03894	+	+	+	[66]
TP53	2.6522	0.04056	+	+	+	[16, 25, 28]
XRCC6	2.6270	0.04263		+		
EZR	2.6213	0.04314			+	[14]
TSC2	2.6116	0.04401	+	+	+	[40]
HGS	2.5730	0.04744	+			

^aThe full names of these proteins according to UniProt database <http://www.uniprot.org/> are listed in Appendix 3

^bThe functional annotations are from the Gene Ontology database <http://www.geneontology.org/> and literatures

^cThe literature evidences indicate that overexpression/dysregulation of the specific protein or mutation of the corresponding gene would result in carcinogenesis

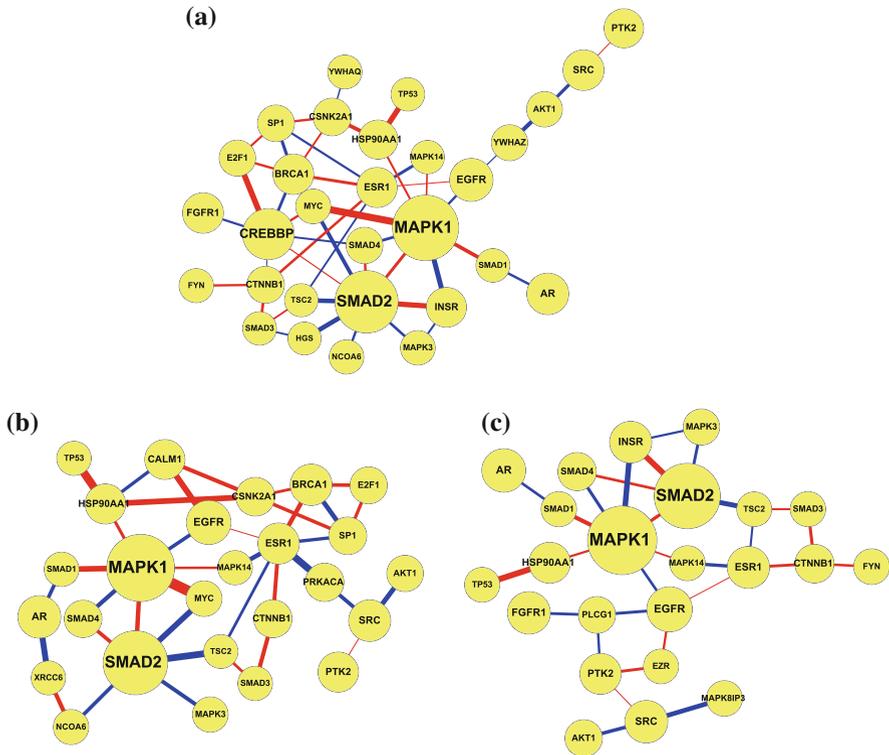


Fig. 1.5 The functional subnetworks of significant proteins identified according to the network biomarker. **a** Cell growth functional subnetwork. **b** Cell survival functional subnetwork. **c** Cell migration functional subnetwork. All the functional subnetworks are extracted from Fig. 1.4

1.3.2 Mechanism Investigation of the Significant Proteins in Lung Carcinogenesis

A total of 40 proteins were determined as significant proteins in lung carcinogenesis using the network biomarker. These 40 significant proteins identified can be divided into three categories according to the functional annotations (Table 1.1) and the three functional subnetworks were shown in Fig. 1.5. The mechanisms for carcinogenesis of the significant proteins were further investigated.

- (1) Cell growth: Cancer is the complex disease of uncontrolled cell growth. Therefore, the proteins responsible for cell growth are likely to play critical roles in lung carcinogenesis. Among the 40 significant proteins identified based on the network biomarker, 30 proteins are annotated with cell growth, reinforcing its significance (Table 1.1). The mitogen-activated protein kinase (MAPK) cascade is a highly conserved module that is relevant to many cancers. Three

MAPK-related proteins annotated with cell growth, MAPK1 (ERK2), MAPK3 (ERK1) and MAPK14 (p38 α), were identified as significant proteins in lung carcinogenesis. Elevated expression of activated MAPK1/3 have been observed in NSCLC [67] and may play a role in lung metastasis [29]. MAPK14 (p38 α) is well known for its important roles in response to inflammation and environmental stress [36] and its protein expression is more than three times lower in human lung tumors compared with normal human lung tissue, suggesting that MAPK14 may function as a negative regulator of lung carcinogenesis [66]. Four proteins from the SMAD protein family (SMAD1, SMAD2, SMAD3 and SMAD4) were also identified as significant proteins in lung carcinogenesis. The SMAD proteins, which consist of three functional classes, are signal transducers and transcriptional modulators [57]. Mutations and altered expression for these four proteins were observed in human cancer [5, 37], indicating the significance of the SMAD protein family in carcinogenesis although the roles they play may be different.

EGFR, FGFR1, and INSR are receptor tyrosine kinases (RTKs) and are bound by epidermal growth factor, fibroblast growth factor, and insulin, respectively. Receptor tyrosine kinases have been shown not only to be key regulators of normal cellular processes but also to critically involve in the development and progression of human cancers [79]. EGFR is one of the most extensively studied proteins in carcinogenesis. It is overexpressed in NSCLC as well as in other common tumors and its increased expression is associated with aggressive tumor growth and therapy resistance [38]. EGFR was also found to be significantly mutated in lung adenocarcinoma [16, 25]. Because of its significance, EGFR becomes a popular therapeutic target for carcinogenesis. Gefitinib (Iressa) and erlotinib (Tarceva) are two targeted therapies that specifically inhibit EGFR tyrosine kinase [25, 38]. FGFR1 and INSR are both shown to be involved in carcinogenesis and thus novel attractive targets for cancer therapeutic strategies like EGFR [4, 13].

Many proteins acted as transcriptional regulators were also identified. CREBBP is a transcriptional co-activator downstream of the TGF β pathway and the mutations and deletions of the CREBBP gene are associated with lung cancer [34]. CTNNB1 (β -catenin) is one of the core components in the Wnt pathway. Mutation of β -catenin, which results in aberrant activation of the Wnt pathway, is a frequent cause in human cancers [3]. E2F1 is one of the significant proteins involved in the cell cycle. Its overexpression has been demonstrated in both NSCLC and SCLC and is induced by its upstream RB protein [30]. MYC is a nuclear phosphoprotein and functions as a transcription factor. It controls cell cycle progression by simulating G1/S transition and may result in loss of cell cycle arrest and uncontrolled tumor growth when dysfunction [77]. Other identified significant proteins such as AR, ESR1, SRC, FYN, YWHAQ, YWHAZ, and HSP90AA1 were also shown to involve in carcinogenesis [7, 10, 20, 24, 39, 64, 70].

- (2) Cell survival: The ability of tumor cell populations to expand in number is determined not only by the rate of cell proliferation but also by the rate of

cell death [22]. In addition, the acquired resistance of programmed cell death, apoptosis, is a hallmark of cancer. Consequently, proteins annotated with cell survival might be important in carcinogenesis. Twenty-seven significant proteins identified by the network biomarker were annotated with cell survival. TP53 (p53) is a well-studied tumor suppressor protein and plays important roles in anti-cancer mechanisms. Its activation is induced by a number of stress signals such as DNA damage, oxidative stress and activated oncogenes. Activated p53 induces cell cycle arrest, apoptosis and inhibition of angiogenesis and metastasis. Once damaged, tumor suppression is severely reduced, resulting in uncontrolled proliferation of the cell. Due to the importance in carcinogenesis, it is not surprise that p53 was found to be significantly mutated in lung adenocarcinoma as well as in squamous cell carcinoma and SCLC [16, 25, 28].

TRAF6 functions as a signal transducer in the NF κ B pathway that activates IKK, in response to proinflammatory cytokines. The identification of TRAF6 by the proposed network biomarker approach reinforces the linking between inflammation and cancer [12, 31]. BRCA1 is a nuclear phosphoprotein that plays a role in genomic stability. The mutant phenotype of BRCA1 predisposes to breast and to ovarian cancer [74]. SP1 is a transcription factor downstream of the TGF β pathway and its overexpression contributes to malignant transformation [15]. Other protein kinases, AKT and CSNK2A1 (CK2), were also indicated to participate in the carcinogenic process [52, 69, 75].

- (3) Cell migration: With the progression of cancer, the malignant tumor cells acquire the ability to migrate and metastasize to distant sites. As a result, the proteins that are relevant to the cell migration capability are crucial for the carcinogenic process. Twenty-three out of 40 significant proteins were annotated with cell migration. PTK2 (FAK), a protein tyrosine kinase in the RTK pathway is an important mediator of cell proliferation, cell survival and cell migration process. Substantial evidence has shown that activated PTK2 leads to tumor growth and metastasis [44], and levels of expression correlates with the invasive potential of tumors [43]. High levels of TSC2 were correlated with increased metastasis and reduced survival in breast cancer patients, revealing a protumorigenic role for TSC2 [40]. The other two significant proteins, PLCG1 and EZR (ezrin), were demonstrated to play critical roles in the metastatic potential of cancer cells but not in the primary tumor growth [14, 54].

For nine out of 40 significant proteins identified, nothing is known in the literature with respect to their roles in lung cancer (see Table 1.1). UBC is a polyubiquitin precursor. Protein ubiquitination is a fundamental regulatory post-translational modification controlling intracellular signaling events. It has been associated with protein degradation, DNA repair, cell cycle regulation, endocytosis, and kinase modification [35]. Dysregulation of ubiquitin-mediated signaling is increasingly implicated in some human diseases. Therefore, UBC may be an important target for further characterization of lung carcinogenesis. CALM1 is calmodulin, which mediates the control of a large number of enzymes and other proteins by Ca²⁺. It is an essential regulator of cell cycle progression and cell survival. Further studies are needed to

examine its relation with carcinogenesis. PRKACA is a cAMP-dependent protein kinase. The identification of PRKACA as a significant protein implies that cAMP signaling might also involve in lung carcinogenesis. MAPK8IP3 function as a scaffold protein in the RTK pathway; NCOA6, HGS, and XRCC6 are annotated with cell growth and/or cell survival. However, no literature evidence indicates their relevance with carcinogenesis until now, which makes them potential targets for further understanding of lung carcinogenesis.

1.3.3 Diagnostic Evaluation of Smokers Suspect with Lung Cancer Using the Network Biomarker

The network biomarker was constructed based on the primary data set of GSE4115. An independent data set (the prospective data set of GSE4115, 18 smokers with lung cancer and 17 smokers without lung cancer) was then used to evaluate the diagnostic performance of the proposed network biomarker. Among the 35 samples, 26 were accurately classified, resulting in the accuracy of 74.29%. The sensitivity and specificity of the proposed approach were also evaluated. The network biomarker can identify the smokers with and without cancer with a high sensitivity of 83.33% and a moderate specificity of 64.71%, enabling the proposed network biomarker to be effective to diagnose the smokers with lung cancer. Thus, it can be used as one kind of screening test, which, with the help of other clinical diagnostic tools, accelerates the whole diagnostic process and improves the diagnostic sensitivity.

The cause of the moderate specificity was further investigated. The reason why the specificity is not as high as the sensitivity is that there are some smokers without cancer misclassified in the cancer category. The misclassification may be due to similar molecular patterns, i.e., the gene expression profiles of smokers with cancer are highly similar to those of smokers without cancer. In order to validate the hypothesis, Pearson correlation coefficients [48] of gene expression profiles for both smokers with and without cancer were calculated. The mean correlation coefficient among smokers with cancer is 0.9616 whereas the mean correlation coefficient among smokers without cancer is 0.9441. In addition, the mean correlation coefficient among smokers with and without cancer is as high as 0.9437, suggesting that the molecular patterns of smokers with and without cancer are indeed highly similar. Because of the highly similar molecular patterns and the fact that cigarette smoking is the main contributor of lung cancer, we may further predict that the smokers without cancer initially would be likely to have lung cancer one day.

In order to validate the predictive performance of the proposed network biomarker, several comparisons were made. First, we tested the predictive performance without the information of protein-protein interactions. The 199 differentially expressed proteins selected by ANOVA using the primary data set were used for classification of the prospective data set. A simple hierarchical clustering was performed, indicating that using the gene expression only cannot accurately classify the prospective data set (65.71% accuracy). The comparison shows that the integration of gene expression

profiles and protein interaction information can improve lung cancer diagnosis. Second, we compared the predictive performance of the proposed method with randomly selected networks. The average accuracy for 100,000 randomly selected networks of 339 proteins is 48.44 %, demonstrating the significance of the proposed network biomarker. Third, in addition to evaluating the predictive performance using an independent data set, a 5-fold cross-validation was applied on the primary data set plus the prospective data set. The accuracy of the cross-validation is similar to 74.29 % computed using independent data set, showing the robustness and reproducibility of the proposed network biomarker approach.

1.4 Discussion

Cancer is a complex disease and carcinogenesis in humans is a multistep process that transforms normal cells into malignant derivatives. Many researchers are investigating the underlying mechanisms which prompt the uncontrolled cell proliferation and metastasis. They successfully identify some key components of different steps in carcinogenesis and some therapeutic interventions have been developed to at least slow down the carcinogenic process. However, because of the complexity, the therapy that targets some specific molecules is only partially effective and tumor-specific. Therefore, investigation of the carcinogenesis from the systems perspective is inevitable. On the other hand, biomarker identification for cancer diagnosis has been the research focus in the biomedical field since the biomarkers could provide early detection for cancer. As a result, in this chapter, a network biomarker approach is proposed for molecular investigation and diagnosis simultaneously. The proposed approach was applied on the sample data obtained from smokers with and without lung cancer and 40 significant proteins were identified in lung carcinogenesis. The network biomarker considers not only differentially expressed proteins but also the protein association network structure. This allows the identification of proteins with low discriminative potential if they are associated with many other significant proteins [11]. This property is important for the discovery of significant protein in lung carcinogenesis and provides mechanistic insights into the process. From the mechanism investigation for the 40 significant proteins identified using the network biomarker, we find that the significant proteins identified are involved in the pathways that are responsible for the cellular processes including proliferation, differentiation, apoptosis, and metastasis. In normal cells, these cellular processes are precisely regulated to achieve the appropriate responses. However, in carcinogenesis, mutations damage the key components and therefore dysregulate the cellular processes, resulting in aberrant expression of the significant proteins and the transformed malignancy. More importantly, from the result presented, we find that the dysregulated signals exist for multiple pathways. Two possibilities are inferred: the genetic mutations are accumulated for components of different pathways or the aberrant signals affect different pathways through the crosstalk mechanisms. Further investigations are needed for the elucidation of these two hypotheses. In addition to the investigation of significant proteins, the network

biomarker can be used as one kind of screening test with high sensitivity. Using the same data (the prospective data set of GSE4115), conventional bronchoscopy was shown to be 44% sensitive for cancer detection [59], which is only half of the proposed network biomarker. As a result, the clinical usefulness of the network biomarker is reinforced.

Although our proposed method is shown to be useful, some limitations or improvements still need to be taken into consideration. In the proposed network biomarker approach, gene expression profiles were overlaid to the corresponding proteins for further analysis. However, levels of mRNA do not always correlate with protein levels and do not provide information on post-translational modification such as phosphorylation that may be critical for regulating protein activity [25]. Consequently, emerging high-throughput proteomic techniques like protein microarrays would be more helpful for our method to significantly improve the detection performance than mRNA microarray data. In addition, if the genome-wide gene expression levels and protein expression levels can be obtained simultaneously, we are able to construct the integrated cellular networks of transcription regulations and protein interactions which provide a more integrated network biomarker to gain insight into the carcinogenic process [71]. The protein-protein interaction data from public databases also plays important roles in the proposed method. Nevertheless, there is a large variation in the coverage of protein interaction data across the interaction databases [42]. Therefore, HPRD and BioGRID databases were integrated for the PPI information here. We believe that the increased quality and coverage of protein interaction data would enhance the proposed network biomarker approach for characterization of lung carcinogenesis. Another limitation of the proposed method is that the size of the protein association network is restricted by the sample size available because of avoiding overfitting in network construction. It results in the exclusion of some well-studied proteins which have relevance to the lung carcinogenesis in the network biomarker, including KRAS, MET, PI3KCA. To overcome the problem, more samples are needed. It is our belief that the diagnostic evaluation using the network biomarker would be improved once the constructed protein association networks can be expanded. Many other groups identified the discriminative subnetworks using different methods, especially graph theory-based methods [53, 63]. For example, Tian et al. proposed a hypergraph-based iterative learning algorithm, which minimizes a cost function under a unified regularization framework, for subnetwork identification [63]. These graph-based methods can also be incorporated to improve the significant protein selection in the proposed method. Further, for the data used, the samples for gene expression profiling are simply divided into two groups, smokers with cancer and without cancer. If more sample data, especially cancer stage-specific samples, are available, we can then determine how the network evolves and changes during cancer progression using the proposed method.

Our network biomarker gives systematic insights into the lung carcinogenic process and provides a good identification method for significant proteins relevant to it. The significant proteins consist of the ones that are shown to be lung cancer-related and many others that have not been previously reported. These proteins not only provide new targets for further studies to understand the mechanisms

in lung carcinogenesis but also are potential targets for therapeutic interventions. The main challenge of cancer research is to find an effective therapeutic approach that specifically kills malignant cells. Conventional chemotherapy acts by killing all rapidly dividing cells, resulting in toxic effects and damage of normal tissues [51]. With the advances in understanding the mechanisms involved in carcinogenic process, the so-called targeted therapy, which is more effective and less harmful to normal cells, is developed to inhibit the specific molecules that play crucial roles in tumor growth. The significant proteins identified by the proposed network biomarker provide suitable molecules to be targeted. For example, gefitinib (Iressa) and erlotinib (Tarceva) are two tyrosine kinase inhibitors that specifically target EGFR. Although the efficacy can be shown, there are still patients not responded well to these drugs [3]. One explanation for the circumstance is that the single-target agents likely result in network compensation and drug resistance [18]. As a result, multi-target therapeutic interventions that impact multiple targets simultaneously might be required for humans to combat against cancer. Multi-target therapeutics can be more efficacious and less vulnerable to adaptive resistance because the biological system is less able to compensate for the action of two or more drugs simultaneously [78]. Hopefully, with the help of significant proteins identified by the proposed network biomarker approach and the pathway information, the multi-target therapeutic interventions that act on different critical pathways in lung carcinogenesis can be developed.

1.5 Conclusions

Lung cancer is the leading cause of cancer deaths worldwide. Understanding the causes and the underlying mechanisms can help fight the disease. In this chapter, a network biomarker approach, which integrated gene expression profiles and protein interaction information, was developed for molecular investigation and diagnosis for lung cancer. The network biomarker constructed shed light on the lung carcinogenic process from the systems perspective and are used for significant protein identification and diagnostic evaluation. The diagnostic results indicate that the network biomarker is sensitive to the diagnosis of smokers with lung cancer and can be used as one kind of screening test. Most importantly, the significant proteins identified by the network biomarker give mechanistic insights into the carcinogenic process and provide potential therapeutic targets to combat against cancer.

Appendix

Appendix 1: Identification of Association Parameters

After the protein association model of the rough PPI network was constructed, the association parameters in Eq. (1.1) were identified using maximum likelihood estimation method [32]. Equation (1.1) can be written in the following regression form.

$$\begin{aligned}
y_i[n] &= [y_{i1}[n] \dots y_{iN_i}[n]] \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{iN_i} \end{bmatrix} + \varepsilon_i[n] \\
&= \phi_i[n] \cdot \theta_i + \varepsilon_i[n]
\end{aligned} \tag{A1}$$

where $\phi_i[n]$ denotes the regression vector which can be obtained from the processing above, θ_i is the parameter vector to be estimated. Suppose that there are M samples for us to estimate the association parameters, $\{y_i[n] \phi_i[n]\}$ for $n \in \{1, 2, \dots, M\}$ can be easily acquired via the same procedure. In this case, Eq. (A1) for different samples can be represented as the following form.

$$\begin{bmatrix} y_i[1] \\ \vdots \\ y_i[M] \end{bmatrix} = \begin{bmatrix} \phi_i[1] \\ \vdots \\ \phi_i[M] \end{bmatrix} \cdot \theta_i + \begin{bmatrix} \varepsilon_i[1] \\ \vdots \\ \varepsilon_i[M] \end{bmatrix} \tag{A2}$$

For simplicity, the notations Y_i , Φ_i , and e_i were defined to express Eq. (A2) as follows

$$Y_i = \Phi_i \cdot \theta_i + e_i \tag{A3}$$

In Eq. (A2), we assumed noises $\varepsilon_i[n]$ for different samples as independent random variables of normal distribution with zero mean and unknown variance σ_i^2 , i.e., the variance of e_i is $\Sigma_i = E\{e_i e_i^T\} = \sigma_i^2 I$, where I is the identity matrix. The probability density function of e_i is given as

$$p(e_i) = \left((2\pi)^M \det \Sigma_i \right)^{-1/2} \exp \left\{ -\frac{1}{2} e_i^T \Sigma_i^{-1} e_i \right\} \tag{A4}$$

Considering Eqs. (A3) and (A4), the likelihood function can be expressed as

$$L(\theta_i, \sigma_i^2) = p(\theta_i, \sigma_i^2) = \left(2\pi \sigma_i^2 \right)^{-M/2} \exp \left\{ -\frac{1}{2\sigma_i^2} (Y_i - \Phi_i \theta_i)^T (Y_i - \Phi_i \theta_i) \right\} \tag{A5}$$

Maximum likelihood estimation method aims at finding θ_i and σ_i^2 to maximize the likelihood function in Eq. (A5). For the simplicity of computation, it is practical to take the logarithm of the likelihood function, and we have the following log-likelihood function

$$\log L(\theta_i, \sigma_i^2) = -\frac{M}{2} \log \left(2\pi \sigma_i^2 \right) - \frac{1}{2\sigma_i^2} \sum_{n=1}^M [y_i[n] - \phi_i[n] \cdot \theta_i]^2 \tag{A6}$$

where $y_i[n]$ and $\phi_i[n]$ are the n -th element of Y_i and Φ_i , respectively. Here, the log-likelihood function is expected to have the maximum at $\theta_i = \hat{\theta}_i$ and $\sigma_i^2 = \hat{\sigma}_i^2$.

The necessary conditions for determining the maximum likelihood estimates $\hat{\theta}_i$ and $\hat{\sigma}_i^2$ must conform to the following two equations.

$$\begin{aligned} \left. \frac{\partial \log L(\theta_i, \sigma_i^2)}{\partial \theta_i} \right|_{\theta_i = \hat{\theta}_i} &= 0 \\ \left. \frac{\partial \log L(\theta_i, \sigma_i^2)}{\partial \sigma_i^2} \right|_{\sigma_i^2 = \hat{\sigma}_i^2} &= 0 \end{aligned} \quad (\text{A7})$$

After some computational deduction, the estimated parameters $\hat{\theta}_i$ and $\hat{\sigma}_i^2$ are

$$\hat{\theta}_i = \left(\Phi_i^T \Phi_i \right)^{-1} \Phi_i^T Y_i \quad (\text{A8})$$

$$\hat{\sigma}_i^2 = \frac{1}{M} \sum_{n=1}^M \left[y_i[n] - \phi_i[n] \cdot \hat{\theta}_i \right]^2 = \frac{1}{M} (Y_i - \Phi_i \hat{\theta}_i)^T (Y_i - \Phi_i \hat{\theta}_i) \quad (\text{A9})$$

Appendix 2: Determination of Significant Protein Associations

When the association parameters were identified, Akaike Information Criterion (AIC) [1, 32] and student's t-test [48], which is used to calculate the p -values of the association abilities, were employed for both model order selection and determination of significant protein associations. The AIC, which attempts to include both the estimated residual variance and model complexity in one statistics, decreases as the residual variance decreases and increases as the number of parameters increases. As the expected residual variance decreases with increasing parameter numbers for nonadequate model complexities, there should be a minimum around the correct parameter number [1, 32]. Therefore, AIC can be used to select model structure based on the association abilities (α_{ik} 's) identified above. Due to computation efficiency, it is impractical to compute the AIC statistics for all possible regression models. Stepwise methods such as forward selection method and backward elimination method are developed to avoid the complexity of exhausted search [27, 48, 55]. However, in the case of backward selection method, a variable once eliminated can never be reintroduced into the model, and in the case of forward selection, once included can never be removed [48, 55]. Thus, the stepwise regression method which combines forward selection method and backward elimination method was applied to compute the AIC statistics. Once the estimated regulatory parameters were examined using the AIC model selection criteria, the student's t-test was employed to calculate the p -values for the association abilities (α_{ik} 's) under the null hypothesis $H_0 : \alpha_{ik} = 0$ [48] to determine the significant protein associations. The p -values computed were then adjusted by Bonferroni correction to avoid a lot of spurious positives [48]. The associations which adjusted p -value ≤ 0.05 were determined as significant associations and be preserved in the protein association network.

Appendix 3: The full Names of the Significant Proteins Identified

MAPK1	Mitogen-activated protein kinase 1
SMAD2	Mothers against decapentaplegic homolog 2
CREBBP	CREB-binding protein
EGFR	Epidermal growth factor receptor
AR	Androgen receptor
UBC	Polyubiquitin-C
SRC	Proto-oncogene tyrosine-protein kinase Src
FGFR1	Basic fibroblast growth factor receptor 1
BRCA1	Breast cancer type 1 susceptibility protein
ESR1	Estrogen receptor
INSR	Insulin receptor
PTK2	Focal adhesion kinase 1
HSP90AA1	Heat shock protein HSP 90-alpha
CALM1	Calmodulin
POLR2A	DNA-directed RNA polymerase II subunit RPB1
CSNK2A1	Casein kinase II subunit alpha
PRKACA	cAMP-dependent protein kinase catalytic subunit alpha
CTNNB1	Catenin beta-1
SP1	Transcription factor Sp1
SMAD4	Mothers against decapentaplegic homolog 4
E2F1	Transcription factor E2F1
YWHAZ	14-3-3 protein zeta/delta
MEPCE	7SK snRNA methylphosphate capping enzyme
AKT1	RAC-alpha serine/threonine-protein kinase
PLCG1	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-1
MYC	Myc proto-oncogene protein
MAPK3	Mitogen-activated protein kinase 3
NCOA6	Nuclear receptor coactivator 6
FYN	Tyrosine-protein kinase Fyn
MAPK8IP3	C-Jun-amino-terminal kinase-interacting protein 3
YWHAQ	14-3-3 protein theta
TRAF6	TNF receptor-associated factor 6
SMAD1	Mothers against decapentaplegic homolog 1
SMAD3	Mothers against decapentaplegic homolog 3
MAPK14	Mitogen-activated protein kinase 14
TP53	Cellular tumor antigen p53
XRCC6	X-ray repair cross-complementing protein 6
EZR	Ezrin
TSC2	Tuberin

References

1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
2. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–511
3. Barker N, Clevers H (2006) Mining the Wnt pathway for cancer therapeutics. *Nat Rev Drug Discov* 5:997–1014
4. Behrens C, Lin HY, Lee JJ, Raso MG, Hong WK (2008) Immunohistochemical expression of basic fibroblast growth factor and fibroblast growth factor receptors 1 and 2 in the pathogenesis of lung cancer. *Clin Cancer Res* 14:6014–6022
5. Bierie B, Moses HL (2006) TGF-beta and cancer. *Cytokine Growth Factor Rev* 17:29–40
6. Brambilla E, Travis WD, Colby TV, Corrin B, Shimosato Y (2001) The new World Health Organization classification of lung tumours. *Eur Respir J* 18:1059–1068
7. Byers LA, Sen B, Saigal B, Diao L, Wang J, Nanjundan M, Cascone T, Mills GB, Heymach JV, Johnson FM (2009) Reciprocal regulation of c-Src and STAT3 in non-small cell lung cancer. *Clin Cancer Res* 15:6852–6861
8. Chang YH, Wang YC, Chen BS (2006) Identification of transcription factor cooperativity via stochastic system model. *Bioinformatics* 22:2276–2282
9. Chen J, Aronow BJ, Jegga AG (2009) Disease candidate gene identification and prioritization using protein interaction networks. *BMC Bioinform* 10:73
10. Chen ZY, Cai L, Bie P, Wang SG, Jiang Y, Dong JH, Li XW (2010) Roles of Fyn in pancreatic cancer metastasis. *J Gastroenterol Hepatol* 25:293–301
11. Chuang HY, Lee E, Liu YT, Lee D, Ideker T (2007) Network-based classification of breast cancer metastasis. *Mol Syst Biol* 3:140
12. Coussens LM, Werb Z (2002) Inflammation and cancer. *Nature* 420:860–867
13. Cox ME, Gleave ME, Zakikhani M, Bell RH, Piura E, Vickers E, Cunningham M, Larsson O, Fazli L, Pollak M (2009) Insulin receptor expression by human prostate cancers. *Prostate* 69:33–40
14. Deng X, Tannehill-Gregg SH, Nadella MV, He G, Levine A, Cao Y, Rosol TJ (2007) Parathyroid hormone-related protein and ezrin are up-regulated in human lung cancer bone metastases. *Clin Exp Metastasis* 24:107–119
15. Deniaud E, Baguet J, Mathieu AL, Pages G, Marvel J, Leverrier Y (2006) Overexpression of Sp1 transcription factor induces apoptosis. *Oncogene* 25:7096–7105
16. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, Sougnez C, Greulich H, Muzny DM, Morgan MB, Fulton L, Fulton RS, Zhang Q, Wendl MC, Lawrence MS, Larson DE, Chen K, Dooling DJ, Sabo A, Hawes AC, Shen H, Jhangiani SN, Lewis LR, Hall O, Zhu Y, Mathew T, Ren Y, Yao J, Scherer SE, Clerc K, Metcalf GA, Ng B, Milosavljevic A, Gonzalez-Garay ML, Osborne JR, Meyer R, Shi X, Tang Y, Koboldt DC, Lin L, Abbott R, Miner TL, Pohl C, Fewell G, Haipek C, Schmidt H, Dunford-Shore BH, Kraja A, Crosby SD, Sawyer CS, Vickery T, Sander S, Robinson J, Winckler W, Baldwin J, Chirieac LR, Dutt A, Fennell T, Hanna M, Johnson BE, Onofrio RC, Thomas RK, Tonon G, Weir BA, Zhao X, Ziaugra L, Zody MC, Giordano T, Orringer MB, Roth JA, Spitz MR (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069–1075
17. Ein-Dor L, Zuk O, Domany E (2006) Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc Natl Acad Sci USA* 103:5923–5928
18. Erler JT, Linding R (2010) Network-based drugs and biomarkers. *J Pathol* 220:290–296
19. Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature* 411:342–348
20. Gallegos Ruiz MI, Floor K, Roepman P, Rodriguez JA, Meijer GA, Mooi WJ, Jassem E, Niklinski J, Muley T, Van Zandwijk N, Smit EF, Beebe K, Neckers I, Ylstra B, Giaccone G (2008) Integration of gene dosage and gene expression in non-small cell lung cancer, identification of HSP90 as potential target. *PLoS One* 3:e0001722

21. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531–537
22. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100:57–70
23. Hecht SS (1999) Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst* 91:1194–1210
24. Henderson BE, Feigelson HS (2000) Hormonal carcinogenesis. *Carcinogenesis* 21:427–433
25. Herbst RS, Heymach JV, Lippman SM (2008) Lung cancer. *N Engl J Med* 359:1367–1380
26. Hirsch FR, Merrick DT, Franklin WA (2002) Role of biomarkers for early detection of lung cancer and chemoprevention. *Eur Respir J* 19:1151–1158
27. Hocking RR (1976) A Biometrics invited paper. The analysis and selection of variables in linear regression. *Biometrics* 32:1–49
28. Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991) p53 mutations in human cancers. *Science* 253:49–53
29. Hu K, Gan YH, Li SL, Gao Y, Wu DC, Wang CY, Yu GY (2009) Relationship of activated extracellular signal-regulated kinase 1/2 with lung metastasis in salivary adenoid cystic carcinoma. *Oncol Rep* 21:137–143
30. Imai MA, Oda Y, Oda M, Nakanishi I, Kawahara E (2004) Overexpression of E2F1 associated with LOH at RB locus and hyperphosphorylation of RB in non-small cell lung carcinoma. *J Cancer Res Clin Oncol* 130:320–326
31. Inoue J, Gohda J, Akiyama T, Semba K (2007) NF-kappaB activation in development and progression of cancer. *Cancer Sci* 98:268–274
32. Johansson R (1993) System modeling and identification. Prentice Hall, Englewood Cliffs
33. Khuder SA (2001) Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung Cancer* 31:139–148
34. Kishimoto M, Kohno T, Okudela K, Otsuka A, Sasaki H, Tanabe C, Sakiyama T, Hiramata C, Kitabayashi I, Minna JD, Takenoshita S, Yokota J (2005) Mutations and deletions of the CBP gene in human lung cancer. *Clin Cancer Res* 11:512–519
35. Komander D (2009) The emerging complexity of protein ubiquitination. *Biochem Soc Trans* 37:937–953
36. Kyriakis JM, Avruch J (2001) Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev* 81:807–869
37. Langenfeld EM, Kong Y, Langenfeld J (2006) Bone morphogenetic protein 2 stimulation of tumor growth involves the activation of Smad-1/5. *Oncogene* 25:685–692
38. Langer CJ (2004) Emerging role of epidermal growth factor receptor inhibition in therapy for advanced malignancy: focus on NSCLC. *Int J Radiat Oncol Biol Phys* 58:991–1002
39. Li Z, Zhao J, Du Y, Park HR, Sun SY, Bernal-Mizrachi L, Aitken A, Khuri FR, Fu H (2008) Down-regulation of 14-3-3zeta suppresses anchorage-independent growth of lung cancer cells through anoikis activation. *Proc Natl Acad Sci USA* 105:162–167
40. Liu H, Radisky DC, Nelson CM, Zhang H, Fata JE, Roth RA, Bissell MJ (2006) Mechanism of Akt1 inhibition of breast cancer cell invasion reveals a protumorigenic role for TSC2. *Proc Natl Acad Sci USA* 103:4134–4139
41. Luscombe NM, Babu MM, Yu H, Snyder M, Teichmann SA, Gerstein M (2004) Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 431:308–312
42. Mathivanan S, Periaswamy B, Gandhi TK, Kandasamy K, Suresh S, Mohmood R, Ramachandra YL, Pandey A (2006) An evaluation of human protein-protein interaction data in the public domain. *BMC Bioinform* 7(Suppl 5):S19
43. McLean GW, Carragher NO, Avizienyte E, Evans J, Brunton VG, Frame MC (2005) The role of focal-adhesion kinase in cancer—a new therapeutic opportunity. *Nat Rev Cancer* 5:505–515
44. Mitra SK, Schlaepfer DD (2006) Integrin-regulated FAK-Src signaling in normal and cancer cells. *Curr Opin Cell Biol* 18:516–523
45. Nibbe RK, Markowitz S, Myeroff L, Ewing R, Chance MR (2009) Discovery and scoring of protein interaction subnetworks discriminative of late stage human colon cancer. *Mol Cell Proteomics* 8:827–845

46. Nibbe RK, Koyuturk M, Chance MR (2010) An integrative -omics approach to identify functional sub-networks in human colorectal cancer. *PLoS Comput Biol* 6:e1000639
47. Oti M, Snel B, Huynen MA, Brunner HG (2006) Predicting disease genes using protein-protein interactions. *J Med Genet* 43:691–698
48. Pagano M, Gauvreau K (2000) Principles of biostatistics. Duxbury, Pacific Grove
49. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. *CA Cancer J Clin* 55:74–108
50. Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC, Yatish AJ, Kavitha MP, Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana R, Chandran S, Krishna S, Joy M, Anand SK, Madavan V, Joseph A, Wong GW, Schiemann WP, Constantinescu SN, Huang L, Khosravi-Far R, Steen H, Tewari M, Ghaffari S, Blobel GC, Dang CV, Garcia JG, Pevsner J, Jensen ON, Roepstorff P, Deshpande KS, Chinnaiyan AM, Hamosh A, Chakravarti A, Pandey A (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res* 13:2363–2371
51. Petrelli A, Giordano S (2008) From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Curr Med Chem* 15:422–432
52. Potter CJ, Pedraza LG, Xu T (2002) Akt regulates growth by directly phosphorylating Tsc2. *Nat Cell Biol* 4:658–665
53. Rapaport F, Zinovyev A, Dutreix M, Barillot E, Vert JP (2007) Classification of microarray data using gene networks. *BMC Bioinform* 8:35
54. Sala G, Dituri F, Raimondi C, Previdi S, Maffucci T, Mazzeletti M, Rossi C, Iezzi M, Lattanzio R, Piantelli M, Iacobelli S, Broggin M, Falasca M (2008) Phospholipase C γ 1 is required for metastasis development and progression. *Cancer Res* 68:10187–10196
55. Seber GAF, Lee AJ (2003) Linear regression analysis. Wiley-Interscience, Hoboken
56. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
57. Shi Y, Massague J (2003) Mechanisms of TGF- β signaling from cell membrane to the nucleus. *Cell* 113:685–700
58. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, Palma J, Brody JS (2004) Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci USA* 101:10143–10148
59. Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P, Sridhar S, Beamis J, Lamb C, Anderson T, Gerry N, Keane J, Lenburg ME, Brody JS (2007) Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 13:361–366
60. Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M (2006) BioGRID: a general repository for interaction datasets. *Nucl Acids Res* 34:D535–D539
61. Takahashi H, Ogata H, Nishigaki R, Broide DH, Karin M (2010) Tobacco smoke promotes lung tumorigenesis by triggering IKK β - and JNK1-dependent inflammation. *Cancer Cell* 17:89–97
62. Taylor IW, Linding R, Warde-Farley D, Liu Y, Pesquita C, Faria D, Bull S, Pawson T, Morris Q, Wrana JL (2009) Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol* 27:199–204
63. Tian Z, Hwang T, Kuang R (2009) A hypergraph-based learning algorithm for classifying gene expression and arrayCGH data with prior knowledge. *Bioinformatics* 25:2831–2838
64. Vander Griend DJ, D'antonio J, Gurel B, Antony I, Demarzo AM, Isaacs JT (2010) Cell-autonomous intracellular androgen receptor signaling drives the growth of human prostate cancer initiating cells. *Prostate* 70:90–99
65. Vanunu O, Magger O, Ruppin E, Shlomi T, Sharan R (2010) Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* 6:e1000641

66. Ventura JJ, Tenbaum S, Perdiguero E, Huth M, Guerra C, Barbacid M, Pasparakis M, Nebreda AR (2007) p38alpha MAP kinase is essential in lung stem and progenitor cell proliferation and differentiation. *Nat Genet* 39:750–758
67. Vicent S, Lopez-Picazo JM, Toledo G, Lozano MD, Torre W, Garcia-Corchon C, Quero C, Soria JC, Martin-Argarra S, Manzano RG, Montuenga LM (2004) ERK1/2 is activated in non-small-cell lung cancer and associated with advanced tumours. *Br J Cancer* 90:1047–1052
68. Vogelstein B, Kinzler KW (2004) Cancer genes and the pathways they control. *Nat Med* 10:789–799
69. Wang H, Davis A, Yu S, Ahmed K (2001) Response of cancer cells to molecular interruption of the CK2 signal. *Mol Cell Biochem* 227:167–174
70. Wang B, Liu K, Lin HY, Bellam N, Ling S, Lin WC (2010a) 14-3-3Tau regulates ubiquitin-independent proteasomal degradation of p21, a novel mechanism of p21 downregulation in breast cancer. *Mol Cell Biol* 30:1508–1527
71. Wang YC, Chen BS (2010) Integrated cellular network of transcription regulations and protein-protein interactions. *BMC Syst Biol* 4:20
72. Wang YC, Lan CY, Hsieh WP, Murillo LA, Agabian N, Chen BS (2010b) Global screening of potential *Candida albicans* biofilm-related transcription factors via network comparison. *BMC Bioinform* 11:53
73. Wang YC, Chen BS (2011) A network-based biomarker approach for molecular investigation and diagnosis of lung cancer. *BMC Med Genomics* 4:2
74. Welsh PL, Owens KN, King MC (2000) Insights into the functions of BRCA1 and BRCA2. *Trends Genet* 16:69–74
75. West KA, Linnoila IR, Belinsky SA, Harris CC, Dennis PA (2004) Tobacco carcinogen-induced cellular transformation increases activation of the phosphatidylinositol 3'-kinase/Akt pathway in vitro and in vivo. *Cancer Res* 64:446–451
76. Yang SK, Wang YC, Chao CC, Chuang YJ, Lan CY, Chen BS (2010) Dynamic cross-talk analysis among TNF-R, TLR-4 and IL-1R signalings in TNFalpha-induced inflammatory responses. *BMC Med Genomics* 3:19
77. Zajac-Kaye M (2001) Myc oncogene: a key component in cell cycle regulation and its implication for lung cancer. *Lung Cancer* 34(Suppl 2):S43–S46
78. Zimmermann GR, Lehar J, Keith CT (2007) Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 12:34–42
79. Zwick E, Bange J, Ullrich A (2001) Receptor tyrosine kinase signalling as a target for cancer intervention strategies. *Endocr Relat Cancer* 8:161–173

Chapter 2

The Impact of Next-Generation Sequencing Technology on Bacterial Genomics

Avantika Lal and Aswin Sai Narain Seshasayee

Abstract For many decades, genomic studies were based on Sanger sequencing or the dideoxy chain termination method of sequencing DNA, along with microarray and hybridization-based techniques to understand genome function. Sanger sequencing was used to sequence the genomes of many organisms, from bacteria to humans. However, in recent years ‘Next-generation’ sequencing technologies have been developed that are cheaper and far more rapid. They produce great sequencing depth, making them applicable to quantitative studies such as gene expression measurements as well. As a result, these technologies have been used extensively to study the sequence, structure, function and evolution of both eukaryotic and bacterial genomes. Here we discuss next-generation sequencing and how it has been used to study a variety of areas from gene expression and protein-DNA interactions to bacterial community function and evolution, at the scale of whole bacterial genomes. We expect that further advances in DNA sequencing technology and methods for managing and analyzing the large volumes of data produced by these approaches will help to answer many more questions in this field.

Keywords Next generation sequencing · DNA · Microarray · Hybridization · Chain termination · Sanger sequencing · Sequence library · Comparative genome hybridization (CGH)

A. Lal · A. S. N. Seshasayee (✉)
National Centre for Biological Sciences, Tata Institute of Fundamental Research,
Bangalore, India
e-mail: aswin@ncbs.res.in

A. Lal
e-mail: avantika@ncbs.res.in

2.1 Introduction

DNA sequencing is the determination of the order of the four nucleotide bases, adenine, guanine, cytosine, and thymine, in a molecule of DNA. Some of the earliest efforts to sequence DNA used laborious methods based on two-dimensional gel chromatography. However, in 1977, Sanger and Coulson described a much easier and more reliable method of DNA sequencing based on chain termination [90], which soon became widespread. In time, their original method was improved by automation and advances in technology and for the next 30 years, ‘Sanger sequencing’ held a monopoly over DNA sequence determination.

The ability to sequence DNA proved to be a turning point in biological sciences. It enabled scientists to understand the genetic basis of many diseases and to trace evolution at the molecular level, among other applications. The genomes, or total cellular DNA contents, of several organisms were sequenced using this method [10, 27, 32]. Over 1990–2004, the International Human Genome Project used Sanger sequencing methods, coupled with the whole genome shotgun technique, to sequence the approximately three billion nucleotides of the human genome [43]. Today DNA sequencing has become a vital tool not only in basic biological research but also in applied fields such as diagnostics and forensics.

As genome sequencing projects matured, it became apparent that further large-scale experimental tools were required to understand the meaning of genome sequences. Several tools were developed to study genomes at the functional level: from gene expression, which is the first stage in the conversion of DNA sequence to a functional readout, to protein-nucleic acid interactions, which enable gene expression. The DNA microarray, developed in the early 1990s, allowed scientists to measure the expression levels of large numbers of genes simultaneously or to genotype multiple regions of a genome. Later developments in DNA microarrays, in the form of genome-tiling arrays, permitted experimental annotation of genomes leading to various descriptions of pervasive transcription in eukaryotic genomes [7, 20], and large numbers of intra-operonic transcriptional initiation events in the simple bacterial genome of *Escherichia coli* [16]. These techniques further allowed large-scale mapping of regulatory networks by interrogating regions of the genome bound by a protein of interest, leading to the emergence of large-scale network biology (See [36, 38, 55, 100] for experiments in yeast; see [3, 4] for reviews); these findings unraveled unanticipated complexity in the binding properties of transcriptional regulators even in model bacterial genomes [34].

However, both Sanger sequencing and DNA microarrays have several drawbacks. The major drawback of Sanger sequencing is that it is too slow and expensive for many applications. The human genome project took over thirteen years and more than two billion dollars to complete. Given these limitations, the ultimate goal of genome sequencing, which for many is the sequencing of personal genomes leading to personalized medicine, is unlikely to be met using Sanger sequencing.

As a result, there has been a large effort in science and industry to bring the cost of high-quality human genome sequencing down to a level that is affordable to

individuals. In 2006, the X PRIZE Foundation announced a \$10 million incentive for the group whose technology would enable a human genome to be sequenced for \$1,000 or less. This has spurred the rise of a number of ‘next-generation’ sequencing (NGS) technologies in the last decade [91]. These produce large quantities of sequence data in a short period of time, and at a reduced cost. This has encouraged ambitious projects such as the 1,000 human genome project [84] and the 10,000 vertebrate genome project [69]. It has also become routine for large sequencing centers to publish a single piece of work describing over 100 bacterial genomes, thus enabling fine-scale genomic studies of bacterial epidemiology and evolution [19, 72].

The second cornerstone of genomic studies in the early 2000s, DNA microarray technology, has the limitation that it can interrogate the properties of only those regions of the genome whose sequence is known. For example, the first generation gene expression microarrays probed only known gene sequences, and were unable to detect transcription from intergenic regions. Further, microarrays cannot be used effectively to investigate the functional properties of non-model organisms whose genomes are not known. Though sequences of related genomes have been used to design microarray probes for the study of non-model organisms, the limitations of such an approach become apparent in the light of the fact that non-conserved portions of a genome have a large effect on an organism’s biology. Moreover, even small variations between the two genomes in the regions probed by the microarrays could lead to unreliable findings. Further, saturation in the measurement of fluorescence limits the dynamic range of microarrays.

The great depth of sequencing afforded by NGS techniques allow us to quantify nucleic acids, thus making them applicable to various applications for which DNA microarrays had been used, while circumventing many of the problems associated with the latter.

In this chapter, we discuss the features and challenges of next-generation sequencing technologies, as well as various types of experimental studies that have been enabled by them. We focus particularly on how these sequencing methods have been used to study bacterial genomics.

2.2 Next-Generation Sequencing Technologies

The term ‘Next-generation sequencing’ applies to several commercially available platforms. The most commonly used NGS platforms are 454 pyrosequencing (Roche/454 Life Sciences), Illumina (erstwhile Solexa sequencing) and SOLiD (Applied Biosystems).

Although these platforms differ from each other in the procedure employed as well as the chemistry of sequencing, their basic strategy is similar. The DNA molecule to be sequenced is fragmented at random positions, and short adaptor sequences are ligated to the ends of each DNA fragment. The resultant set of molecules is called the sequencing library. Each molecule in this library is amplified to generate a cluster of amplicons. Each cluster of identical DNA molecules is spatially separated from the

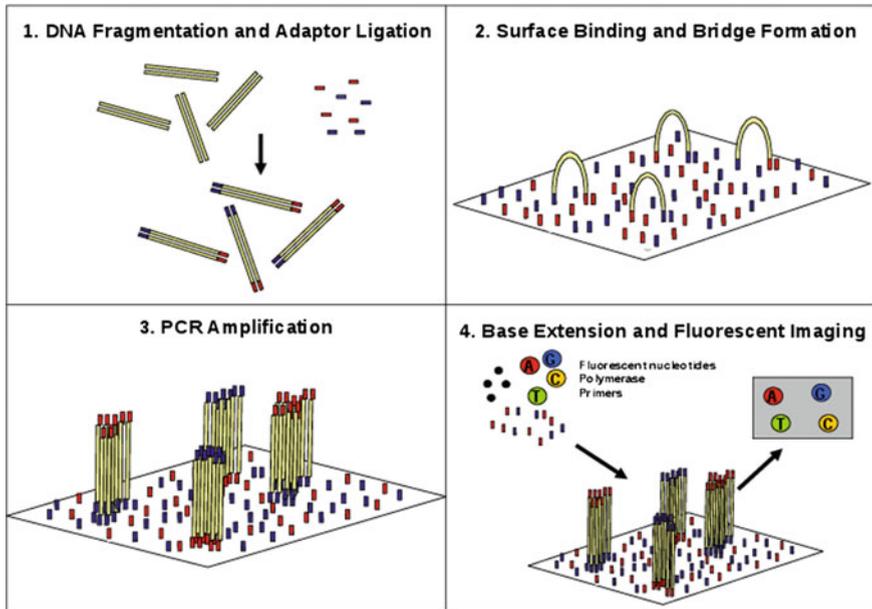


Fig. 2.1 The process of Illumina Sequencing. **1** The DNA to be sequenced is fragmented and short adaptor sequences are ligated to the ends of the fragments. **2** The DNA fragments are attached to a substrate coated with primers. **3** Bridge PCR amplifies each fragment to produce spatially separated clusters of fragments. **4** Fluorescently labeled nucleotides, polymerase and primers are added to synthesize complementary strands of DNA. After each step of nucleotide addition, the fluorescent signal from each cluster is read to generate the base sequence

others by tethering them to separate locations on a substrate. Primers complementary to the adaptor sequences at the end of each DNA fragment are added. The DNA molecules are then sequenced by extending these primers to produce DNA strands complementary to the template. Sequence data is acquired by imaging of the full array at the end of each cycle of nucleotide addition. These technologies have been described elsewhere [64, 107]. Figure 2.1 shows the process of DNA sequencing using Illumina technology.

The process of Sanger sequencing also begins with breaking the DNA molecule into smaller fragments. However, Sanger sequencing then requires each DNA fragment to be cloned into a vector and amplified in host cells (usually *E. coli*). This is time-consuming and expensive, and also, many cloned sequences are not stably maintained in the host [62]. In NGS, the sequencing library is constructed and amplified entirely in vitro, saving the trouble of cloning and colony picking.

Further, as the effective size of next-generation sequencing features can be in microns, millions of sequencing reads can be obtained in parallel by imaging of a small surface area. This makes the NGS strategy both faster and cheaper than Sanger sequencing. Also, because sequencing features are immobilized to a planar surface,

they can be enzymatically manipulated by a single reagent volume. This results in a much lower effective reagent volume per feature, and therefore lowers cost [93].

Perhaps the greatest drawback of NGS technologies is their lower read length (35–400 bp, depending on the platform) compared to Sanger sequencing (650–800 bp). The raw accuracy of sequencing is also currently lower for NGS than Sanger sequencing. However, it is expected that these problems will reduce with further improvements in technology.

2.3 NGS in the Study of Bacterial Communities, Evolution and Epidemiology

Bacteria are the most predominant form of free-living life on earth. Many bacteria are disease causing and understanding their evolution and function might help in developing intervention procedures. From a basic science standpoint, they are excellent systems for studying adaptation, both at the level of genome content and at the level of controlling gene expression and protein activity in response to changing conditions. In the following sections, we explore how recent literature has explored these aspects of bacterial biology using NGS technologies.

2.3.1 Genome Sequencing and Re-sequencing

Despite the increased speed and lower cost of sequencing, *de novo* sequencing of a genome using NGS is challenging due to the low read length of these instruments. This makes it difficult to assemble the sequenced fragments into a complete genome. The low read-length of the sequencing data is offset by great depth; a single lane of sequencing on an Illumina HiSeq 1000 sequencer, with a 12x multiplexing of samples, will provide 3Gb of sequence data per sample, thus giving a 1,000-fold coverage of an ‘average’ bacterial genome. Our experience working with genomes of *Staphylococcus aureus* shows that extremely stringent filtering of sequencing reads could easily give >200-fold coverage [81]. Compare this with the ~10-fold coverage that was typically achieved with Sanger sequencing projects, after laborious experimental work!

These unique characteristics of NGS data, namely an extremely large number of short reads with unique error characteristics, have led to the development of dedicated software, such as Velvet [119], SOAPdenovo [57], and ALLPATHS [11], for genome assembly. These methods have developed to such an extent that with sufficient coverage, short-read sequencing data can be used to produce first-pass mammalian genome assemblies that are comparable to those obtained with traditional Sanger sequencing [31].

Smith et al. [97] first used pyrosequencing to sequence the ~4Mb genome of the human pathogen *Acinetobacter baumannii*, the causative agent of several infections including pneumonia and meningitis. To overcome the limitation of short DNA reads (~100 nt on average), they obtained more than 21-fold coverage of the genome. More recently, Bos et al. [8] used Illumina to sequence DNA samples from victims of the 'Black Death' pandemic that spread through Europe in the fourteenth century, and reconstructed the ancient genome of the bacterium *Yersinia pestis* that was responsible for the pandemic.

A more common application of NGS is genome resequencing, i.e. sequencing the genome of a member of a species for which a reference genome is already available (Figure 2.2a). This is often done to catalog variations such as single nucleotide polymorphisms (SNPs), insertions and deletions relative to a reference genome, for example to identify those that might lead to interesting phenotypes. Before the advent of NGS, such studies were pursued using DNA microarrays, designed on the basis of one or more fully-sequenced reference genomes, following a technique dubbed comparative genome hybridization (CGH).

CGH is a tool to detect variations in DNA copy number between a test and reference genome. In this method, samples of genomic DNA from the test and reference cells are hybridized to microarrays. If, for instance, a gene from the reference genome is absent in the test organism, then the fluorescence intensity from its corresponding probe will be considerably less in the test than in the reference. For example, Willenbrock et al. [115] designed a microarray with probes covering the total content of 32 *E. coli* genomes to characterize novel *E. coli* strains based on their genomic content. McCarthy et al. [67] developed a 62-strain *S. aureus* microarray and used it to compare the genomes of different isolates of *S. aureus* from pigs and humans. They found that while the core genomes of these isolates did not vary much, the distribution of mobile genetic elements was variable and several mobile elements were host-specific. McCarthy et al. [68] used the same method to compare mobile genetic elements between 40 isolates of methicillin-resistant *S. aureus* from a hospital and found a diverse range of MGEs, virulence and resistance genes in the population.

The CGH technique, however, suffers from the drawback of interrogating only genomic regions present in the reference genome(s). Further, it is hard to distinguish between regions that are absent in the test genome and those that are merely divergent in the probed loci. Though these can be partially overcome by the adoption of a larger number of reference genomes in the design of the microarray, the process becomes quickly complicated, in particular for genomes with what is called an open pan-genome [101]. However, the advent of NGS has circumvented these problems, leading to several studies on the genetic variation and evolution of pathogens.

He et al. [42] used 454 and Sanger sequencing to sequence the genomes of thirty isolates of *Clostridium difficile*, which causes diarrheal disease. A phylogenetic analysis of these genomes suggested that both horizontal gene transfer and large-scale recombination played a significant role in the evolution of this species, and that virulence evolved independently in multiple lineages. In another such study,

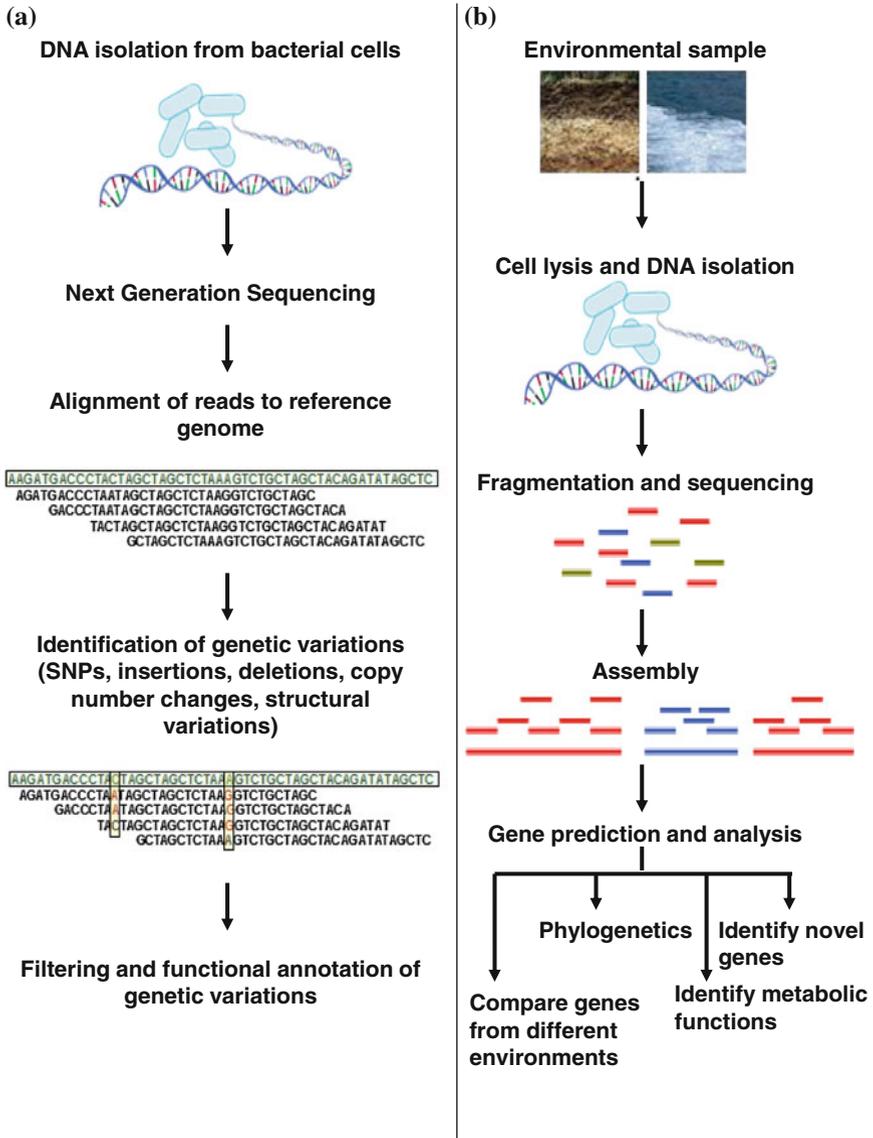


Fig. 2.2 a Identifying genome variations by resequencing. b Metagenomics

Mutreja et al. [72] studied the evolution of *Vibrio cholerae*, which causes millions of cases of cholera every year, and has caused seven recorded pandemics. They sequenced the genomes of 136 isolates of *V. cholerae* and identified SNPs to construct a phylogeny of this species. Their analysis suggested that the seventh cholera pandemic spread from the Bay of Bengal in at least three independent but overlapping

waves with a common ancestor in the 1950s, and identified several transcontinental transmission events. Similar genome-scale epidemiological studies have been carried out for *S. aureus* [39, 51, 66], and Casali et al. [12] sequenced 34 isolates of *Mycobacterium tuberculosis* to study the evolution of drug resistance in this pathogen. Such studies have given us valuable insights into the evolution and spread of human diseases, and may help us understand how to control epidemics in the future.

The ability to sequence large numbers of genomes also enabled studies in which the genomes of bacteria were sequenced over the course of long-term evolution experiments to identify regions that underwent evolutionary change. For example, Wielgoss et al. [114] sequenced 19 *E. coli* genomes from a 40,000-generation evolution experiment and estimated the point-mutation rate in the *E. coli* genome based on the accumulation of synonymous substitutions to be 8.9×10^{-11} per base pair per generation.

Other studies have examined the evolution of pathogens in their hosts. Yang et al. [118] studied the evolutionary dynamics of *Pseudomonas aeruginosa* as it adapted to its human hosts over 200,000 generations. These authors found that the population underwent limited genotypic diversification—most of which occurred early in the form of a few pleiotropic mutations followed by a landscape dominated by negative selection—despite the complex host environment. This was reported to be in contrast to *in vitro* studies, which documented continuous positive selection. Lieberman et al. [59] sequenced the genomes of 112 isolates of *Burkholderia dolosa* collected from human hosts over 16 years. They identified a set of genes that acquired nonsynonymous mutations in several individuals, suggesting that they experienced strong selection pressure during pathogenesis. These genes were involved in processes important for pathogenicity, such as antibiotic resistance and membrane biosynthesis, and might represent possible targets for therapy.

The acquisition of antibiotic resistance by bacterial pathogens is a growing problem, and NGS has been used to investigate this phenomenon. Zhang et al. [120] grew *E. coli* in the presence of the antibiotic ciprofloxacin and sequenced the genomes of the bacteria that survived, to identify mutations that gave rise to resistance. Whole-genome sequencing revealed that four single-nucleotide polymorphisms, including one in the *gyrA* gene encoding gyrase, were fixed in the resistant population. Similarly, Toprak et al. [102] grew *E. coli* with several different drugs and studied the evolution of resistance over 20 days. Sequencing the genomes of the resistant populations revealed mutations that conferred resistance to specific drugs and to multiple drugs.

2.3.2 Metagenomics

Metagenomics, or community genomics, is an approach to analyze the total genomic content of a microbial community. The total DNA from a population is isolated, sequenced and compared with previously known sequences. Metagenomic studies allow researchers to discover new species, and also to identify the types of biological

processes that occur in a specific environment. (Figure 2.2b) These methods permit bird's eye-level genetic characterization of unculturable bacteria, which represent most of the bacterial populations on the planet.

Early metagenomic explorations were based on painstaking Sanger sequencing experiments [30, 106, 111]. However, the main difficulty was presented by the sheer number of microbes in an environmental sample—it is estimated that there are $\sim 10^6$ bacterial species per gram of soil. The DNA of bacteria in the human gut, which has been a subject of keen interest among biologists and medics alike, can be expected to harbor much more diversity than the genome of its host. This represents a vast amount of DNA to sequence. Further, sampling rarer constituents of the microbial population requires great depth of sequencing, which is difficult to achieve with traditional sequencers. The high-throughput capability, relatively low cost and depth of next-generation sequencing makes such an approach much easier. Next-generation sequencers have been used to sequence the metagenome of diverse environments such as soil, oceanic communities, and the human gut.

Edwards et al. [26] used pyrosequencing to sequence genomes from two adjacent but chemically and geologically different sites in an iron mine in Minnesota. The microbial communities at the two sites were found to be functionally and metabolically different from each other, in pathways such as carbon utilization, iron acquisition, nitrogen assimilation, and respiration. Dinsdale et al. [22] used pyrosequencing to compare microbial and viral DNA sequences from nine biomes including marine, freshwater, subterranean, and host-associated, and found strongly discriminatory metabolic profiles across different environments.

It is estimated that up to 100 trillion microbial cells reside in an average human body [5]. Most of these microbes are present in the gut, where they are thought to influence human physiology, nutrition, and health [17, 87]. To understand and exploit the functioning of the gut microbial community it is necessary to understand its content and diversity. Qin et al. [83] used Illumina to sequence the total genomic DNA from faecal samples of 124 European individuals. They assembled and characterized 3.3 million microbial genes from 576.7 Gb of sequence, and found that each individual harbored at least 160 bacterial species, which were largely common across individuals. Further, they compared samples from healthy individuals and inflammatory bowel disease patients, and showed that the gut microbiomes of the two groups differed in terms of overall bacterial diversity as well as the relative abundances of various species.

'Metatranscriptomic' studies aim to sequence the total RNA expressed by microbes in an environmental sample, instead of DNA. This is particularly interesting as bacteria have several untranslated small RNAs that regulate environmentally important processes, including amino acid biosynthesis, starvation responses, and quorum sensing [99]. Because studies on sRNAs have focused on a few model microorganisms, the diversity and ecological function of sRNAs in natural communities is little understood. Compared to protein-coding sequences, these are also difficult to identify from DNA sequences. Shi et al. [94] analyzed metatranscriptomic data sets from ocean water and found that a large fraction of cDNA sequences detected comprised small RNAs. They also identified several new classes of putative sRNAs.

Metagenomics has given scientists access to unculturable microbial communities and their activities in a wide variety of systems. Though this approach can tell us about the species and functions present in a microbial community, it is very difficult to assemble individual genomes in the community from such data. However, Iverson et al. [45] were able to reconstruct a nearly complete genome of the uncultured marine group II *Eurarchaeota* entirely *de novo* from the metagenome of seawater samples.

2.3.3 Single-Cell Genomics

The various limitations of metagenomics, particularly difficulties in accessing rare components of the microbiota and ability to map genes to individual species or organisms can be overcome by single-cell genome sequencing. Methods have now been developed to isolate single cells and amplify DNA for sequencing. The higher throughput of NGS makes it possible to finish several single-cell genomes in a reasonable time, and single-cell sequencing has now been applied to many environmental microbes.

Marcy et al. [63] developed a microfluidic device for isolating single cells and amplifying their genomes, and used it to isolate bacteria of a little-understood phylum from the human oral microbiota. They were able to assemble the sequences of over 1,000 genes by pyrosequencing, providing insight into the physiology of members of this phylum. Woyke et al. [116] isolated DNA from individual cells of two marine flavobacteria from the Gulf of Maine that were phylogenetically distant from existing cultured strains. With a combination of pyrosequencing and Sanger sequencing they recovered 91% and 78% of the two genome sequences, and analyzed the genome content, metabolic adaptations, and biogeography of these taxa.

2.4 Studying Genome Function

2.4.1 Gene Expression Control in Bacteria

Bacteria do not express all the genes present on their genome all of the time. Instead, they produce those gene products that are important for them to survive in the environment they face. Bacteria may accomplish this by modulating any step of a gene's expression and protein activity, from transcription to translation to post-translational modification of a protein, though they most commonly control gene expression by regulating the level of transcription.

Transcription is regulated at multiple levels, including: (a) variations in the sequence of the promoter to which RNA polymerase, the enzyme responsible for transcription, binds, thus ensuring that different genes have different inherent ability to be transcribed; (b) three-dimensional topology of the DNA where DNA supercoiling controls many DNA transactions including transcription, by controlling the extent to which the DNA is unwound; (c) sigma factors, which tightly associate with the

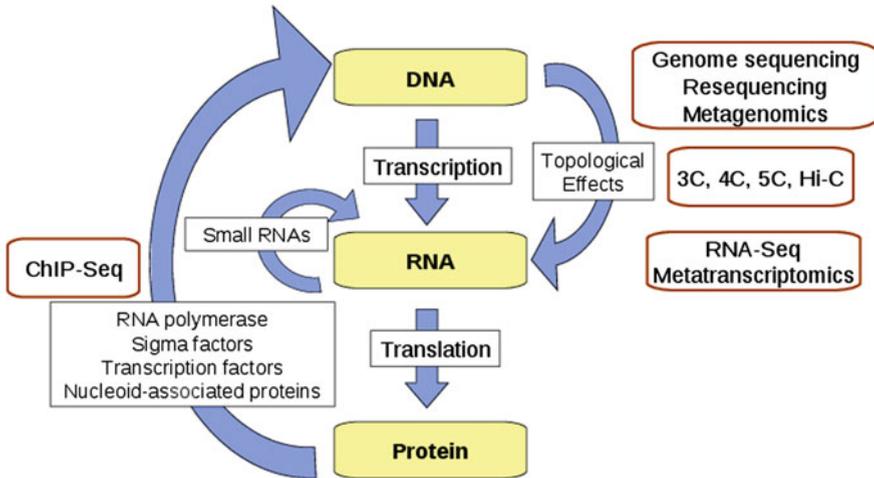


Fig. 2.3 NGS—based techniques are used to understand bacterial genomics at multiple levels

RNA polymerase and recognize promoter sequences; (d) transcription factors, which are proteins that alter the affinity of a promoter to the RNA polymerase by binding close to the promoter and either enabling or blocking access of the promoter to the RNA polymerase; (e) various small molecule and RNA-based ligands that bind to components of the RNA polymerase thus altering its availability to transcribe genes; (f) DNA modifications such as methylation.

Details of these regulatory mechanisms have been reviewed elsewhere [9, 44, 73]. Here we review research that has investigated the above aspects of gene regulation in bacteria using techniques based on NGS (See Figure 2.3). Where applicable, we point to relevant research based on microarrays indicating that similar work can be pursued using sequencing as well.

2.4.2 Describing Transcriptomes

The transcriptome of a cell is the total content of RNA transcripts expressed in the cell at a given time. Studying the transcriptome of a bacterium helps us to understand how it responds to different environmental conditions. For example, Nicolas et al. [74] investigated the transcriptomes of *Bacillus subtilis* grown under 104 conditions (nutrients, aerobic and anaerobic growth, stresses, etc.) that the bacterium might encounter in nature.

Before the development of NGS, most methods to study gene expression required hybridization of specific oligonucleotides to the loci of interest—either primers binding to complementary cDNA in quantitative reverse transcription polymerase chain reactions (qRT-PCR), or labeled probes binding to RNA in Northern blotting, or hybridization of cDNA to probes on microarray chips. Of these, only DNA

microarrays had a high throughput. However, sequencing of RNA by next-generation sequencing (RNA-seq) is different, in that fragments of DNA are matched to genes by sequence alignment instead of physical hybridization. (See [29] for a review of bacterial transcriptomics). These avoid problems typical of microarrays, such as (a) a limitation to interrogate only regions of the chromosome that have been used as probes; (b) saturation of fluorescence signals that limit the dynamic range; (c) background noise in fluorescence; (d) artifacts due to probe characteristics such as base composition.

Global transcriptional analyses using RNA-Seq have been carried out for many bacteria. Güell et al. [35] analyzed the transcriptome of one of the smallest self-replicating organisms, *Mycoplasma pneumoniae*, using a combination of microarrays and sequencing. They found that even this simple bacterium has several antisense transcripts, alternative transcripts, and multiple regulators per gene. This suggests a dynamic and regulated transcriptome, more similar to that of eukaryotes than was previously thought.

However, it is not possible to directly locate the transcription start sites of genes using RNA-Seq, as it is not possible to distinguish between primary transcripts and processed transcripts. Sharma et al. [92] used a differential or dRNA-seq approach to discriminate these two. Primary transcripts have a 5' triphosphate group, whereas processed transcripts such as ribosomal and transfer RNAs have a 5' monophosphate. They carried out 454 pyrosequencing of two cDNA libraries—one prepared from untreated total bacterial RNA, and the other enriched for primary transcripts by treatment with an exonuclease that degrades 5' monophosphate but not 5' triphosphate RNA. They were able to map ~217 million bases of cDNA to the *H. pylori* genome, and construct a genome-wide map of *H. pylori* transcriptional start sites and operons. They discovered hundreds of transcriptional start sites within operons, and opposite to annotated genes, indicating that antisense transcription takes place throughout the genome. They also discovered ~60 small RNAs of different classes.

RNA-Seq involves reverse transcribing single-stranded RNA into double stranded cDNA, which is then sequenced. The result is the sequences of both DNA strands of the gene that encodes the RNA. Hence RNA-Seq does not tell us which strand of the DNA is transcribed into RNA, which is important to resolve overlapping genetic features and detect antisense transcription. However, there are methods to identify the directionality of transcription. These generally involve modifying the RNA molecules before reverse transcription, or modifying the first cDNA strand before the synthesis of its complementary strand. Croucher et al. [18] developed strand-specific cDNA sequencing, in which they reverse transcribed the RNA into only one strand of cDNA and directly sequenced the library of single-stranded cDNA molecules. Perkins et al. [79] used this to analyze the transcriptome of *Salmonella enterica* serovar Typhi in a strand-specific manner. This allowed them to identify many transcribed regions within prophages, pseudogenes, and UTRs of other genes.

As no pre-existing knowledge of the RNA sequence to be detected is necessary, RNA-Seq has been especially useful for discovering new species of RNA. Sitka et al. [96] carried out a transcriptome analysis in *Salmonella enterica* serovar Typhimurium, and also sequenced the small RNAs associated with the regulatory

protein Hfq. They discovered several novel small RNAs, and found that Hfq regulates the expression of nearly one-fifth of all *Salmonella* genes, including several pathogenicity islands.

Mraheil et al. [71] sequenced cDNAs from RNA less than 500 nucleotides long, in the intracellular pathogen *Listeria monocytogenes*, during intracellular and extracellular growth. They discovered 150 putative regulatory RNAs, including 29 that were expressed only during intracellular growth. Some of these were found to be required for efficient intracellular growth and infection by this pathogen. Lasa et al. [54] sequenced long and short (<50 nt) RNAs from *S. aureus*. They found short RNAs that were produced by RNase III digestion of double stranded RNAs formed by overlapping sense and antisense transcripts throughout the genome. This suggested that antisense transcription is used to suppress expression of some genes by producing double-stranded RNA that is degraded.

2.4.3 Promoter Sequences and Their Affinity to RNA Polymerase

Cho et al. [16] performed a ChIP-chip experiment—in which fragments of DNA bound to a protein of interest are isolated using an antibody and the resulting DNA fragment hybridized to a microarray—for RNA polymerase in the presence of the antibiotic rifampicin which blocks transcription elongation. This was shown to provide a static picture of RNA polymerase occupancy at promoters. Though this was used primarily to define promoters and transcription start sites, one can envisage these data being used as a measure of the inherent affinity of a promoter to the RNA polymerase.

2.4.4 Structure of the DNA and Its Effect on Transcription

The three-dimensional structure of the bacterial genome both reflects and regulates the functional state of the cell. However, until recently it had not been possible to study the three-dimensional conformation of the chromosome on a genomic-scale with high resolution. Recent techniques under the general category of “chromosome conformation capture” (3C [21], 4C [95], 5C [25] and Hi-C [60]), build interaction maps in which spatially proximal regions of the chromosome are linked together. This network is subsequently used to build a three-dimensional model of the chromosome.

Umberger et al. [105] used 5C to construct a 3D model of the *Caulobacter crescentus* genome in wild and genetically modified strains. They found the chromosome to be ellipsoidal with periodically arranged arms, and identified a short region of the genome that affected the orientation of the entire chromosome.

An important topological property of the chromosome that affects transcription is supercoiling. Though it has not been possible to define local supercoiling at high resolution on a genomic scale yet, Peter et al. [80] used DNA microarrays to identify genes that respond to perturbations to the global supercoiling levels. They found that

negative supercoiling activates expression of ~200 genes in the *E. coli* chromosome, while repressing that of ~100. Genes that were activated by negative supercoiling tended to have higher G/C content than average, whereas the opposite was true of those that were repressed. This has potential implications for transcription during stationary phase.

2.4.5 *Sigma Factors and Transcription*

Bacterial RNA polymerase is a multisubunit enzyme. The core RNA polymerase, composed of five subunits, is capable of transcribing DNA. However, this core polymerase is not capable of binding tightly and specifically to promoter sequences. This ability is conferred by the sigma subunit [9].

The number of sigma subunits varies between bacterial species. *E. coli* has seven, of which RpoD is the 'housekeeping' sigma factor that transcribes most of the cellular genes in growing cells. The other sigma subunits are activated under different environmental conditions and direct the transcription of genes needed to survive in those conditions. For example, RpoH is activated under heat stress and transcribes chaperones and other genes involved in the heat stress response. Nicolas et al. [74] examined the transcriptome of *B. subtilis*, and concluded that approximately 66% of the variance in gene expression between different environmental conditions can be explained by variation in the expression of different sigma factors.

Wade et al. [109] used ChIP-chip to identify the genomic binding sites of RpoH in *E. coli*. Interestingly, a quarter of the RpoH targets were found within coding regions. Also, most of the targets overlapped with those of RpoD, suggesting extensive overlap between the functions of different sigma factors. However this result remains controversial [110]. Patten et al. [78] and Weber et al. [113] used microarrays to compare gene expression of wild-type *E. coli* with a strain lacking a functional RpoS sigma factor. This sigma factor controls the expression of many stationary-phase genes. They identified hundreds of genes that are regulated by RpoS during the transition into stationary phase.

NGS can be a powerful tool to study the functions of sigma factors in detail. Recently, Dong and Mekalanos [24] used ChIP-Seq and RNA-Seq to define the regulon of the alternative sigma factor RpoN in *Vibrio cholerae*. They identified a consensus sequence for RpoN binding and showed that RpoN regulates the expression of flagellar genes and secreted proteins.

2.4.6 *Global Transcription Factors*

Many crucial processes in the cell, including chromosome organization, replication, and regulation of gene expression are orchestrated through the interaction of proteins with their binding sites on the bacterial genome. Understanding these processes on a global scale requires mapping of protein-DNA interactions across the entire genome.

Vora et al. [108] profiled the binding sites of all proteins on the entire *E. coli* genome using a technique they named *in vivo* protein occupancy display (IPOD). They isolated protein-DNA complexes and used DNA microarrays to identify protein-bound domains on the *E. coli* chromosome. They found extensive (longer than 1 kb) protein occupancy domains (EPODs), many of which were located in highly curved and transcriptionally silent regions of the genome. They suggested that these EPODs bind nucleoid-associated proteins and act as organizing centers that isolate the domains of the chromosome. However, this technique does not allow identification of the protein bound at each site, which has been addressed with ChIP (described earlier).

As with transcriptomes, ChIP studies were first carried out on a genomic scale using DNA microarrays. However, in a manner similar to RNA-Seq, ChIP followed by NGS (ChIP-Seq) can produce significantly better data than microarrays, including single base-pair resolution when appropriately modified (ChIP-exo), lower noise, a larger dynamic range, and of course it is not limited by fixed probe sequences. Although the short reads (~35 bp) of NGS platforms are disadvantageous for applications like *de novo* genome assembly, they are acceptable for ChIP-Seq.

ChIP-Seq has been used to study the DNA binding of many proteins that are involved in global gene regulation, including RNA polymerase, nucleosomes and transcription factors, in eukaryotes [6, 13] and to a much lesser extent in prokaryotes.

Bacterial transcription factors can be broadly classified into local and global transcription factors. Local transcription factors regulate the expression of a small number of genes, generally directed toward a single cellular function. On the other hand, global transcription factors regulate a large number of genes which belong to multiple functional categories, act under several different environmental conditions, and bind extensively to chromosomal DNA. For example, there are 187 known transcription factors in *E. coli*, but nine of these regulate 63% of the target genes and are responsible for 52% of the regulatory interactions [89]. Seven of these are considered to be global transcription factors: CRP, FNR, ArcA, LRP, FIS, IHF and H-NS [65].

Genome-wide approaches are necessary to understand the function of global transcription factors, as these proteins influence transcription across the entire genome. Grainger et al. [34] used ChIP-chip and microarrays to study the binding pattern of the global transcription factor CRP across the *E. coli* genome and identify its regulated genes. They found that while CRP does not have many strong binding sites (around 70), it binds to several weaker sites throughout the genome. Cho et al. [14] carried out ChIP-chip to identify binding sites of the *E. coli* transcription factor Lrp and RNA polymerase along with a comparison of gene expression in wild-type and an Lrp knockout. They showed that Lrp regulates a large number of genes involved in diverse functions and has three modes of regulation at different promoters.

More recently, NGS techniques have been used to study transcription factor function. Kahramanoglou et al. [50] used ChIP-Seq to map the binding sites of the nucleoid-associated proteins H-NS and Fis throughout the *E. coli* chromosome. These proteins were previously studied using ChIP-chip [15, 33, 76]. Fis affects the expression of over 20% of all genes, mostly by increasing transcription. However, there was little correlation between Fis binding regions and regions where the

protein affected transcription. But, regions where such a correlation existed were characterized by multiple Fis binding sites in operon-upstream regions. This in turn was correlated with high A/T content of the binding region and possible DNA bending. H-NS binds to longer stretches of DNA, and shows mild or strong repression of its target genes depending on the length of the binding site. Prieto et al. [82] identified the binding sites of HU and IHF in *E. coli* on a genome-wide scale using ChIP-seq, along with microarray analysis of gene expression in single- and double-deletion mutants of each protein. They found that the binding of IHF was sequence specific and included ~30% of all operons in the genome, demonstrating its role as a global regulator. HU was seen to bind non-specifically to the chromosome, though with a preference for A/T-rich DNA.

2.4.7 DNA Modifications and Transcription

Methylation of the nucleotides of DNA is another important means of regulation in a cell. In bacteria, the Dam methylase adds a methyl group to the adenine residue in the sequence 5'-GATC-3'. Dam methylation regulates DNA replication, mismatch repair and transcription in bacteria, by modulating protein-DNA binding. A microarray study by Robbins-Manke et al. [85] found an upregulation of over 200 genes in the absence of Dam in *E. coli*. They suggested that this might be due to Dam changing the binding sites of transcription factors and RNA polymerase and hence modulating the binding of these proteins.

A second DNA methylase in *E. coli* is DNA Cytosine Methylase or Dcm, which methylates the internal cytosine in CCWGG sequences. Kahramanoglou et al. [49] carried out bisulfite sequencing of *E. coli* genomic DNA using Illumina to identify sites of cytosine methylation. They also compared gene expression of wild-type *E. coli* with a Dcm knockout strain and identified over 500 differentially expressed genes. Methylation by Dcm progressively increases from exponential to stationary phase, and Dcm may also regulate the stationary phase sigma subunit RpoS. Another study [70] had previously shown that Dcm regulates the expression of ribosomal proteins in stationary phase.

2.5 Computational Challenges of Next-Generation Sequencing

Because of the higher parallelism and lower cost of sequencing, the widespread use of NGS in biology has made massive amounts of sequence data available. Also, as single-cell genomics becomes more widely used, it is likely that even greater amounts of genomic data will become available. The large datasets produced by NGS experiments require large amounts of storage space. Apart from the sequence data itself, an NGS experiment initially produces terabytes of raw image files. Once base calling is done to convert these images of fluorescent light into DNA sequences,

these images are discarded due to lack of storage space. It would be an interesting challenge to store and use these images, possibly to improve base calling algorithms. Several tools and algorithms have been designed to analyze data generated by Next-Generation Sequencing experiments (See Table 2.1).

2.5.1 Reference Mapping

For genome resequencing as well as techniques like ChIP-Seq and RNA-Seq, the short sequence reads obtained from NGS have to be mapped back to their position on a reference genome. This remains a challenge for large and complex genomes like that of humans. Commonly used alignment algorithms like BLAST have drawbacks for this application. Primarily, NGS data may contain many millions of short reads, which BLAST would be very slow to align [62]. Therefore there was a need for methods that were designed to work with short sequences and save time by operating on compressed data.

The Burrows-Wheeler Transform is an algorithm used to permute the order of characters in a sequence, which allows sequence data to be greatly compressed. This technique also allows searching for subsequences in the original sequence while operating on the compressed file, making this technique suitable for mapping NGS data to reference genomes. BWA [58] and BOWTIE [53] are successful short-read aligners which are based on this technique.

2.5.2 Genome Assembly

De novo genome assembly is done by piecing together sequence fragments to join them into contigs or contiguous sequences. Many algorithms designed for whole-genome sequence assembly from Sanger sequencing data use an approach of representing each read as a node and each overlap between sequences as an arc between the two nodes. However, NGS data contains a much larger number of very short reads. The number of reads would make such an overlap graph very large and difficult to compute [119].

A different approach is based on de Bruijn graphs, which are not based on reads, but on k -mers (words that are k nucleotides long). Reads are mapped as paths through the graph, going from one word to the next word in a determined order. Velvet is a short read assembler based on de Bruijn graphs that has been used to assemble bacterial genomes [119].

2.5.3 Analysis of ChIP-Seq Data

When ChIP-Seq is carried out for global transcription factors, sequences over the entire genome are sampled. These sequences are mapped back to their position on a

Table 2.1 Computational tools to analyze NGS data

Category	Program	Type	Reference
de novo Genome assembly	1. Velvet	1–3. De Bruijn assemblers	1. Zerbino and Birney (2008) [119]
	2. ALLPATHS	4–6. Greedy extension assemblers	2. Butler et al. (2008) [11]
	3. SOAPdenovo		3. Li et al. (2010) [61]
	4. SSAKE	7. Overlap-layout-consensus assembler	4. Warren et al. (2007) [112]
	5. SHARCGS		5. Dohm et al. (2007) [23]
	6. VCAKE		6. Jeck et al. (2007) [46]
	7. Edena		7. Hernandez et al. (2008) [41]
Reference mapping	1. MAQ	1–4. Burrows-Wheeler transform based	1. Li et al. (2008) [56]
	2. BWA		2. Li and Durbin (2009) [58]
	3. Bowtie	5. Hash table based	3. Langmead et al. (2009) [53]
	4. SOAP		4. Li et al. (2008) [57]
	5. SSAHA		5. Ning et al. (2001) [75]
RNA-Seq	1. Scripture	1. Transcriptome reconstruction.	1. Guttman et al. (2010) [37]
	2. Cufflinks	2. Transcript assembly, estimation, differential expression testing	2. Trapnell et al. (2010) [104]
	3. TopHat		3. Trapnell et al. (2009) [103]
ChIP-Seq	4. SpliceMap	3, 4. Splice junction discovery	4. Au et al. (2010) [2]
	1. SISSRS	1–4. Peak identification	1. Jothi et al. (2008) [48]
	2. MACS		2. Zhang et al. (2008) [121]
	3. BayesPeak	5, 6. Visualization, peak detection, gene-peak association	3. Spyrou et al. (2009) [98]
	4. PeakSeq		4. Rozowsky et al. (2009) [88]
	5. CisGenome		5. Ji et al. (2008) [47]
	6. CASSys		6. Alawi et al. (2011) [1]
Conformation capture	1. my5C	A web tool for design, visualization and analysis of 5C studies	1. Lajoie et al. (2009) [52]

reference genome. After reference mapping, the next step in ChIP-Seq data analysis is peak calling, i.e. identifying the genomic regions that produced a relatively high number of sequence reads, indicating that they are the binding sites of the protein of interest. However, ChIP-Seq data often contains a high amount of noise, and a major challenge in analysis is how to distinguish peaks from background noise.

One way to reduce this problem is to perform a ‘mock IP’ or control experiment in which the procedure is carried out without using specific antibodies [77]. If a peak in the experimental data co-localizes with a peak in the control data, it would lower the likelihood of having detected a binding site. Several algorithms have been developed to carry out peak calling for Chip-Seq data, including BayesPeak [98] and MACS [121].

2.5.4 Analysis of RNA-Seq Data

In an RNA-Seq experiment, the abundance of a gene in the sample is measured by the number of reads that map to that gene. The number of reads for a gene in the raw sequence data is generally normalized by the length of a gene and by the total number of reads. However, the number of reads of a given gene depends not only on its abundance, but also on the abundance of other genes. For example, if an RNA sample includes a very highly expressed gene, most of the reads may be taken up by its transcripts, leaving very little sequencing space for less abundant transcripts. The number of reads for a transcript in such conditions may not reflect its actual abundance, and this can be a difficulty in comparing transcriptomes of different samples. Robinson and Oshlack [86] proposed a normalization method using a weighted trimmed mean of the log expression ratio to estimate the ratio of RNA production.

2.5.5 Chromosome Conformation

Yaffe et al. [117] studied several sources of experimental bias in Hi-C experiments. These were: (1) Ligation occurs between nonspecific cleavage sites as well as restriction fragment ends. (2) The efficiency of ligation of restriction fragments may depend on their length. (3) The GC content near the ends of the ligated fragments may influence the processing and sequencing of the DNA. (4) The mappability or genomic uniqueness of the fragment ends affects the estimated probability of contact between sequences.

There is a need to minimize these experimental biases and incorporate them into algorithms to analyze chromosome conformation capture data.

2.6 The Future of Sequencing Technologies

Just as NGS technologies did in the last decade, a number of emerging DNA sequencing technologies promise to influence the study of genomes in the coming years. A major advance in this field is single-molecule sequencing, which allows the sequence of a single molecule of DNA to be read without the need for any amplification step at all.

One example of a single-molecule sequencing technology is Single molecule real time sequencing (SMRT). In this, a single DNA polymerase enzyme is presented with a template DNA molecule and substrate nucleotides. Each of the four DNA bases is attached to a different fluorescent dye. The fluorescent signal of the incorporated nucleotides is read to generate the sequence. SMRT data was first published by Harris et al. [40], who used this method to resequence the M13 phage genome. This method can also be used to detect DNA methylation, and, in theory, other DNA modifications [28].

Another promising recent method is nanopore sequencing. A nanopore is a hole with diameter of the order of 1 nm. When a nanopore is immersed in a conducting fluid and a potential is applied across it, an electric current passes through it. The DNA to be sequenced is forced through the pore one base at a time. As it does so, each nucleotide of the DNA obstructs the nanopore. The degree of obstruction, and hence the amount of current through the nanopore, varies depending on whether the nucleotide blocking the nanopore is an A, C, G or T. The change in current through the nanopore as the DNA passes through it can therefore be used to directly read the DNA sequence.

At present, Oxford Nanopore Technologies is developing a commercial nanopore sequencing system. They have recently announced that they have used this technology to sequence the 5.4 kb genome of the Phi X phage in one continuous read.

2.7 Lessons Learnt

1. Next-Generation technologies to sequence DNA are significantly faster and cheaper than Sanger sequencing. However, they have lower read length and accuracy of sequencing.
2. NGS technologies have made a great difference to the field of bacterial genomics by allowing whole genomes and large sets of DNA to be sequenced at a reasonable speed and cost. This has allowed sequencing of many bacterial genomes and metagenomic sequencing of DNA from whole bacterial communities. Ecological and evolutionary studies based on NGS have given us valuable insights.
3. Bacteria respond to changing environmental conditions by regulating the expression of their genes at many levels. Methods of gene regulation used by bacteria include DNA topology, promoter sequences, sigma subunits, transcription factors, RNA polymerase ligands and small RNAs.

4. NGS has enabled studies of gene regulation on a global scale in bacteria, for example by mapping the topology of the bacterial chromosome, studying the transcriptome of bacteria under different conditions, and by mapping sites at which transcription factors bind to the DNA.
5. NGS leads to the generation of vast amounts of sequence data. The storage and analysis of this data is difficult. Several algorithms have been developed specifically for analyzing NGS data.
6. Several new and more advanced techniques for DNA sequencing are being developed, which may lead to further and more interesting developments in the field of genomics.

Acknowledgments We thank Dr. Dasaradhi P. for proofreading the manuscript.

References

1. Alawi M, Kurtz S, Beckstette M (2011) CASSys: an integrated software-system for the interactive analysis of ChIP-seq data. *J Integr Bioinf* 8(2):155
2. Au KF, Jiang H, Lin L, Xing Y, Wong WH (2010) Detection of splice junctions from paired-end RNA-seq data by splicemap. *Nucleic Acids Res* 38(14):4570–4578
3. Babu MM, Teichmann SA, Aravind L (2006) Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol* 358(2):614–633
4. Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA (2004) Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* 14(3):283–291
5. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. *Science* 307(5717):1915–1920
6. Baski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129(4):823–837
7. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M (2004) Global identification of human transcribed sequences with genome tiling arrays. *Science (New York, N.Y.)* 306(5705):2242–2246
8. Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S, Wood J, Earn DJD, Herring DA, Bauer P, Poinar HN, Krause J (2011) A draft genome of *Yersinia pestis* from victims of the black death. *Nature* 478(7370):506–510
9. Browning DF, Busby SJ (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2(1):57–65
10. Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb J-F, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kaine BP, Borodovsky M, Klenk H-P, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *methanococcus jannaschii*. *Science* 273(5278):1058–1073
11. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB (2008) ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res* 18(5):810–820

12. Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, Bentley SD, Parkhill J, Nejentsev S, Hoffner SE, Horstmann RD, Brown T, Drobniewski F (2012) Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res* 22(4):735–745
13. Chen X, Han X, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh Y-HH, Yeo HCC, Yeo ZXX, Narang V, Govindarajan KRR, Leong B, Shahab A, Ruan Y, Bourque G, Sung W-KK, Clarke ND, Wei C-LL, Huck-Hui HNg (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6):1106–1117
14. Cho B-K, Barrett CL, Knight EM, Park YS, Bernhard (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci* 105(49):19462–19467
15. Cho B-K, Knight EM, Barrett CL, Bernhard (2008) Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res* 18(6):900–910
16. Cho B-KK, Zengler K, Qiu Y, Park YSS, Knight EM, Barrett CL, Gao Y, Palsson BØ (2009) The transcription unit architecture of the *Escherichia coli* genome. *Nat Biotechnol* 27(11):1043–1049
17. Clemente JC, Ursell LK, Parfrey LW, Knight R (2012) The impact of the gut microbiota on human health: an integrative view. *Cell* 148(6):1258–1270
18. Croucher NJ, Fookes MC, Perkins TT, Turner DJ, Marguerat SB, Keane T, Quail MA, He M, Assefa S, Bähler J, Kingsley RA, Parkhill J, Bentley SD, Dougan G, Thomson NR (2009) A simple method for directional transcriptome sequencing using Illumina technology. *Nucleic Acids Res* 37(22):e148
19. Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, McGee L, von Gottberg A, Song JHH, Ko KSS, Pichon B, Baker S, Parry CM, Lambertsen LM, Shahinas D, Pillai DR, Mitchell TJ, Dougan G, Tomasz A, Klugman KP, Parkhill J, Hanage WP, Bentley SD (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science (New York, N.Y.)* 331(6016):430–434
20. David L, Huber W, Granovskaia M, Toedling J, Curtis PJ, Bofkin L, Jones T, Ronald DW, Lars SM (2006) A high-resolution map of transcription in the yeast genome. *Proc Natl Acad Sci* 103(14):5320–5325
21. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* 295(5558):1306–1311
22. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F (2008) Functional metagenomic profiling of nine biomes. *Nature* 452(7187):629–632
23. Dohm JC, Lottaz C, Borodina T, Himmelbauer H (2007) SHARCGS, a fast and highly accurate short-read assembly algorithm for de novo genomic sequencing. *Genome Res* 17(11):1697–1706
24. Dong TG, Mekalanos JJ (2012) Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52. *Nucleic Acids Res* 40(16):7766–7775
25. Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J (2006) Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* 16(10):1299–1309
26. Edwards R, Brito BR, Wegley L, Haynes M, Breitbart M, Peterson D, Saar M, Alexander S, Alexander EC, Rohwer F (2006) Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* 7(1):57+
27. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269(5223):496–512
28. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 7(6):461–465

29. GÅell M, Yus E, Lluch-Senar M, Serrano L (2011) Bacterial transcriptomics: what is beyond the RNA horix-ome? *Nat Rev Micro* 9(9):658–669
30. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE (2006) Metagenomic analysis of the human distal gut microbiome. *Science (New York, N.Y.)* 312(5778):1355–1359
31. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci* 108(4):1513–1518
32. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. *Science* 274(5287):546–567
33. Grainger DC, Hurd D, Goldberg MD, Busby SJW (2006) Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res* 34(16):4642–4652
34. Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJW (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci USA* 102(49):17693–17698
35. Güell M, van Noort V, Yus E, Chen W-H, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kühner S, Rode M, Suyama M, Schmidt S, Gavin A-C, Bork P, Serrano L (2009) Transcriptome complexity in a genome-reduced bacterium. *Science* 326(5957):1268–1271
36. Guelzim N, Bottani S, Bourgine P, Kepes F (2002) Topological and causal structure of the yeast transcriptional regulatory network. *Nat Genetics* 31(1):60–63
37. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, Fan L, Koziol MJ, Gnirke A, Nusbaum C, Rinn JL, Lander ES, Regev A (2010) Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol* 28(5):503–510
38. Harbison CT, Gordon B, Lee TII, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne J-BB, Reynolds DB, Yoo J, Jennings EG, Zeitlinger J, Pokholok DK, Kellis M, Rolfe PA, Takusagawa KT, Lander ES, Gifford DK, Fraenkel E, Young RA (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* 431(7004):99–104
39. Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD (2010) Evolution of MRSA during hospital transmission and intercontinental spread. *Science (New York, N.Y.)* 327(5964):469–474
40. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, DiMeo J, Efcavitch JW, Giladi E, Gill J, Healy J, Jarosz M, Lapen D, Moulton K, Quake SR, Steinmann K, Thayer E, Tyurina A, Ward R, Weiss H, Xie Z (2008) Single-molecule DNA sequencing of a viral genome. *Science* 320(5872):106–109
41. Hernandez D, François P, Farinelli L, Østerås M, Schrenzel J (2008) De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* 18(5):802–809
42. He M, Sebahia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, Holt KE, Seth-Smith HMB, Quail MA, Rance R, Brooks K, Churcher C, Harris D, Bentley SD, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren BW, Dougan G, Parkhill J (2010) Evolutionary dynamics of *clostridium difficile* over short and long time scales. *Proc Natl Acad Sci* 107(16):7527–7532
43. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431(7011):931–945
44. Ishihama A (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Ann Rev Microbiol* 54:499–518
45. Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV (2012) Untangling genomes from metagenomes: revealing an uncultured class of marine euryarchaeota. *Science* 335(6068):587–590

46. Jeck WR, Reinhardt JA, Baltrus DA, Hickenbotham MT, Magrini V, Mardis ER, Dangl JL, Jones CD (2007) Extending assembly of short DNA sequences to handle error. *Bioinformatics* 23(21):2942–2944
47. Ji Hongkai, Jiang Hui, Ma Wenxiu, Johnson David S, Myers Richard M, Wong Wing H (2008) An integrated software system for analyzing CHIP-chip and CHIP-seq data. *Nat Biotechnol* 26(11):1293–1300
48. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from CHIP-Seq data. *Nucleic Acids Res* 36(16):5221–5231
49. Kahramanoglou C, Prieto AI, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM, Luscombe NM, Seshasayee ASN (2012) Genomics of DNA cytosine methylation in *Escherichia coli* reveals its role in stationary phase transcription. *Nat Commun* 3:886+
50. Kahramanoglou C, Seshasayee ASN, Prieto AI, Ibberson D, Schmidt S, Zimmermann J, Benes V, Fraser GM, Luscombe NM (2011). Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res* 39(6):2073–2091
51. Köser CU, Holden MTG, Ellington MJ, Cartwright EJP, Brown NM, Ogilvy-Stuart AL, Hsu LY, Chewapreecha C, Croucher NJ, Harris SR, Sanders M, Enright MC, Dougan G, Bentley SD, Parkhill J, Fraser LJ, Betley JR, Schulz-Trieglaff OB, Smith GP, Peacock SJ (2012) Rapid Whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 366(24):2267–2275
52. Lajoie BR, van Berkum NL, Sanyal A, Dekker J (2009) My5C: web tools for chromosome conformation capture studies. *Nat Methods* 6(10):690–691
53. Langmead B, Trapnell C, Pop M, Salzberg S (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25–10
54. Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penadés JR, Valle J, Solano C, Gingeras TR (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Nat Acad Sci* 108(50):20172–20177
55. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne J-B, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 298(5594):799–804
56. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18(11):1851–1858
57. Li R, Li Y, Kristiansen K, Wang J (2008) SOAP: short oligonucleotide alignment program. *Bioinformatics* 24(5):713–714
58. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760
59. Lieberman TD, Michel J-B, Aingaran M, Potter-Bynoe G, Roux D, Davis MR, Skurnik D, Leiby N, LiPuma JJ, Goldberg JB, McAdam AJ, Priebe GP, Kishony R (2011) Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat Genet* 43(12):1275–1280
60. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozcy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326(5950):289–293
61. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 20(2):265–272
62. MacLean D, Jones JDG, Studholme DJ (2009) Application of 'next-generation' sequencing technologies to microbial genetics. *Nat Rev Microbiol* 7(4):287–296
63. Marcy Y, Ouverney C, Bik EM, Lösekann T, Ivanova N, Martin HG, Szeto E, Platt D, Hugenholtz P, Relman DA, Quake SR (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc Nat Acad Sci* 104(29):11889–11894

64. Mardis ER (2008) The impact of next-generation sequencing technology on genetics. *Trends Genet (TIG)* 24(3):133–141
65. Martínez-Antonio A, Collado-Vides J (2003) Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr Opin Microbiol* 6(5):482–489
66. McAdam PR, Templeton KE, Edwards GF, Holden MTG, Feil EJ, Aanensen DM, Bargawi HJA, Spratt BG, Bentley SD, Parkhill J, Enright MC, Holmes A, Girvan EK, Godfrey PA, Feldgarden M, Kearns AM, Rambaut A, Robinson DA, Fitzgerald JR (2012) Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. *Proc Nat Acad Sci* 109(23):9107–9112
67. McCarthy AJ, Witney AA, Gould KA, Moodley A, Guardabassi L, Voss A, Denis O, Broens EM, Hinds J, Lindsay JA (2011) The distribution of mobile genetic elements (MGEs) in MRSA CC398 is associated with both host and country. *Genome Biol Evol* 3:1164–1174
68. McCarthy AJ, Breathnach AS, Lindsay JA (2012) Detection of mobile-genetic-element variation between colonizing and infecting hospital-associated methicillin-resistant *Staphylococcus aureus* isolates. *J Clin Microbiol* 50(3):1073–1075
69. McGuire J et al (2009) Genome 10K community of scientists. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J Hered* 100(6):659–674
70. Militelio KT, Simon RD, Qureshi M, Maines R, Horne ML, Hennick SM, Jayakar SK, Pounder S (2012) Conservation of Dcm-mediated cytosine DNA methylation in *Escherichia coli*. *FEMS Microbiol Lett* 328(1):78–85
71. Mraheil MA, Billion A, Mohamed W, Mukherjee K, Kuenne C, Pischmarov J, Krawitz C, Retey J, Hartsch T, Chakraborty T, Hain T (2011) The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. *Nucleic Acids Res* 39(10):4235–4248
72. Mutreja A, Kim DWW, Thomson NR, Connor TR, Lee JHH, Kariuki S, Croucher NJ, Choi SYY, Harris SR, Lebens M, Niyogi SKK, Kim EJJ, Ramamurthy T, Chun J, Wood JL, Clemens JD, Czerkinsky C, Nair GB, Holmgren J, Parkhill J, Dougan G (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* 477(7365):462–465
73. Narain AS, Seshasayee S (2011) An overview of prokaryotic transcription factors: a summary of function and occurrence in bacterial genomes. *Subcell Biochem* 52:7–23
74. Nicolas P, Mäder U, Dervyn E, Rochat T, Leduc A, Pigeonneau N, Bidnenko E, Marchadier E, Hoebeke M, Aymerich S, Becher D, Bisicchia P, Botella E, Delumeau O, Doherty G, Denham EL, Fogg MJ, Fromion V, Goelzer A, Hansen A, Härtig E, Harwood CR, Homuth G, Jarmer H, Jules M, Klipp E, Le Chat L, Lecointe F, Lewis P, Liebermeister W, March A, Mars RAT, Nannapaneni P, Noone D, Pohl S, Rinn B, Rügheimer F, Sappa PK, Samson F, Schaffer M, Schwikowski B, Steil L, Stülke J, Wiegert T, Devine KM, Wilkinson AJ, van Dijl JM, Hecker M, Völker U, Bessières P, Noirot P (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* 335(6072):1103–1106
75. Ning Z, Cox AJ, Mullikin JC (2001) SSAHA: a fast search method for large DNA databases. *Genome Res* 11(10):1725–1729
76. Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N (2006) *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res Int J Rapid Publ Rep Genes Genomes* 13(4):141–153
77. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genetics* 10(10):669–680
78. Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Mol Genet Genomics* 272(5):580–591
79. Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, Assefa SA, He M, Croucher NJ, Pickard DJ, Maskell DJ, Parkhill J, Choudhary J, Thomson NR, Dougan G (2009) A strand-specific RNA-seq analysis of the transcriptome of the typhoid bacillus salmonella typhi. *PLoS Genet*, 5(7):e1000569+
80. Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol* 5(11):R87+

81. Prabhakara S, Khedkar S, Loganathan RM, Chandana S, Gowda M, Arakere G, Seshasayee ASN (2012) Draft genome sequence of *Staphylococcus aureus* 118 (ST772), a major disease clone from India. *J Bacteriol* 194(14):3727–3728
82. Prieto AI, Kahramanoglou C, Ali RM, Fraser GM, Seshasayee ASN, Luscombe NM (2011) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res*
83. Qin J, Li R, Raes J, Arumugam M, Burgdorf KSS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-MM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Consortium M, Bork P, Ehrlich SD, Wang J (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464(7285):59–65
84. Reid J (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061–1073
85. Robbins-Manke JL, Zdraveski ZZ, Marinus M, Essigmann JM (2005) Analysis of global gene expression and double-strand-break formation in DNA adenine methyltransferase- and mismatch repair-deficient *Escherichia coli*. *J Bacteriol* 187(20):7027–7037
86. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11(3):R25+
87. Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 9(5):313–323
88. Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, Carriero N, Snyder M, Gerstein MB (2009) PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat Biotechnol* 27(1):66–75
89. Salgado H, Gama-Castro S, Peralta-Gil M, Díaz-Peredo E, Sánchez-Solano F, Santos-Zavaleta A, Martínez-Flores I, Jiménez-Jacinto V, Bonavides-Martínez C, Segura-Salazar J, Martínez-Antonio A, Collado-Vides J (2006) RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res* 34(suppl 1):D394–D397
90. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Nat Acad Sci USA* 74(12):5463–5467
91. Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nat Methods* 5(1):16–18
92. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeisz S, Sittka A, Chabas S, Reiche K, Hackermuller J, Reinhardt R, Stadler PF, Vogel J (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464(7286):250–255
93. Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* 26(10):1135–1145
94. Shi Y, Tyson GW, DeLong EF (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459(7244):266–269
95. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* 38(11):1348–1354
96. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JCD, Vogel J (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* 4(8):e1000163+
97. Smith MG, Gianoulis TA, Pukatzki S, Mekalanos JJ, Ornston LN, Gerstein M, Snyder M (2007) New insights into *Acinetobacter baumannii* pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis. *Genes Dev* 21(5):601–614
98. Spyrou C, Stark R, Lynch AG, Tavaré S (2009) BayesPeak: bayesian analysis of ChIP-seq data. *BMC Bioinf* 10(1):299+
99. Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 43(6):880–891

100. Tanay A, Sharan R, Kupiec M, Shamir R (2004) Revealing modularity and organization in the yeast molecular network by integrated analysis of highly heterogeneous genomewide data. *Proc Natl Acad Sci USA* 101(9):2981–2986
101. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, y Ros IM, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc Natl Acad Sci USA* 102(39):13950–13955
102. Toprak E, Veres A, Michel J-B, Chait R, Hartl DL, Kishony R (2012) Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat Genet* 44(1):101–105
103. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* 25(9):1105–1111
104. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat biotechnol* 28(5):511–515
105. Umbarger MA, Toro E, Wright MA, Porreca GJ, Baù D, Hong S-HH, Fero MJ, Zhu LJ, Marti-Renom MA, McAdams HH, Shapiro L, Dekker J, Church GM (2011) The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. *Mol Cell* 44(2):252–264
106. Venter CJ, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers Y-H, Smith HO (2004) Environmental genome shotgun sequencing of the Sargasso sea. *Science* 304(5667):66–74
107. Voelkerding KV, Dames SA, Durtschi JD (2009) Next-generation sequencing: from basic research to diagnostics. *Clin Chem* 55(4):641–658
108. Vora T, Hottes AK, Tavazoe S (2009) Protein occupancy landscape of a bacterial genome. *Mol Cell* 35(2):247–253
109. Wade JT, Roa DC, Grainger DC, Hurd D, Busby SJW, Struhl K, Nudler E (2006) Extensive functional overlap between σ factors in *Escherichia coli*. *Nat Struct Mol Biol* 13(9):806–814
110. Waldminghaus T, Skarstad K (2010) ChIP on chip: surprising results are often artifacts. *BMC Genomics* 11(1):414+
111. Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang X, Hernandez M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450(7169):560–565
112. Warren RL, Sutton GG, Jones SJ, Holt RA (2007) Assembling millions of short DNA sequences using SSAKE. *Bioinformatics* (Oxford, England) 23(4):500–501
113. Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigma S-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* 187(5):1591–1603
114. Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Médigue C, Lenski RE, Schneider D (2011) Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. *G3: Genes Genomes Genet* 1(3):183–186
115. Willenbrock H, Hallin P, Wassenaar T, Ussery D (2007) Characterization of probiotic *Escherichia coli* isolates with a novel pan-genome microarray. *Genome Biol* 8(12):R267+
116. Woyke T, Xie G, Copeland A, González JM, Han C, Kiss H, Saw JH, Senin P, Yang C, Chatterji S, Cheng J-F, Eisen JA, Sieracki ME, Stepanauskas R (2009) Assembling the marine metagenome, one cell at a time. *PLoS ONE* 4(4):e5299+

117. Yaffe E, Tanay A (2011) Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat Genet* 43(11):1059–1065
118. Yang L, Jelsbak L, Marvig RL, Damkiær S, Workman CT, Rau MH, Hansen SK, Folkesson A, Johansen HK, Ciofu O, Højby N, Sommer MOA, Molin S (2011) Evolutionary dynamics of bacteria in a human host environment. *Proc Nat Acad Sci* 108(18):7481–7486
119. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18(5):821–829
120. Zhang Q, Lambert G, Liao D, Kim H, Robin K, Tung C, Pourmand N, Austin RH (2011) Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science* 333(6050):1764–1767
121. Zhang Y, Liu T, Meyer C, Eeckhoutte J, Johnson D, Bernstein B, Nusbaum C, Myers R, Brown M, Li W, Liu XS (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* 9(9):R137+

Chapter 3

Sensitivity Analysis of Circadian Entrainment in the Space of Phase Response Curves

Pierre Sacré and Rodolphe Sepulchre

Abstract Sensitivity analysis is a classical and fundamental tool to evaluate the role of a given parameter in a given system characteristic. Because the phase response curve is a fundamental input–output characteristic of oscillators, we developed a sensitivity analysis for oscillator models in the space of phase response curves. The proposed tool can be applied to high-dimensional oscillator models without facing the curse of dimensionality obstacle associated with numerical exploration of the parameter space. Application of this tool to a state-of-the-art model of circadian rhythms suggests that it can be useful and instrumental to biological investigations.

Keywords Circadian entrainment · Sensitivity analysis · Input–output · Phase response curve (PRC)

3.1 Introduction

Circadian entrainment is a biological process at the core of most living organisms which need to adapt their physiological activity to the 24h environmental cycle associated with earth’s rotation (e.g. variations in light or temperature condition). This process relies on the robust interaction between an autonomous molecular oscillator and its environment (Fig. 3.1a). Experimental observations have shown that the system is capable to exhibit oscillations with a period close to 24h in constant

P. Sacré (✉)

Department of Biomedical Engineering, Johns Hopkins University,
3400 N Charles Street, Baltimore, MD 21218, USA
e-mail: p.sacre@jhu.edu

R. Sepulchre

Department of Engineering, University of Cambridge, Trumpington Street,
Cambridge CB2 1PZ, UK
e-mail: r.sepulchre@eng.cam.ac.uk

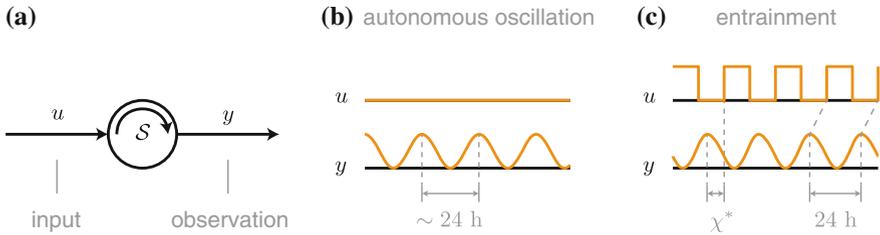


Fig. 3.1 **a** Circadian oscillators are viewed as open dynamical systems with input u and output y . **b** The unforced system exhibits autonomous rhythms that occur with a period close to 24 h. **c** The periodically forced system adapts the organism rhythms through entrainment (1:1 phase-locking) with the 24-h stimulus associated with earth’s rotation

environmental condition (unforced system, Fig. 3.1b) and to lock its oscillations (in frequency and phase) to an environmental cue with a period equal to 24 h (periodically forced system, Fig. 3.1c). This locking phenomenon is often called (circadian) entrainment [23]. Moreover, this biological process is known to be very robust, that is, it maintains its performance (its period and its locking) despite internal or external perturbations (e.g. genetic mutations, molecular noise, variability of the environmental condition, etc.).

With recent experimental advances in biology, the molecular bases of circadian rhythms has been increasingly unfolded in various organisms. In most eukaryotic organisms (e.g. fungus, fly, or mouse), the core mechanism relies on analogous interacting positive and negative feedback loops with several minor alterations [10]. However, even though the architecture of those biological clocks is better known, the specific design and robustness mechanisms implemented in those architectures remain unknown [21, 28].

Starting with the pioneering work of Winfree [32, 33], the Phase Response Curve (PRC) has emerged as a fundamental input–output characteristic of oscillators. Analogously to the static gain of a transfer function, the PRC measures a steady-state (asymptotic) property of the system response induced by an impulsive input. For the static gain, the measured property is the integral of the response; for the PRC, the measured property is the phase shift between the unperturbed and perturbed responses. Because of the periodic nature of the steady-state, this phase shift depends on the phase at which the system receives the impulsive input. The PRC is thus a curve rather than a scalar. In many situations, the PRC can be determined experimentally and provides unique data for the model identification of the oscillator. Likewise, numerical methods exist to compute the PRC from a state-space model of the oscillator. Finally, the PRC contains the fundamental mathematical information required to reduce a n -dimensional state-space model to the one-dimensional (phase) center manifold of a hyperbolic limit cycle.

In this chapter, we review (local) sensitivity tools that provide numerical and mathematical grounds to the robustness analysis of oscillator state-space models in connection with experimentally available observations like the PRC or the period.

We then illustrate how these tools can be used to make physiologically relevant predictions from mathematical models of circadian rhythms. We apply our sensitivity analysis to a state-of-the-art model [16] of 16 states and 52 parameters and exploit the results to extract the parameters and circuits that determine the robustness of entrainment.

The *local* proposed approach is systematic and computationally tractable. It provides a rapid screening of all parameters, even in high-dimensional models with a large number of parameters. It complements *nonlocal* analyses often used to assess the robustness of parameters, such as bifurcation analysis [17] or parameter space exploration [8, 28].

The chapter is organized as follows. Section 3.2 reviews the notion of PRCs characterizing the input–output behavior of an oscillator model in the neighborhood of a stable limit cycle. Section 3.3 develops the sensitivity analysis for oscillators in terms of the sensitivity of its periodic orbit, its PRC, and its entrainment (phase-locking). Section 3.4 provides scalar robustness measures based on this sensitivity analysis. Section 3.5 illustrates how those tools permit to address system-theoretic questions meaningful for the robustness analysis of circadian entrainment.

3.2 Open Oscillator Models: From State-Space to Phase Models

In this section, we provide a short introduction to oscillators viewed as open dynamical systems, that is, as dynamical systems that interact with their environment [26]. We first recall basic definitions about stable periodic orbits in n -dimensional state-space models (see [5, 13] for details). We then introduce (finite and infinitesimal) phase response curves as fundamental input–output mathematical information required for the model reduction. We finally summarize the standard phase reduction procedure which concentrates the phase behavior information of n -dimensional state-space models into one-dimensional phase models characterized by its angular frequency, its PRC, and a measurement map (see [11, 15] for details).

3.2.1 State-Space Models: Periodic Orbits and Phase Maps

We consider open dynamical systems described by nonlinear (single-input and single-output¹) time-invariant state-space models

$$\dot{x} = f(x) + g(x)u, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}, \quad (3.1a)$$

$$y = h(x), \quad y \in \mathbb{R}, \quad (3.1b)$$

¹ For presentation convenience, we consider single-input and single-output systems. All developments are easily generalizable to multiple-input and multiple-output systems.

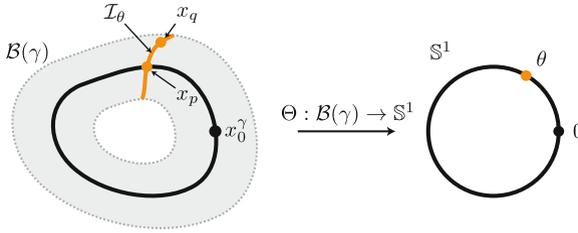


Fig. 3.2 The asymptotic phase map $\Theta : \mathcal{B}(\gamma) \rightarrow \mathbb{S}^1$ associates with each point x_q in the basin $\mathcal{B}(\gamma)$ a scalar phase $\Theta(x_q) = \theta$ on the unit circle \mathbb{S}^1 such that $\lim_{t \rightarrow +\infty} \|\phi(t, x_q, 0) - \phi(t, x_p, 0)\|_2 = 0$ with $x_p = x^\gamma(\theta/\omega)$

where the *vector fields* f and g , and the *measurement map* h support all the usual smoothness conditions that are necessary for existence and uniqueness of solutions. We denote by $\phi(\cdot, x_0, u)$ the solution to the initial value problem (3.1a) from the initial condition $x_0 \in \mathbb{R}^n$ at time 0, that is, $\phi(0, x_0, u) = x_0$.

An oscillator is an open dynamical system whose zero-input steady-state behavior is periodic rather than constant. Formally, we assume that the zero-input system $\dot{x} = f(x)$ admits a *locally hyperbolic stable periodic orbit* $\gamma \subseteq \mathbb{R}^n$ with period T (and corresponding angular frequency $\omega = 2\pi/T$). Picking an initial condition x_0^γ on the periodic orbit γ , this latter is described by the (nonconstant) T -periodic trajectory $\phi(\cdot, x_0^\gamma, 0) =: x^\gamma(\cdot)$, such that $x^\gamma(\cdot) = x^\gamma(\cdot + T)$. The *basin of attraction* of γ is the maximal open set from which the periodic orbit γ attracts. (Main notations are illustrated on Fig. 3.2.)

Since the periodic orbit γ is a one-dimensional manifold in \mathbb{R}^n , it is homeomorphic to the unit circle \mathbb{S}^1 . It is thus naturally parametrized in terms of a single scalar *phase*. The smooth bijective *phase map* $\Theta : \gamma \rightarrow \mathbb{S}^1$ associates with each point x_p on the periodic orbit γ its phase $\Theta(x_p) =: \vartheta_p$ on the unit circle \mathbb{S}^1 , such that,

$$x_p - x^\gamma(\vartheta_p/\omega) = 0. \quad (3.2)$$

This mapping is constructed such that the image of the reference point x_0^γ is equal to 0 (i.e. $\Theta(x_0^\gamma) = 0$) and the progression along the periodic orbit (in absence of perturbation) produces a constant increase in ϑ . The *phase variable* $\vartheta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{S}^1$ is defined along each zero-input trajectory $\phi(\cdot, x_0, 0)$ starting from a point x_0 on the periodic orbit γ , as $\vartheta(t) := \Theta(\phi(t, x_0, 0))$ for all times $t \geq 0$. The phase dynamics are thus given by $\dot{\vartheta} = \omega$.

For hyperbolic stable periodic orbit, the notion of phase can be extended to any point x_q in the basin $\mathcal{B}(\gamma)$ by defining the concept of *asymptotic phase*. The *asymptotic phase map* $\Theta : \mathcal{B}(\gamma) \rightarrow \mathbb{S}^1$ associates with each point x_q in the basin $\mathcal{B}(\gamma)$ its asymptotic phase $\Theta(x_q) =: \theta_q$ on the unit circle \mathbb{S}^1 , such that,

$$\lim_{t \rightarrow +\infty} \|\phi(t, x_q, 0) - \phi(t, x^\gamma(\theta_q/\omega), 0)\|_2 = 0. \quad (3.3)$$

Again, this mapping is constructed such that the image of x_0^γ is equal to 0 and such that the progression along any orbit in $\mathcal{B}(\gamma)$ (in absence of perturbation) produces a constant increase in θ . The *asymptotic phase variable* $\theta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{S}^1$ is defined along each zero-input trajectory $\phi(\cdot, x_0, 0)$ starting from a point x_0 in the basin of attraction of γ , as $\theta(t) := \Theta(\phi(t, x_0, 0))$ for all times $t \geq 0$. The asymptotic phase dynamics are thus given by $\dot{\theta} = \omega$.

The notion of asymptotic phase variable can be extended to a nonzero-input trajectory $\phi(\cdot, x_0, u)$ provided that it stays in the basin of attraction of γ . In this case, the asymptotic phase variable is defined as $\theta(t) := \Theta(\phi(t, x_0, u))$ for all times $t \geq 0$. Thus the variable $\theta(t_*)$ at an instant $t_* \geq 0$ evaluates the asymptotic phase of the point $\phi(t_*, x_0, u)$ such that

$$\lim_{t \rightarrow +\infty} \|\phi(t, \phi(t_*, x_0, u), 0) - \phi(t, x^\gamma(\theta(t_*)/\omega), 0)\| = 0. \quad (3.4)$$

The asymptotic phase dynamics in the case of a nonzero input is often hard to derive.

For presentation convenience, we introduce the map $\tilde{x}^\gamma : \mathbb{S}^1 \rightarrow \gamma$ which associates to each phase θ a point $\phi(\theta/\omega, x_0^\gamma, 0) = \tilde{x}^\gamma(\theta)$ on the periodic orbit. This map corresponds to a reparametrization of the periodic solution $x^\gamma(\cdot)$.

The 2π -periodic steady-state solution $\tilde{x}^\gamma(\cdot)$ and the angular frequency ω can be calculated by solving the boundary value problem [1, 27]

$$(\tilde{x}^\gamma)'(\theta) - \frac{1}{\omega} f(\tilde{x}^\gamma(\theta)) = 0 \quad (3.5a)$$

$$\tilde{x}^\gamma(2\pi) - \tilde{x}^\gamma(0) = 0 \quad (3.5b)$$

$$\psi(\tilde{x}^\gamma(0), \omega) = 0 \quad (3.5c)$$

(where the prime \cdot' denotes the derivative with respect to θ). The boundary conditions are given by the periodicity condition (3.5b) which ensures the periodicity of the map $\tilde{x}^\gamma(\cdot)$ and the phase condition (3.5c) which anchors a reference position $\tilde{x}^\gamma(0)$ along the periodic orbit. The phase condition $\psi : \mathbb{R}^n \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$ is chosen such that it defines an isolated point on the periodic orbit (see [27] for details). Numerical algorithms to solve this boundary value problem are reviewed in [25, Appendix].

3.2.2 Phase Response Curves: Local Information About the Phase Map

For many oscillators, the structure of the asymptotic phase map is very complex. This often makes its analytical computation impossible and even its numerical computation intractable (or at least very expensive, in particular for high-dimensional oscillator models). However, in many situations, the global knowledge of the asymptotic phase map is not required to study oscillator dynamics. Instead, it is sufficient to consider a local phase information also known as the phase response curve.

Starting with the pioneering work of Winfree [32, 33], the phase response curve of an oscillator has proven a useful input–output tool to study oscillator dynamics. It indicates how the timing of inputs affects the timing (steady-state phase shift) of oscillators. Phase response curves are directly related to asymptotic phase maps but capture only partial (local) information about them.

Definition 1 The *Phase Response Curve (PRC)* corresponding to an impulsive input of finite amplitude ε (i.e. $u(\cdot) := \varepsilon\delta(\cdot)$ where $\delta(\cdot)$ is the Dirac delta function) is the map $q_\varepsilon : \mathbb{S}^1 \rightarrow (-\pi, \pi]$ defined as

$$q_\varepsilon(\theta) := \Delta\Theta(\tilde{x}^\gamma(\theta)) = \lim_{t \rightarrow 0^+} \underbrace{\Theta(\phi(t, \tilde{x}^\gamma(\theta), \varepsilon\delta(\cdot)))}_{\text{post-stimulus phase}} - \underbrace{\Theta(\phi(t, \tilde{x}^\gamma(\theta), 0))}_{\text{pre-stimulus phase}}. \quad (3.6)$$

It associates with each point on the periodic orbit (parametrized by its phase θ) the phase shift induced by the input.

In many situations, the PRC can be determined experimentally (in particular for circadian rhythms). Moreover, it can be computed numerically by simulating the nonlinear state-space model and comparing the asymptotic phase shift between perturbed and unperturbed trajectories.

A mathematically more abstract—yet very useful—tool is the infinitesimal phase response curve. It records essentially the same information as the finite phase response curve but for infinitesimally small Dirac delta input ($\varepsilon \ll 1$).

Definition 2 The (input) *infinitesimal Phase Response Curve (iPRC)* is the map $q : \mathbb{S}^1 \rightarrow \mathbb{R}$ defined as the directional derivative

$$q(\theta) := D\Theta(\tilde{x}^\gamma(\theta))[g(\tilde{x}^\gamma(\theta))] \quad (3.7)$$

where

$$D\Theta(x)[\eta] := \lim_{\varepsilon \rightarrow 0} \frac{\Theta(x + \varepsilon\eta) - \Theta(x)}{\varepsilon}. \quad (3.8)$$

The directional derivative can be computed as the inner product

$$D\Theta(x)[g(x)] = \langle \nabla_x \Theta(x), g(x) \rangle \quad (3.9)$$

where $\nabla_x \Theta(x)$ is the gradient of Θ at x . The map $q_x : \mathbb{S}^1 \rightarrow \mathbb{R}^n : \theta \mapsto \nabla_x \Theta(\tilde{x}^\gamma(\theta)) =: q_x(\theta)$ is known as the state infinitesimal phase response curve.

The (state) iPRC $q_x(\cdot)$ can be calculated by solving the boundary value problem [4, 15, 18–20]

$$q'_x(\theta) + \frac{1}{\omega} f_x(\tilde{x}^\gamma(\theta))^T q_x(\theta) = 0 \quad (3.10a)$$

$$q_x(2\pi) - q_x(0) = 0 \quad (3.10b)$$

$$\langle q_x(\theta), f(\tilde{x}^\gamma(\theta)) \rangle - \omega = 0 \quad (3.10c)$$

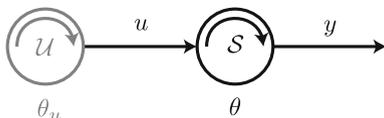


Fig. 3.3 Entrainment is studied by applying weakly connected oscillator theory to the feedforward interconnection between an artificial oscillator generating the input and the actual oscillator

(where the notation A^T stands for the transpose of the matrix A). The boundary condition (3.10b) imposes the periodicity of $q_x(\cdot)$ and the normalization condition (3.10c) ensures a linear increase at rate ω of the phase variable θ along zero-input trajectories. Numerical methods to solve this boundary value problem as a by-product of the periodic orbit computation are presented in [25, Appendix].

Remark 1 For small values of ε (i.e. $\varepsilon \ll 1$), the PRC for impulsive input of finite amplitude is well approximated by the iPRC, that is, $q_\varepsilon(\cdot) = \varepsilon q(\cdot) + \mathcal{O}(\varepsilon^2)$.

3.2.3 Phase Models: Entrainment

In the weak perturbation limit, that is, for small inputs

$$u(t) = \varepsilon \bar{u}(t), \quad \varepsilon \ll 1, \quad |\bar{u}(t)| \leq 1 \quad \text{for all } t, \quad (3.11)$$

any solution $\phi(t, x_0, u)$ of the oscillator model which starts in the neighborhood of the hyperbolic stable periodic orbit γ stays in its neighborhood. The n -dimensional state-space model can thus be approximated by a one-dimensional (continuous-time) phase model [4, 15, 18–20]

$$\dot{\theta} = \omega + \varepsilon q(\theta) \bar{u}(t) \quad \theta \in \mathbb{S}, \quad \bar{u} \in \mathbb{R}, \quad (3.12a)$$

$$y = \tilde{h}(\theta) \quad y \in \mathbb{R}. \quad (3.12b)$$

The phase model is fully characterized by its angular frequency $\omega > 0$, its infinitesimal phase response map $q : \mathbb{S}^1 \rightarrow \mathbb{R}$, and its measurement map $\tilde{h} : \mathbb{S}^1 \rightarrow \mathbb{R}$.

To study entrainment through weak coupling, we can apply weakly connected oscillator theory [11, Chap. 9] by considering the input $u(t)$ as generated by an artificial oscillator described by the trivial phase model $\dot{\theta}_u = \omega_u$, $y_u = \tilde{h}_u(\theta_u)$, where we denote by ω_u the input angular frequency and we choose the artificial oscillator output map \tilde{h}_u such that $y_u(t) = \bar{u}(t)$ for all times $t \geq 0$. Moreover, the network interconnection in this case is a feedforward interconnection from the artificial oscillator generating the input to the studied oscillator (see Fig. 3.3).

The interconnected phase dynamics are thus given by

$$\dot{\theta}_u = \omega_u \quad (3.13a)$$

$$\dot{\theta} = \omega + \varepsilon q(\theta) \tilde{h}_u(\theta_u). \quad (3.13b)$$

Following the weakly connected oscillator theory, we decompose the angular frequencies as $\omega = \Omega + \Delta$ and $\omega_u = \Omega_u + \Delta_u$ with $\Omega - \Omega_u = 0$, and the phase variables as $\theta = \Omega t + \varphi$ and $\theta_u = \Omega_u t + \varphi_u$ where φ and φ_u are slow phase deviations from the fast oscillations Ωt and $\Omega_u t$. The phase deviation dynamics are given by

$$\dot{\varphi}_u = \Delta_u \quad (3.14a)$$

$$\dot{\varphi} = \Delta + \varepsilon q(\Omega t + \varphi) \tilde{h}_u(\Omega_u t + \varphi_u). \quad (3.14b)$$

Assuming that $\Delta, \Delta_u, \varepsilon \ll 1$, standard averaging techniques yield

$$\dot{\varphi}_u = \Delta_u \quad (3.15a)$$

$$\dot{\varphi} = \Delta + \varepsilon \Gamma(\varphi - \varphi_u) \quad (3.15b)$$

where the coupling function is given by

$$\Gamma(\varphi - \varphi_u) = \lim_{\tilde{T} \rightarrow +\infty} \frac{1}{\tilde{T}} \int_0^{\tilde{T}} q(\Omega t + \varphi - \varphi_u) \tilde{h}_u(\Omega_u t) dt. \quad (3.16)$$

Introducing the phase difference $\chi = \varphi - \varphi_u$, we have

$$\dot{\chi} = \Delta - \Delta_u + \Gamma(\chi) =: V(\chi). \quad (3.17)$$

A stable equilibrium χ^* of (3.17), that is,

$$\chi^* \in \mathbb{S}^1 : V(\chi^*) = 0 \quad \text{and} \quad V'(\chi^*) < 0 \quad (3.18)$$

correspond to a stable 1:1 phase-locking behavior (or entrainment) for (3.13), that is,

$$\theta(t) - \theta_u(t) = \chi^* \quad \text{for all times } t. \quad (3.19)$$

In this section, we saw that a state-space oscillator model may be reduced to a phase model characterized by its angular frequency (or period) and its (infinitesimal) phase response curve. In addition, phase models are very useful to study entrainment. It is then very natural to study the sensitivity of oscillators with an emphasis on those characteristics.

3.3 Sensitivity Analysis for Oscillators

Sensitivity analysis for oscillators has been widely studied in terms of sensitivity analysis of periodic orbits [12, 14, 24, 31]. Because the phase response curve is an important oscillator characteristic, we recently proposed a sensitivity analysis of oscillator models in the space of phase response curves [25]. Moreover, the sensitivity analysis in the space of PRC can be exploited to predict the sensitivity of the entrainment.

We summarize those developments for nonlinear time-invariant state-space models with one parameter²

$$\dot{x} = f(x, \lambda) + g(x, \lambda)u \quad (3.20a)$$

$$y = h(x, \lambda) \quad (3.20b)$$

where the constant parameter λ belongs to \mathbb{R} .

3.3.1 Sensitivity Analysis of a Periodic Orbit

The periodic orbit γ of an oscillator model is characterized by its angular frequency ω which measures the ‘speed’ of a solution along the orbit and by the 2π -periodic steady-state solution $\tilde{x}^\gamma(\cdot)$ which describes the locus of this orbit in the state space. The sensitivity of both characteristics is important.

Given a nominal parameter value λ_0 , the sensitivity of the angular frequency is the scalar $S_\omega \in \mathbb{R}$ defined as

$$S_\omega := \left. \frac{d\omega}{d\lambda} \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{\omega|_{\lambda_0+h} - \omega|_{\lambda_0}}{h} \quad (3.21)$$

where the notation $\star|_\lambda$ emphasizes the parameter value λ at which the model characteristic \star is evaluated. Likewise, the sensitivity of the 2π -periodic steady-state solution is the 2π -periodic function $Z_{\tilde{x}} : \mathbb{S}^1 \rightarrow \mathbb{R}^n$ defined as

$$Z_{\tilde{x}}(\cdot) := \left. \frac{d\tilde{x}^\gamma}{d\lambda}(\cdot) \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{\tilde{x}^\gamma(\cdot)|_{\lambda_0+h} - \tilde{x}^\gamma(\cdot)|_{\lambda_0}}{h} \quad (3.22)$$

where the explicit dependence of the 2π -periodic steady-state solution in λ is given by

$$\tilde{x}^\gamma(\cdot)|_\lambda = \phi(\cdot/\omega|_\lambda, x_0^\gamma|_\lambda, 0)|_\lambda. \quad (3.23)$$

² For presentation convenience, we consider systems with a one-dimensional parameter space. All developments are easily generalizable to systems with a q -dimensional parameter space.

From (3.5), we have, taking derivatives with respect to λ ,

$$Z'_{\tilde{x}}(\theta) - \frac{1}{\omega} A(\theta) Z_{\tilde{x}}(\theta) + \frac{1}{\omega^2} \tilde{v}(\theta) S_{\omega} - \frac{1}{\omega} b(\theta) = 0 \quad (3.24a)$$

$$Z_{\tilde{x}}(2\pi) - Z_{\tilde{x}}(0) = 0 \quad (3.24b)$$

$$\psi_x Z_{\tilde{x}}(0) + \psi_{\omega} S_{\omega} + \psi_{\lambda} = 0 \quad (3.24c)$$

where we use the following short notations

$$A(\cdot) := \frac{\partial f}{\partial x}(\tilde{x}^{\gamma}(\cdot), \lambda_0), \quad \psi_x := \frac{\partial \psi}{\partial x}(x_0^{\gamma}, \omega, \lambda_0), \quad (3.25)$$

$$b(\cdot) := \frac{\partial f}{\partial \lambda}(\tilde{x}^{\gamma}(\cdot), \lambda_0), \quad \psi_{\omega} := \frac{\partial \psi}{\partial \omega}(x_0^{\gamma}, \omega, \lambda_0), \quad (3.26)$$

$$\tilde{v}(\cdot) := f(\tilde{x}^{\gamma}(\cdot), \lambda_0), \quad \psi_{\lambda} := \frac{\partial \psi}{\partial \lambda}(x_0^{\gamma}, \omega, \lambda_0). \quad (3.27)$$

Remark 2 In the literature, the sensitivity of the period is often used instead of the sensitivity of the angular frequency. It is the scalar $S_T \in \mathbb{R}$ defined as

$$S_T := \left. \frac{dT}{d\lambda} \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{T|_{\lambda_0+h} - T|_{\lambda_0}}{h}. \quad (3.28)$$

Both sensitivity measures are equivalent up to a change of sign and a scaling factor. The following relationship holds

$$S_T/T = -S_{\omega}/\omega. \quad (3.29)$$

3.3.2 Sensitivity Analysis of a Phase Response Curve

Given a nominal parameter value λ_0 , the sensitivity of the (input) infinitesimal phase response curve is the 2π -periodic function $Z_q : \mathbb{S}^1 \rightarrow \mathbb{R}$ defined as

$$Z_q(\cdot) := \left. \frac{dq}{d\lambda}(\cdot) \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{q(\cdot)|_{\lambda_0+h} - q(\cdot)|_{\lambda_0}}{h}. \quad (3.30)$$

From (3.7), we have, taking derivatives with respect to λ ,

$$Z_q(\cdot) = \left\langle Z_{q_x}(\cdot), g(\tilde{x}^{\gamma}(\cdot), \lambda_0) \right\rangle + \left\langle q_x(\cdot), \frac{\partial g}{\partial x}(\tilde{x}^{\gamma}(\cdot), \lambda_0) Z_{\tilde{x}}(\cdot) + \frac{\partial g}{\partial \lambda}(\tilde{x}^{\gamma}(\cdot), \lambda_0) \right\rangle \quad (3.31)$$

where the 2π -periodic function $Z_{q_x} : \mathbb{S}^1 \rightarrow \mathbb{R}^n$ is the sensitivity of the (state) infinitesimal phase response curve defined as

$$Z_{q_x}(\cdot) := \left. \frac{dq_x}{d\lambda}(\cdot) \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{q_x(\cdot)|_{\lambda_0+h} - q_x(\cdot)|_{\lambda_0}}{h}. \quad (3.32)$$

From (3.10), we have, taking derivatives with respect to λ ,

$$Z'_{q_x}(\theta) + \frac{1}{\omega} A(\theta)^T Z_{q_x}(\theta) + \frac{1}{\omega} C(\theta)^T q_x(\theta) = 0 \quad (3.33a)$$

$$Z_{q_x}(2\pi) - Z_{q_x}(0) = 0 \quad (3.33b)$$

$$\langle Z_{q_x}(\theta), \tilde{v}(\theta) \rangle + \langle q_x(\theta), Z_{\tilde{v}}(\theta) \rangle - S_\omega = 0 \quad (3.33c)$$

where elements of the matrix $C(\cdot)$ are given by

$$\begin{aligned} C_{ij}(\cdot) := & \sum_{k=1}^n \frac{\partial^2 f_i}{\partial x_j \partial x_k}(\tilde{x}^\gamma(\cdot), \lambda_0) (Z_x)_k(\cdot) \\ & + \frac{\partial^2 f_i}{\partial x_j \partial \lambda}(\tilde{x}^\gamma(\cdot), \lambda_0) - \frac{1}{\omega} \frac{\partial f_i}{\partial x_j}(\tilde{x}^\gamma(\cdot), \lambda_0) S_\omega, \end{aligned} \quad (3.34)$$

and where the 2π -periodic function $Z_{\tilde{v}} : \mathbb{S}^1 \rightarrow \mathbb{R}^n$ is the sensitivity of the vector field evaluated along the periodic orbit defined as

$$Z_{\tilde{v}}(\cdot) := \left. \frac{d\tilde{v}}{d\lambda}(\cdot) \right|_{\lambda_0} = \lim_{h \rightarrow 0} \frac{\tilde{v}(\cdot)|_{\lambda_0+h} - \tilde{v}(\cdot)|_{\lambda_0}}{h}. \quad (3.35)$$

Given the explicit dependence of the 2π -periodic vector field in λ

$$\tilde{v}(\cdot) = f(\tilde{x}^\gamma(\cdot)|_\lambda, \lambda), \quad (3.36)$$

we have, taking derivatives with respect to λ ,

$$Z_{\tilde{v}}(\cdot) = \frac{\partial f}{\partial x}(\tilde{x}^\gamma(\cdot), \lambda_0) Z_{\tilde{x}}(\cdot) + \frac{\partial f}{\partial \lambda}(\tilde{x}^\gamma(\cdot), \lambda_0). \quad (3.37)$$

3.3.3 Sensitivity Analysis of the 1:1 Phase-Locking

Given a nominal parameter value λ_0 , the sensitivity of the phase difference χ^* is the scalar $S_{\chi^*} \in \mathbb{R}$ defined as

$$S_{\chi^*} := \lim_{h \rightarrow 0} \frac{\chi^*|_{\lambda_0+h} - \chi^*|_{\lambda_0}}{h}. \quad (3.38)$$

From $V(\chi^*) = 0$, we have, taking derivatives of with respect to λ and using (3.17),

$$S_{\chi^*} = - \left[V'(\chi^*|_{\lambda_0})|_{\lambda_0} \right]^{-1} \times \left[S_V(\chi^*|_{\lambda_0}) \right] \quad (3.39)$$

$$= - \left[\Gamma'(\chi^*|_{\lambda_0})|_{\lambda_0} \right]^{-1} \times \left[S_{\Delta} + S_{\Gamma}(\chi^*|_{\lambda_0}) \right] \quad (3.40)$$

where $S_{\Delta} := \lim_{h \rightarrow 0} [\Delta|_{\lambda_0+h} - \Delta|_{\lambda_0}]/h$ and $S_{\Gamma}(\cdot) := \lim_{h \rightarrow 0} [\Gamma(\cdot)|_{\lambda_0+h} - \Gamma(\cdot)|_{\lambda_0}]/h$. Considering that $\omega|_{\lambda} = \Omega + \Delta|_{\lambda}$ is the sum of a parameter independent term Ω and a parameter dependent term Δ , we have that $S_{\omega} = S_{\Delta}$. In addition, from (3.16), we have, taking derivatives with respect to λ ,

$$S_{\Gamma}(\cdot) = \lim_{\tilde{T} \rightarrow +\infty} \frac{1}{\tilde{T}} \int_0^{\tilde{T}} S_q(\Omega t + \cdot) \tilde{h}_u(\Omega u t) dt. \quad (3.41)$$

The sensitivity of the phase difference has thus two distinct contributions:

$$S_{\chi^*} = S_{\chi^*|\omega} + S_{\chi^*|\Gamma} \quad (3.42)$$

where $S_{\chi^*|\omega} := -[\Gamma'(\chi^*|_{\lambda_0})|_{\lambda_0}]^{-1} \times S_{\omega}$ denotes the contribution of the angular frequency sensitivity and $S_{\chi^*|\Gamma} := -[\Gamma'(\chi^*|_{\lambda_0})|_{\lambda_0}]^{-1} \times S_{\Gamma}(\chi^*|_{\lambda_0})$ denotes the contribution of the coupling function sensitivity at χ^* , the latter being closely related to the iPRC.

3.3.4 Numerics of Sensitivity Analysis

Numerical algorithms to solve boundary value problems (3.24) and (3.33) are reviewed in [25, Appendix]. We stress that existing algorithms that compute periodic orbits and iPRCs are easily adapted to compute their sensitivity curves, essentially at the same numerical cost. All numerical tests in Sect. 3.5 have been obtained with a MATLAB numerical code available from the authors.

The proposed approach is systematic and computationally tractable but it only provides a *local* sensitivity analysis in the parameter space, around a nominal set of parameter values. It complements more *global*—but less tractable—tools such as bifurcation analysis or parameter space exploration. Studying the bifurcation diagram associated with a given parameter [17] or using sampling methods in the full parameter space [8, 28] are classical ways to assess the robustness of an oscillator: the parameter range over which the oscillation exists is a (nonlocal) indicator of the sensitivity of the oscillator to the parameters. The limitation of those approaches

is that they are univariate (only one direction of the parameter space is explored in a particular bifurcation diagram) and that the exploration of the parameter space rapidly becomes formidable as the number of parameters grows.

3.4 Scalar Robustness Measures for Oscillators

Testing the robustness of a model against parameter variations is a basic system-theoretic question. In a number of situations, the very purpose of modeling is to identify those parameters that influence a given system property. In the literature, robustness analysis of circadian rhythms mostly studies the zero-input steady-state behavior such as the period or the amplitude of oscillations [6, 28, 29] and (empirical) phase-based performance measures [2, 7, 9, 22]. In this section, we propose scalar robustness measures to quantify the sensitivity of the angular frequency, the infinitesimal phase response curve, and the 1:1 phase-locking to parameters.

3.4.1 Robustness Measure of the Angular Frequency

The angular frequency ω is a positive scalar number. The sensitivity of ω with respect to the parameter λ is thus also a scalar number S_ω , leading to a scalar robustness measure R_ω defined as

$$R_\omega := |S_\omega| \quad (3.43)$$

where $|\cdot|$ denotes the real absolute value function.

3.4.2 Robustness Measure of the Infinitesimal Phase Response Curve

In contrast, the iPRC $q : \mathbb{S}^1 \rightarrow \mathbb{R}$ belongs to an infinite-dimensional space \mathcal{Q} . The sensitivity of q with respect to the parameter λ is thus a vector S_q which belongs to the tangent space $T_q\mathcal{Q}$ at q . A scalar robustness measure R_q is defined as

$$R_q := \|S_q\|_q = \sqrt{g_q(S_q, S_q)} \quad (3.44)$$

where $\|\cdot\|_q$ denotes the norm induced by a Riemannian metric $g_q(\cdot, \cdot)$ at q . In this chapter, we use the simplest metric for signals in $\mathcal{L}_2(\mathbb{S}^1, \mathbb{R})$, that is the standard inner product,

$$g_q(\xi_q, \zeta_q) := \langle \xi_q, \zeta_q \rangle = \int_{\mathbb{S}^1} \xi_q(\theta) \overline{\zeta_q(\theta)} d\theta. \quad (3.45)$$

In our previous work [25], we proposed further metrics which capture equivalence properties in the space of phase response curves. This is motivated by the fact that, in many applications, it is not meaningful to distinguish among PRCs that are related by a scaling factor and/or a phase shift.

3.4.3 Robustness Measure of the 1:1 Phase-Locking

The stable phase difference χ^* is a scalar phase on the unit circle \mathbb{S}^1 . The sensitivity of χ^* with respect to the parameter λ is a scalar number S_{χ^*} , leading to a scalar robustness measure R_{χ^*} defined as

$$R_{\chi^*} := |S_{\chi^*}|. \quad (3.46)$$

3.4.4 Normalized Robustness Measures

When analyzing a model with several parameters ($\lambda \in \Lambda \subseteq \mathbb{R}^q$), all robustness measures R_{\star} (where \star stands for any characteristic of the oscillator) are q -dimensional vectors. Each element of those vectors represents to the scalar robustness measure corresponding to each parameter. The normalized robustness measure

$$\bar{R}_{\star} = \frac{R_{\star}}{\|R_{\star}\|_{\infty}} \quad (3.47)$$

has all its components in the unit interval $[0, 1]$. This normalized measure allows to rank model parameters according to their relative ability to influence the characteristic \star .

Remark 3 Note that $\bar{R}_T = \bar{R}_{\omega}$.

3.5 Application to a Model of Circadian Rhythms

We illustrate our sensitivity analysis on the genetic oscillator model of [16] (see Fig. 3.4). This model accounts for several regulatory processes identified in circadian rhythms of mammals. A negative autoregulatory feedback loop established by the *per* (period) and *cry* (cryptochrome) genes is at the heart of the circadian oscillator. The PER and CRY proteins form a complex PER–CRY that indirectly represses the activation of the *Per* and *Cry* genes. The PER–CRY complexes exert their repressive effect by binding to a complex of two proteins CLOCK–BMAL1. This latter, formed by the products of *Clock* and *Bmal1* genes, activates *Per* and *Cry*

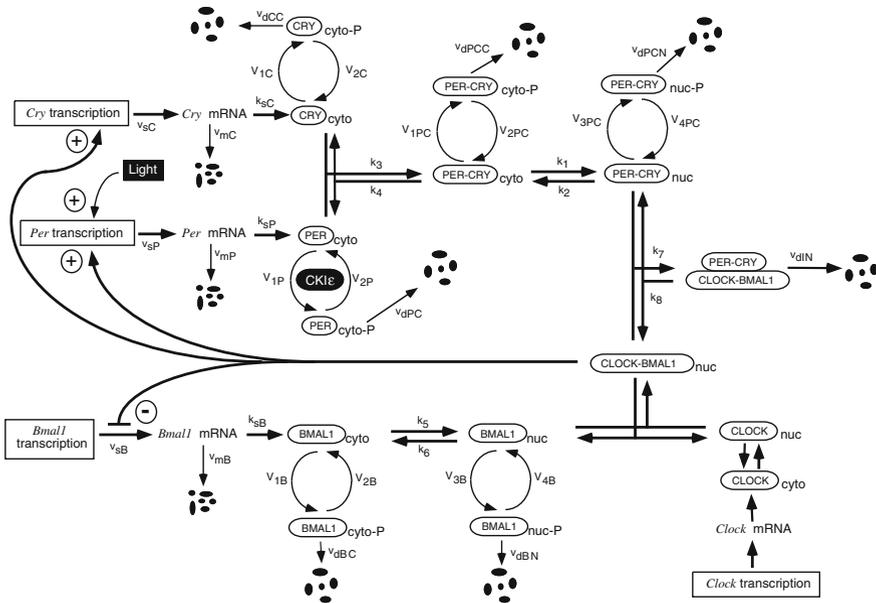


Fig. 3.4 The Leloup-Goldbeter model accounts for several regulatory processes identified in circadian rhythms of mammals (Figure is modified, with permission, from [16]. © (2003) National Academy of Sciences, USA)

transcription. In addition to this negative autoregulation, an (indirect) positive regulatory feedback loop is also involved. Indeed, the *Bmal1* expression is subjected to negative autoregulation by CLOCK–BMAL1, through the product of the *Rev-Erba* gene. The complex PER–CRY enhances *Bmal1* expression in an indirect manner by binding to CLOCK–BMAL1, and thereby reducing the transcription of the *Rev-Erba* gene. Finally, environmental periodic cycles associated with earth’s rotation are mediated through light–dark cycles. Light acts on the system by inducing the expression of the *Per* gene.

The detailed computational model of [16] possesses 16 state variables and 52 parameters. State-space model equations and nominal parameter values are available in [16, Supporting Text]. The effect of light is incorporated through periodic square-wave variations in the maximal rate of *Per* expression (i.e. the value of the parameter v_{sP} goes from a constant low value during dark phase to a constant high value during light phase). Parameters values remain to be determined experimentally and have been chosen semiarbitrarily in physiological ranges in order to satisfy experimental observations.

Each parameter of the model describes a single regulatory mechanism such as transcription and translation control of mRNAs, degradation of mRNAs or proteins, transport reaction, and phosphorylation/dephosphorylation of proteins. The analysis of single-parameter sensitivities reveals thus the importance of individual regulatory processes on the function of the oscillator.

However, in order to enlighten the potential role of circuits rather than single-parameter properties, we grouped model parameters according to the mRNA loop to which they belong. Each group of parameters is associated with a different color: *Per*-loop in blue, *Cry*-loop in red, and *Bmal1*-loop in green. In addition, we gathered parameters associated with interlocked loops in a last group represented in gray.

In the following, we consider sensitivities to relative variations of parameters. We write without distinction about period sensitivities and angular frequency sensitivities due to their direct proportional relationship (see remarks 2 and 3).

3.5.1 Sensitivity Analysis of the Period and the Phase Response Curve

The period and the PRC are two intrinsic characteristics of the circadian oscillator with physiological significance. We use the sensitivity analysis of the period and the PRC to measure the influence of regulatory processes on tuning the period and shaping the PRC.

A two-dimensional $(\bar{R}_\omega, \bar{R}_q)$ scatter plot in which each point corresponds to a parameter of the model reveals the shape and strength of the relationship between both normalized robustness measures \bar{R}_ω (angular frequency or, equivalently, period) and \bar{R}_q (PRC). It enables to identify which characteristic is primarily affected by perturbations in individual parameters: parameters corresponding to points situated below the dashed bisector influence mostly the period; those above the dashed bisector influence mostly the PRC (see Fig. 3.5).

At a coarse level of analysis, the scatter plot reveals that the period and the PRC exhibit a low sensitivity to most parameters (most points are close to the origin); the period and the PRC display a medium or high sensitivity to only few parameters, respectively.

At a finer level of analysis, the scatter plot reveals that the parameters associated with each of the three mRNA loops have distinct sensitivities:

- the *Bmal1*-loop parameters have a strong influence on the period and a medium influence of the PRC (regression line below the bisector);
- the *Per*-loop parameters have a medium influence on the period and a high influence on the PRC (regression line above the bisector);
- the *Cry*-loop parameters have a low influence on the period and a high influence on the PRC (regression line above the bisector, close to the vertical axis).

In each feedback loop, the three more influential parameters represent the three same biological functions: the maximum rates of mRNA synthesis (v_{sB} , v_{sP} , and v_{sC}), the maximum rate of mRNA degradation (v_{mB} , v_{mP} , and v_{mC}), and the inhibition (I) or activation (A) constants for the repression or enhancement of mRNA expression by BMAL1 (K_{IB} , K_{AP} , and K_{AC}).

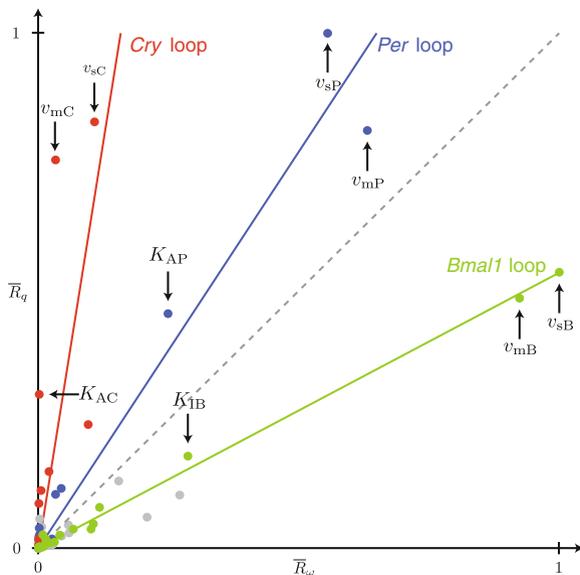


Fig. 3.5 Normalized robustness measures \bar{R}_ω (angular frequency) and \bar{R}_q (iPRC) reveal the distinct sensitivity of three distinct genetic circuits (*Cry*, *Per*, and *Bmal1*). Each point is associated to a particular parameter. The *three lines* are regression over the parameters of the three gene loops. The *dashed* bisector indicates the positions at which both measures of robustness are identical. Only parameters associated with the *Cry*-loop exhibit low angular frequency and high iPRC sensitivities. The *color code* corresponds to different subsets of parameters associated to different loops (see the text for details)

The small number of highly influential parameters is in agreement with the robust nature of the circadian clock and the concentration of fragilities in some specific locations of the architecture [28]. Our analysis suggests that the transcriptional and translational control of mRNA (i.e. the control of both biological steps required to synthesize a protein) has to be regulated by specific mechanisms (not included in the model) in order to avoid failures in the clock function. While the topology of *Per*- and *Cry*-loops are identical, the asymmetry introduced by the choice of parameter values leads to different sensitivity for those loops. Both loops have a similar high sensitivity of the PRC (while the light acts only on the maximum rate of *Per* mRNA synthesis) but a different sensitivity of the period, the *Per*-loop being more influential than the *Cry*-loop. The high sensitivity of the period for parameters associated with the *Bmal1*-loop has also been identified in [17]. However, this last prediction of the model (high sensitivity of the period to *Bmal1*-loop) is not in agreement with experimental observations in [3, 30]. This observation may encourage the biologist and the modeler to design of new experiments to enlighten biological mechanisms responsible for this discrepancy between the experiment and the model.

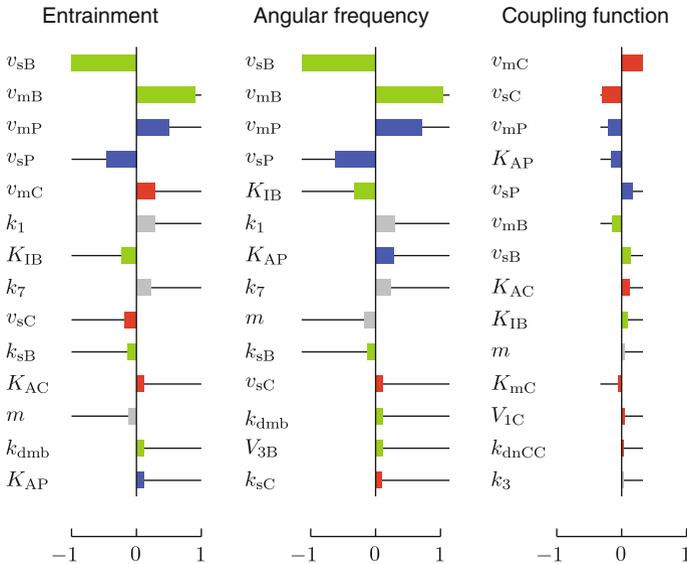


Fig. 3.6 Normalized sensitivity measures $S_{\chi^*} / \|S_{\chi^*}\|_{\infty}$ (entrainment) are due to two contributions: $S_{\chi^*|\omega} / \|S_{\chi^*}\|_{\infty}$ (angular frequency) and $S_{\chi^*|\Gamma} / \|S_{\chi^*}\|_{\infty}$ (coupling function). Each (thick) horizontal bar corresponds to a sensitivity measure with respect to a particular parameter. The (thin) horizontal lines indicate (in absolute value) the maximal sensitivity (among all parameters) and may be useful to compare the sensitivity of a parameter to the maximal sensitivity. The color code corresponds to different subsets of parameters associated to different loops (see the text for details)

3.5.2 Sensitivity Analysis of the Entrainment

Entrainment is an important characteristic of the circadian model. In Section 3.3.3, we have seen that the entrainment sensitivity S_{χ^*} is mathematically given by the summation of two terms: a term $S_{\chi^*|\omega}$ proportional to the period sensitivity and a term $S_{\chi^*|\Gamma}$ proportional to the coupling function sensitivity at χ^* . Those two terms correspond to two biologically distinct mechanisms by which the entrainment properties of the circadian clock can be regulated: a modification of the period or a modification of the coupling function (resulting from the modification of the iPRC or the input signal).

Bar plots of $S_{\chi^*} / \|S_{\chi^*}\|_{\infty}$, $S_{\chi^*|\omega} / \|S_{\chi^*}\|_{\infty}$, and $S_{\chi^*|\Gamma} / \|S_{\chi^*}\|_{\infty}$ in which each bar corresponds to a parameter allows to identify the most influential parameters for entrainment and to quantify³ the respective contribution of both mechanisms in the entrainment sensitivity (see Fig. 3.6). For each bar plot, we sorted parameters

³ The entrainment sensitivity and the contributing terms are normalized by $\|S_{\chi^*}\|_{\infty}$ (the same maximal value of the entrainment sensitivity) such that the summation of normalized terms is equal to the normalized entrainment sensitivity.

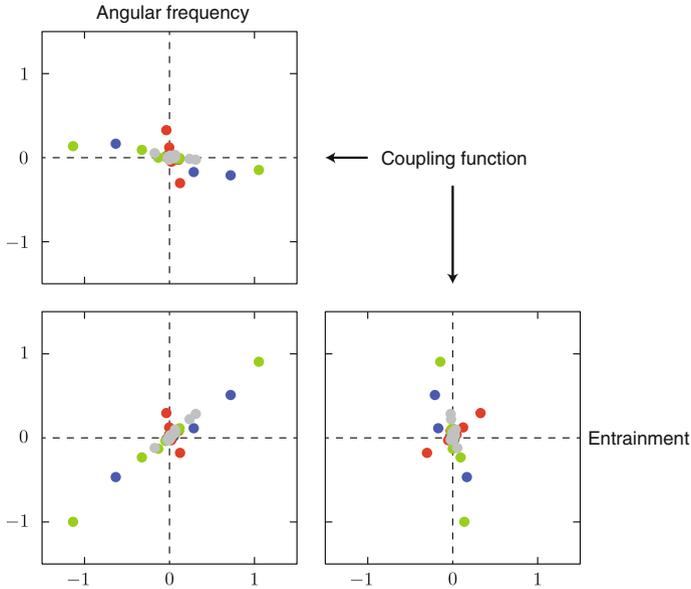


Fig. 3.7 Normalized sensitivity measures $S_{\chi^*}/\|S_{\chi^*}\|_{\infty}$ (entrainment), $S_{\chi^*|\omega}/\|S_{\chi^*}\|_{\infty}$ (angular frequency), and $S_{\chi^*|\Gamma}/\|S_{\chi^*}\|_{\infty}$ (coupling function) exhibit particular correlation shapes. The *top graph* represents the $(S_{\chi^*|\omega}/\|S_{\chi^*}\|_{\infty}, S_{\chi^*|\Gamma}/\|S_{\chi^*}\|_{\infty})$ -plan; the *bottom-left* graph represents the $(S_{\chi^*|\omega}/\|S_{\chi^*}\|_{\infty}, S_{\chi^*}/\|S_{\chi^*}\|_{\infty})$ -plan; and the *bottom-right* graph represents the $(S_{\chi^*|\Gamma}/\|S_{\chi^*}\|_{\infty}, S_{\chi^*}/\|S_{\chi^*}\|_{\infty})$ -plan. Each point is associated to a particular parameter. The *color code* corresponds to different subsets of parameters associated to different loops (see the text for details). Those correlations support the competitive nature of both mechanisms (modification of the period or the coupling function) leading to the entrainment sensitivity

by absolute magnitude and restricted the plot to the 14 parameters with the highest sensitivity measure (the number 14 results from our choice to keep the parameters with an entrainment sensitivity greater than 0.1). Those plots allow to identify the parameters which play an important role in the entrainment sensitivity. We note that the parameter orders for $S_{\chi^*}/\|S_{\chi^*}\|_{\infty}$ and $S_{\chi^*|\omega}/\|S_{\chi^*}\|_{\infty}$ are almost identical, except for parameters associated with the *Cry*-loop. Those parameters appear in the highest ones for $S_{\chi^*|\Gamma}/\|S_{\chi^*}\|_{\infty}$.

Figure 3.7 (top) reveals the competitive and complementary nature of both contributions to entrainment sensitivity. For most parameters, both contributions have opposite signs, that is, points are located in the second and fourth quadrants. In addition, both mechanisms are well decoupled such that, when one mechanism is active, the other is almost inactive (points are located close to the horizontal and vertical axes). Parameters associated with *Cry*-loop seem to influence the entrainment sensitivity through a modification of the coupling function (points close to the vertical axis); others parameters associated with *Per*-loop and *Bmal1*-loop seem to

influence the entrainment sensitivity through a modification of the period (points close to the horizontal axis).

The different mechanisms leading to entrainment sensitivity are also observed in both other scatter plots (see Fig. 3.7 bottom-left and -right). In those plots, parameters associated with points close to the bisector of the first and third quadrants influence the entrainment sensitivity through a modification of the period (bottom-left) or the coupling function (bottom-right), respectively. Again, only parameters associated with the *Cry*-loop seem to affect the entrainment through a variation of the PRC.

Two of the parameters belonging to the *Cry*-loop (with high coupling function and low period sensitivities) have been identified by numerical simulations as important for entrainment properties of the model without affecting the period: K_{AC} in [16] and v_{mC} in [17]. Our approach supports the importance of those two parameters and identifies the potential importance of a third one (v_{sC}).

We stress that the sensitivity analysis in [16, 17] is a *global* approach that relies on exploring the parameter space through numerical simulations of the model to determine the system behavior under constant and periodic environmental conditions while varying one parameter at a time. In contrast, the proposed analysis is *local* but systematic and computationally tractable. In the particular model studied here and in [17], the predictions of the (local) sensitivity analysis match the predictions of the (nonlocal) analysis.

To evaluate the nonlocal nature of our local predictions, we plot in Fig. 3.8 the time behavior of solutions for different finite (nonlocal) parameter changes. The left plots illustrate the autonomous oscillation of the isolated oscillator whereas the right plots illustrate the steady-state solution entrained by a periodic light input. Parameter perturbations are randomly taken in a range of $\pm 10\%$ around the nominal parameter value. Each panel corresponds to the perturbation of a different group of parameters (the black time-plot corresponds to the nominal system behaviors for nominal parameter values).

- A. Perturbations of three most influential parameters of *Cry*-loop (v_{sC} , v_{mC} , and K_{AC}) lead to small variations (mostly shortening) of the autonomous period and (not structured) large variations of the phase-locking. This observation is consistent with the low sensitivity of the period and the high sensitivity of the PRC.
- B. Perturbations of three most influential parameters of *Bmal1*-loop (v_{sB} , v_{mB} , and K_{IB}) lead to medium variations of the autonomous period and medium variations of the phase-locking. The variations of the phase-locking exhibit the same structure as variations of the period, suggesting that the change in period is responsible for the change of phase-locking for those parameters. This observation is consistent with the high sensitivity of the period and the medium sensitivity of the PRC.
- C. Perturbations of three most influential parameters of *Per*-loop (v_{sP} , v_{mP} , and K_{AP}) exhibit an intermediate behavior between the situations A and B.

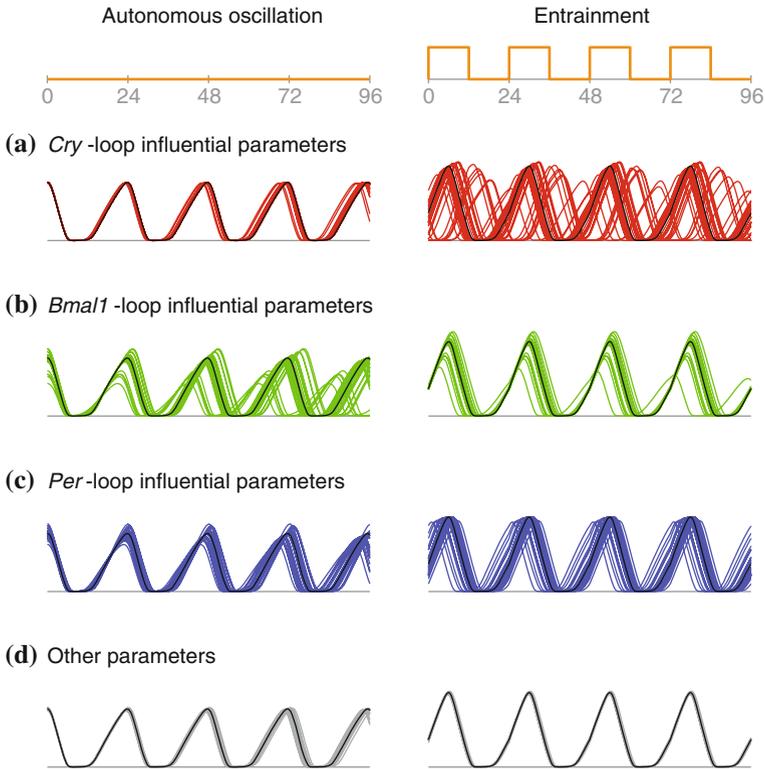


Fig. 3.8 Steady-state behaviors for the nominal model and different finite (nonlocal) parameter perturbations are illustrated by time-plots of the state variable M_P under constant environmental conditions (autonomous oscillation, *left*) and periodic environmental conditions (entrainment, *right*). Each *panel* (or *row*) corresponds to the perturbation of a different group of parameters, the black time-plot corresponding to system behaviors for nominal parameter values. Perturbations are randomly taken in a range of $\pm 10\%$ around the nominal parameter value (for one parameter at a time). **a** Perturbations of three most influential parameters of *Cry*-loop (v_{sC} , v_{mC} , and K_{AC}) lead to small variations of the autonomous period and (not structured) large variations of the phase-locking. **b** Perturbations of three most influential parameters of *Bmal1*-loop (v_{sB} , v_{mB} , and K_{IB}) lead to larger variations of the autonomous period and medium variations of the phase-locking. **c** Perturbations of three most influential parameters of *Per*-loop (v_{sP} , v_{mP} , and K_{AP}) exhibit an intermediate behavior between the situations A and B. **d** Perturbations of parameters of interlocked loops lead to small variations of the autonomous period and the phase-locking

D. Perturbations of parameters of interlocked loops lead to small variations of the autonomous period and the phase-locking, which is consistent with their low sensitivity.

Those (nonlocal) observations are thus well predicted by the classification of parameters suggested by the (local) sensitivity analysis (see Fig. 3.5).

3.6 Conclusion

This chapter proposes (local) sensitivity tools to analyze oscillator models as open dynamical systems. We showed that, under the weak perturbation assumption, state-space models can be reduced to phase models characterized by their angular frequency and their phase response curve. Those phase models are then useful to study the entrainment (or phase-locking) to a periodic input. We then introduced the sensitivity analysis for oscillators and their phase-locking behavior.

The application of this approach to a detailed computational model of circadian rhythms provides physiologically relevant predictions. It enlightens the distinct role of different circuits in the robustness of entrainment and it selects 3 out of 52 parameters as parameters that strongly affect the phase response curve while barely affecting the period. The importance of two of these parameters was previously identified in the literature through simulations of the model.

3.7 Lessons Learnt

Sensitivity analysis is a classical and fundamental tool to evaluate the role of a given parameter in a given system characteristic. Because the phase response curve is a fundamental input–output characteristic of oscillators, we developed a sensitivity analysis for oscillator models in the space of phase response curves. The proposed tool can be applied to high-dimensional oscillator models without facing the curse of dimensionality obstacle associated with numerical exploration of the parameter space. Application of this tool to a state-of-the-art model of circadian rhythms suggests that it can be useful and instrumental to biological investigations.

Acknowledgments This chapter presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors. The work was completed while P. Sacré was a PhD student at the University of Liège. He was a Research Fellow with the Belgian Fund for Scientific Research (F.R.S.-FNRS).

References

1. Ascher UM, Mattheij RMM, Russell RD (1988) Numerical solution of boundary value problems for ordinary differential equations. Prentice Hall, Englewood Cliffs
2. Bagheri N, Stelling J, Doyle FJ III (2007) Quantitative performance metrics for robustness in circadian rhythms. *Bioinformatics* 23(3):358–364
3. Bunger MK, Wilsbacher LD, Moran SM, Clendenin C, Radcliffe LA, Hogenesch JB, Simon MC, Takahashi JS, Bradfield CA (2000) Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* 103(7):1009–1017
4. Ermentrout GB, Kopell N (1984) Frequency plateaus in a chain of weakly coupled oscillators I. *SIAM J Math Anal* 15(2):215–237
5. Farkas M (1994) Periodic motions, applied mathematical sciences, vol 104. Springer, New York

6. Gonze D, Halloy J, Goldbeter A (2002) Robustness of circadian rhythms with respect to molecular noise. *Proc Natl Acad Sci USA* 99(2):673–678
7. Gunawan R, Doyle FJ III (2007) Phase sensitivity analysis of circadian rhythm entrainment. *J Biol Rhythms* 22(2):180–194
8. Hafner M, Koepl H, Hasler M, Wagner A (2009) ‘Glocal’ robustness analysis and model discrimination for circadian oscillators. *PLoS Comput Biol* 5(10):e1000534
9. Hafner M, Sacré P, Symul L, Sepulchre R, Koepl H (2010) Multiple feedback loops in circadian cycles: robustness and entrainment as selection criteria. In: *Proceedings of the 7th international workshop computational systems biology, Luxembourg, Luxembourg*, pp 51–54
10. Hastings MH (2000) Circadian clockwork: two loops are better than one. *Nat Rev Neurosci* 1(2):143–146
11. Hoppensteadt FC, Izhikevich EM (1997) *Weakly connected neural networks*, applied mathematical sciences, vol 126. Springer, New York
12. Ingalls BP (2004) Autonomously oscillating biochemical systems: parametric sensitivity of extrema and period. *Syst Biol* 1(1):62–70
13. Khalil HK (2002) *Nonlinear systems*, 3rd edn. Prentice Hall, Upper Saddle River
14. Kramer MA, Rabitz H, Calo JM (1984) Sensitivity analysis of oscillatory systems. *Appl Math Model* 8(5):328–340
15. Kuramoto Y (1984) *Chemical oscillations, waves, and turbulence*. Springer series in synergetics, vol 19, 1st edn. Springer, Berlin
16. Leloup JC, Goldbeter A (2003) Toward a detailed computational model for the mammalian circadian clock. *Proc Natl Acad Sci USA* 100(12):7051–7056
17. Leloup JC, Goldbeter A (2004) Modeling the mammalian circadian clock: sensitivity analysis and multiplicity of oscillatory mechanisms. *J Theor Biol* 230(4):541–562
18. Malkin IG (1949) *The methods of Lyapunov and Poincare in the theory of nonlinear oscillations*. Gostexizdat, Moscow
19. Malkin IG (1956) *Some problems in the theory of nonlinear oscillations*. Gostexizdat, Moscow
20. Neu JC (1979) Coupled chemical oscillators. *SIAM J Appl Math* 37(2):307–315
21. Novák B, Tyson JJ (2008) Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol* 9(12):981–991
22. Pfeuty B, Thommen Q, Lefranc M (2011) Robust entrainment of circadian oscillators requires specific phase response curves. *Biophys J* 100(11):2557–2565
23. Pittendrigh CS (1981) Circadian systems: entrainment. In: *Aschoff J (ed) Biological rhythms*. Plenum Press, New York, pp 95–124
24. Rosenwasser E, Yusupov R (1999) *Sensitivity of automatic control systems*. CRC Press, Boca Raton
25. Sacré P, Sepulchre R (2012) System analysis of oscillator models in the space of phase response curves. *arXiv math.DS*
26. Sepulchre R (2006) Oscillators as systems and synchrony as a design principle. In: *Current trends in nonlinear systems and control*. In honor of Petar Kokotović and Turi Nicosia. Birkhäuser, Boston, MA, pp 123–141
27. Seydel R (2010) *Practical bifurcation and stability analysis*. Interdisciplinary Applied Mathematics, vol 5, 3rd edn. Springer, New York
28. Stelling J, Gilles ED, Doyle III FJ (2004) Robustness properties of circadian clock architectures. *Proc Natl Acad Sci USA* 101(36):13210–13215
29. Wilkins AK, Barton PI, Tidor B (2007) The Per2 negative feedback loop sets the period in the mammalian circadian clock mechanism. *PLoS Comput Biol* 3(12):e242
30. von Gall C, Noton E, Lee C, Weaver DR (2003) Light does not degrade the constitutively expressed BMAL1 protein in the mouse suprachiasmatic nucleus. *Eur J Neurosci* 18(1):125–133
31. Wilkins AK, Tidor B, White J, Barton PI (2009) Sensitivity analysis for oscillating dynamical systems. *SIAM J Sci Comput* 31(4):2706–2732
32. Winfree AT (1967) Biological rhythms and the behavior of populations of coupled oscillators. *J Theor Biol* 16(1):15–42
33. Winfree AT (1980) *The geometry of biological time*. Biomathematics, vol 8, 1st edn. Springer, New York

Chapter 4

Modelling and Analysis of Feedback Control Mechanisms Underlying Osmoregulation in Yeast

Francesco Montefusco, Ozgur E. Akman, Orkun S. Soyer
and Declan G. Bates

Abstract Biological systems display complex dynamics emerging from intricate networks of interacting molecular components: cells use signalling pathways and regulatory control mechanisms to coordinate multiple processes, allowing them to respond and adapt to an ever-changing environment. Many structural and dynamical features of biological control systems can also be found in engineered control systems and, hence, feedback control theory can provide a useful approach for the analysis and design of complex biological systems. In this chapter we provide a control theoretic analysis of the osmoregulation system in *Saccharomyces cerevisiae* (see [8, 24, 26, 40]), where a complex biochemical signalling and regulatory network allows cells to maintain homeostasis in the face of osmotic shock.

Keywords Osmoregulation · Signalling pathway · Signaling pathway · Homeostasis · Osmosis · Osmoadaptation · Eukaryotic · Yeast · High osmolarity glycerol (HOG) · Integral feedback

F. Montefusco (✉)

Department of Information Engineering, University of Padova, Via Gradenigo
6/b, 35131 Padova, Italy
e-mail: montefusco@dei.unipd.it

O. E. Akman

Centre for Systems, Dynamics and Control, College of Engineering, Mathematics and Physical
Sciences, University of Exeter, Exeter EX4 4QF, UK
e-mail: o.e.akman@exeter.ac.uk

O. S. Soyer

School of Life Sciences, University of Warwick, Gibbet Hill Campus, Coventry CV4 7AL, UK
e-mail: O.Soyer@warwick.ac.uk

D. G. Bates

School of Engineering, University of Warwick, Coventry CV4 7AL, UK
e-mail: D.Bates@warwick.ac.uk

4.1 Introduction

Osmosis is the diffusion of water through a semipermeable membrane (permeable to the solvent, but not the solute), from the compartment containing a low concentration (hypotonic) solution to the one at high concentration (hypertonic). The chemical potential of water is central in this process and can be considered as a measure of the effective water concentration in a given area. The water potential is influenced by two factors, [17]: the osmotic potential and the pressure potential. The first is approximately proportional to the concentration of dissolved molecules of solutes: when the concentration of solute molecules increases, the water potential decreases. The second takes into account the hydrostatic pressure, the pressure exerted by a fluid at equilibrium due to the force of gravity. For two regions of water with different potentials and separated from each other by a semipermeable membrane, there is a water flow to the region of lower potential by osmosis: the movement of the fluid from the hypotonic to the hypertonic solution, while decreasing the concentration difference, increases the pressure of the hypertonic solution with respect to the hypotonic, thus producing a force that counteracts the osmosis. When these two effects balance each other, the osmotic equilibrium is reached: there is no net movement of solvent and the pressure required to maintain an equilibrium is defined as the osmotic pressure.

Osmosis is particularly important for cells, since many biological membranes are permeable to small molecules like water, but impermeable to larger molecules and ions. Osmosis provides the primary means by which water is transported into and out of cells. Typically, a cell has a higher intra cellular osmotic pressure (P_i) than extra cellular osmotic pressure (P_e). The main reason for this difference is that highly charged macromolecules and metabolites attract many small inorganic ions to the cell interior (the Donnan effect, see [1]). Due to this difference, water will flow into the cell, leading to swelling and potentially to cell rupture. The yeast *Saccharomyces cerevisiae* prevents the fundamental problem of water inflow and cell swelling by its cell wall, which is less elastic than the plasma membrane. The cell wall resists the expansion of the cell and creates an inward pressure on the cell contents, Gervais and Beney [9]. This pressure is called the turgor pressure P_t , defined as the difference in the hydrostatic pressure between the inside and the outside of the cell. At equilibrium (equil.), the water potential is equal inside and outside of the cell and the turgor pressure balances the difference in osmotic pressures, as in [33],

$$P_i = P_e + P_t \quad (\text{equil.}). \quad (4.1)$$

Osmotic shocks arise due to a sudden rise (for example the addition of salt to the cell medium) or fall in the concentration of a solute in the cell's environment, resulting in rapid movements of water through the cell's membrane. These movements can produce dramatic consequences for the cell, since loss of water inhibits the transport of substrates and cofactors into the cell, while the uptake of large quantities of water can lead to swelling, rupture of the cell membrane or apoptosis. Due to their more

direct contact with their environment, single-celled organisms are generally more vulnerable to osmotic shock. However, cells in large animals such as mammals also suffer similar stresses under certain conditions, Ho [12].

Osmoadaptation is the mechanism by which cells sense and respond to various changes in their environmental conditions to avoid the aforementioned dramatic consequences. Organisms have evolved a variety of mechanisms to respond to osmotic shock. Typically, cells recognise changes in the osmolarity of their surroundings by using surface sensors which generate signals by activating signal transduction networks. These pathways are found in all eukaryotic organisms and are important in coordinating the response from the cell membrane into the cell, Rep et al. [30]. Recent experimental research indicates that most eukaryotic cells use the mitogen activated protein (MAP) kinase pathways for this purpose, Klitz and Burg [16].

4.2 Osmoregulation Process in Yeast

In recent years, the osmoregulatory response in yeast has emerged as an important model system for studying adaptive, homeostatic responses to environmental disturbances (see [8, 15]). The underlying molecular control system is well characterized in *Saccharomyces cerevisiae* (see [26, 40]), where it comprises three separate mechanisms that act to adjust the glycerol production in order to keep the cell's turgor pressure and volume constant in the face of environmental changes: (1) the regulation of the membrane protein Fps1 determining the glycerol export rate; (2) the transcription of several genes, whose proteins are involved in glycerol production, by the activation of the high osmolarity glycerol (HOG) mitogen-activated protein kinase (MAPK) signaling pathway and (3) the HOG kinase dependent regulation of the glycerol via non-transcriptional mechanisms. Despite its biochemical complexity (see Fig. 4.1), the osmoregulation system in yeast can be naturally abstracted as a feedback control system comprised of distinct branches as described above. This approach was taken in recent studies, which aimed to use standard engineering control models to capture the experimentally observed responses of yeast to osmotic shock and to further predict its structural and dynamic features (see [8, 24, 26]). Genemark et al. [8] combined proportional controllers to model the above-described biochemical branches. Mettetal et al. [24] developed a concise model by using linear systems theory, and then revised this model arguing for the necessity of at least one branch of the system to implement integral control to achieve the experimentally observed adaptive responses in the system, Muzzey et al. [26]. The role of integral feedback in perfectly adaptive systems is by now well-studied in the Systems Biology literature (see [26, 27]), and it is highly likely that the osmoregulation system in yeast does indeed include a biochemical implementation of integral feedback, as seen in other systems (see [6, 27, 39]). It is still unclear, however, exactly how biological control systems such as osmoregulation might have evolved to use integral feedback control, and whether other alternative mechanisms might produce similar (or better) performance properties.

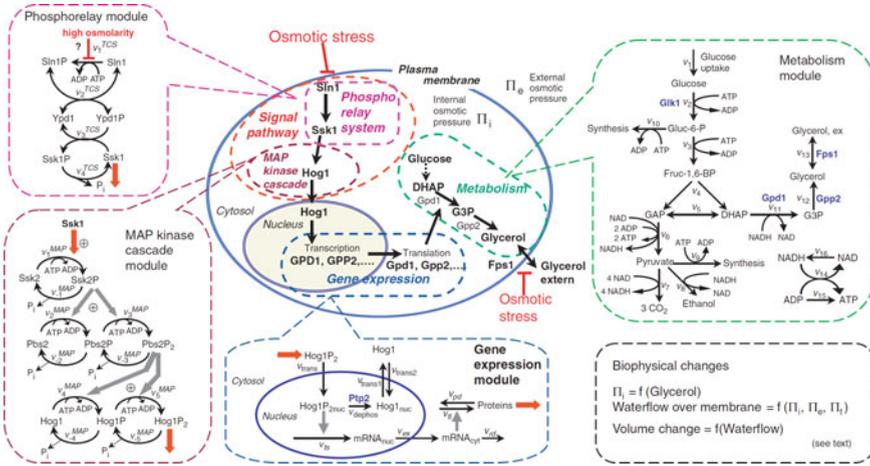


Fig. 4.1 Overview of the response of yeast to osmotic shock, the figure has been taken from [15] and reproduced with permission of Nature Publishing Group

As a first step towards answering this question, we recently extended the proportional controller model devised in [8] with the implementation of an ultrasensitive controller, Montefusco et al. [25]. Ultrasensitivity describes a particular form of sensitivity in biological systems, where the system does not respond to incoming signals outside of a certain regime, but responds in a highly sensitive manner within this regime. Such an input-output relationship (i.e. ultrasensitivity) can be described by a specific nonlinear function, is shown to be a ubiquitous feature in several biological systems, and can be biochemically implemented through a variety of mechanisms such as phosphorylation cycles and cooperative binding (see [4, 11]). The MAPK systems, which are also found in osmoregulation, are theoretically shown to be capable of embedding ultrasensitivity (see [3, 13]), and bistability [21]. Starting from the proportional control model developed by [8], we explore the consequences of such potential ultrasensitivity and show that it significantly increases system performance in achieving homeostasis to osmotic perturbations.

In the following sections we present the model devised in [8], then we focus our attention on the results presented in [24, 26]. Finally, we provide an updated description of the recent results first presented in [25].

4.3 A Proportional Control Based Model of the Osmoregulation in Yeast

In this section we describe the model presented in [8], where the authors devised a simple ordinary differential equation (ODE) model of the adaptive response to an osmotic shock in *S. cerevisiae*. They abstracted several elements to yield a reduced

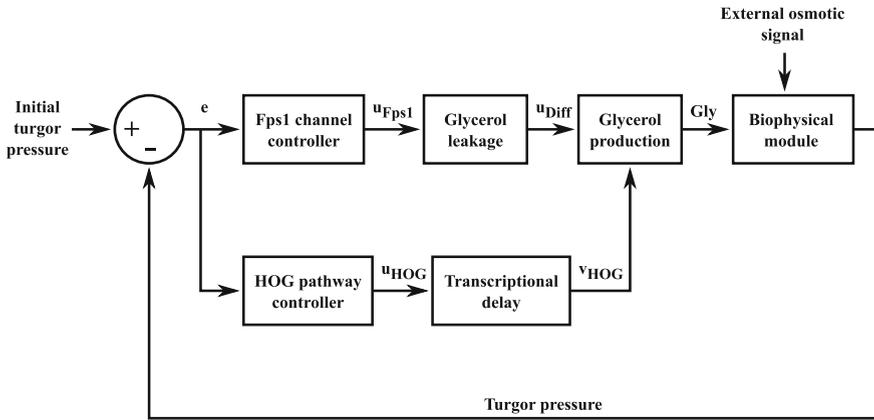


Fig. 4.2 Mathematical model of the osmoregulation process. This figure has been adapted from [8]

representation of the system, starting from the detailed model developed in [15] (see the diagram in Fig. 4.2). The model, in particular, consists of two main components. First, a biophysical model describing how the cell volume and the turgor pressure are affected by varying extra-cellular osmolarity. Second, the two parallel mechanisms for controlling the biophysical system in order to keep turgor pressure and volume constant: one by controlling the production of glycerol via the HOG pathway and the other by controlling the outflow of glycerol via the Fps1 channel. The complete model consists of 4 ODE's, 3 algebraic equations and 10 parameters, that have been estimated using experimental data on glycerol. The authors have validated the model by predicting the behaviour of modified strains and input functions.

4.3.1 The Mathematical Model

The mathematical model presented in [8] is described in the following paragraphs.

4.3.1.1 The Biophysical Module

The biophysical system is modelled by considering the dependencies between cell volume V , the turgor pressure P_t , the intra-cellular osmotic pressure P_i and the extra-cellular osmotic pressure P_e . At any given time t , $P_i(t)$, $P_e(t)$ and $P_t(t)$ are determining the flow of water across the cell membrane, which is proportional to $(P_i(t) - P_e(t) - P_t(t))$. Assuming that the cell volume is only affected by the inflow and outflow of water, then the change in volume can be expressed as

$$\frac{dV}{dt} = k_{p1}(P_i(t) - P_e(t) - P_t(t)), \tag{4.2}$$

with k_{p1} denoting a hydraulic water permeability constant. At equilibrium (equil.), i.e. constant volume and no net flow of water over the membrane, the Eq. (4.2) reduces to (4.1).

The only osmolyte considered explicitly in the model is glycerol (*Gly*) and, hence, ions and other small molecules, changing upon osmotic shock, Sunder et al. [35], are not considered. This assumption is motivated by experimental results from [29], where the authors found that glycerol counter-balances approximately 80% of applied NaCl in *S. cerevisiae*. Therefore, the intra-cellular osmotic pressure, according to van't Hoff's law, is expressed as

$$P_i(t) = \frac{s + Gly(t)}{V(t) - V_b}, \quad (4.3)$$

with s being the concentration of the sum of osmolytes (assumed constant) other than glycerol present in the cell, and V_b being the non-osmotic volume of the cell, subsuming non-polar cellular components, such as membranes. According to Eq. (4.3), the intra-cellular osmotic pressure increases with the glycerol, which can be used to control the turgor pressure of the cell. The extra-cellular osmotic pressure is only modified by the input signal, for example applied salt stress, and is then independent of changes in other variables. The turgor pressure is linearly dependent on the volume according to [17], in the following manner:

$$P_t(t) = \varepsilon \left(\frac{V(t)}{V(0)} - 1 \right) + P_t(0), \quad (4.4)$$

where $V(0)$ is the initial volume, $P_t(0)$ is the initial turgor pressure, and ε is the volumetric elastic modulus. By expressing the volume at which $P_t = 0$ with the notation $V^{P_t=0}$, (4.4) can be rewritten as

$$P_t(t) = \begin{cases} P_t(0) \frac{V(t) - V^{P_t=0}}{V(0) - V^{P_t=0}}, & V(t) > V^{P_t=0} \\ 0, & \text{otherwise.} \end{cases}$$

4.3.1.2 The Controller Modules

There are two branches of control in the model: the first represents the closure of Fps1 glycerol transporter channels as a reaction to osmotic shock, and the second the activation of the HOG pathway, leading to glycerol production after a time delay. The input signal e arriving at the controllers is expressed as

$$e(t) = P_t(0) - P_t(t), \quad (4.5)$$

which is the difference in turgor pressure. The output of the Fps1 branch, which corresponds to the response of the transporter channels, is given by

$$u_{Fps1}(t) = \begin{cases} k_{p2} \frac{P_t(0) - e(t)}{P_t(0)}, & e(t) > 0 \\ k_{p2}, & \text{otherwise.} \end{cases} \quad (4.6)$$

The function u_{Fps1} returns real values in the interval $[0, k_{p2}]$, where 0 corresponds to completely closed and where k_{p2} is the glycerol permeability coefficient in a completely open Fps1 channel.

The output of the HOG branch, which corresponds to the HOG pathway dependent glycerol production, is expressed as

$$u_{HOG}(t) = \begin{cases} k_{HOG} \cdot e(t), & e(t) > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (4.7)$$

where k_{HOG} is the gain of this branch.

The time delay accounting for transcription and translation in the HOG pathway is approximated by

$$\frac{dv_{HOG}}{dt} = \frac{1}{T_d}(u_{HOG}(t) - v_{HOG}(t)),$$

with $v_{HOG}(t)$ being the time delayed variable and T_d being the amount of time delay considered. As reported in [8], very simple proportional controllers have been used in order to reduce the complexity of the model, even though it is known that, for example, MAPK signalling pathways often exhibit a switch-like behaviour, Huang and Ferrell [13]. In the last section of this chapter we compare the dynamics of this model with those obtained by using a HOG controller implementing ultrasensitivity, Montefusco et al. [25] (see Sect. 4.6).

4.3.1.3 The Glycerol Module

The exchange of internal and external glycerol, u_{Diff} over the Fps1 channel is modelled by using Fick's first law of diffusion as

$$u_{Diff}(t) = u_{Fps1}(t) \left(\frac{Gly(t)}{V(t) - V_b} - \frac{Gly_e(t)}{V_e} \right),$$

with V_e being the extra-cellular volume and Gly_e being the glycerol concentration in the extra-cellular compartment. Intra-cellular glycerol Gly production is expressed, combining the output of the two controllers described above, as

$$\frac{dGly}{dt} = v_{HOG}(t) - u_{Diff}(t)$$

Table 4.1 Proportional control based model parameters: all volumes are scaled such that the initial volume of the cell is 1

Parameters		Bounds
k_{p1}	Water perm. coeff.	[0.0052, 160] Osm ⁻¹
k_{p2}	Fps1 control const.	[0, ∞]
T_d	Time delay	[5, 30] min
k_{HOG}	HOG control const.	[0, ∞] Osm ⁻¹
$Gly(0)$	Initial Gly	[1.1 5] × 10 ⁻⁴
$P_i(0)$	Initial P_i	[0.6 0.7] Osm
$P_e(0)$	Initial P_e	[0.24 0.25] Osm
V_b	Non osmotic volume	[0.31 0.46]
$V^{P_i=0}$	V when $P_i = 0$	[0.5 0.99]
V_e	External volume	[0.5 5] × 10 ³
Dependent parameters		Value
$V(0)$	Initial V - relative volume	1
$Gly_e(0)$	Initial Gly_e	$\frac{V_e Gly(0)}{(V(0) - V_b)}$
$P_i(0)$	Initial P_i	$P_i(0) - P_e(0)$
s	No. of osmolytes other than Gly	$P_i(0)(V(0) - V_b)$ $- Gly(0)$

Both Gly and Gly_e represent number of molecules (mol scaled by $V(0)$)

and extra-cellular glycerol, depending only on the diffusion over the Fps1 channel, is described by

$$\frac{dGly_e}{dt} = u_{Diff}(t).$$

4.3.2 Parameter Estimation and Results

The model contains 14 parameters, 4 of which are dependent, as given in Table 4.1. In [8], the other parameters are estimated by simulating the model and minimising the error defined as the sum of the squares of the difference between simulated, $X(t)$, and experimental time series data, $\hat{X}(t_i)$, for intra-cellular and total glycerol. The error for one time series is calculated as

$$\text{error} = \sum_i (X(t_i) - \hat{X}(t_i))^2. \quad (4.8)$$

The best parameters found are given in Table 4.2. To find a possible global minimum point of the error function, the authors in [8] evaluated several randomly chosen starting points in the feasible region of the parameter space. The research was continued for the sets of parameter with sufficient low error by using the function *fmincon* from the MATLAB Optimization Toolbox, MATLAB [22]. Figures 4.3 and 4.4 show

Table 4.2 Optimized parameters by using the relation (4.8): the values of P_e and k_{p1} were fixed while the remaining 8 parameters were estimated from time series data

Parameters		Value
k_{p1}	Water perm. coeff.	1 Osm^{-1}
k_{p2}	Fps1 control const.	0.316
T_d	Time delay	8.61 min
k_{HOG}	HOG control const.	0.416 Osm^{-1}
$Gly(0)$	Initial Gly	2×10^{-4}
$P_i(0)$	Initial P_t	0.636 Osm
$P_e(0)$	Initial P_e	0.240 Osm
V_b	Non osmotic volume	0.368
$V^{P_t=0}$	V when $P_t = 0$	0.99
V_e	External volume	4.79×10^3

Gly represents number of molecules (mol scaled by $V(0)$)

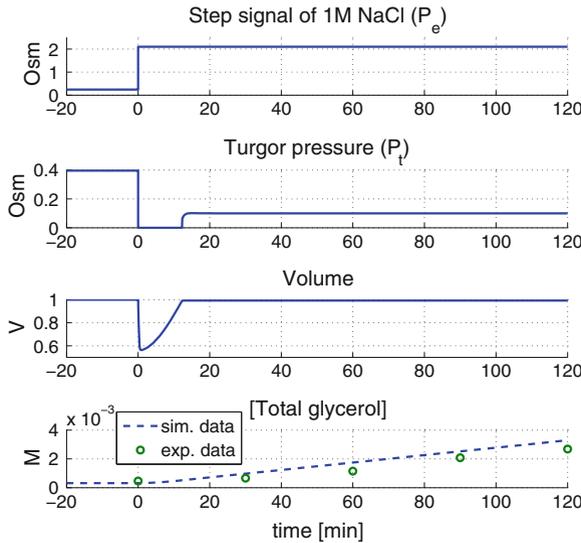


Fig. 4.3 Simulation of a step signal of 1M NaCl at $t = 0$. *Upper plot* external osmotic signal. *Second plot* the turgor pressure. *Third plot* the volume response. *Lower plot* the total glycerol concentration for the simulated (sim.) and experimental (exp.) data taken from [8]

the simulated data using the simple model devised in [8] and the parameter set given in Table 4.2 and a comparison with the time-series experimental data. Figure 4.3 shows the simulation of the model by applying an osmotic stress of 1M NaCl at time $t = 0$, corresponding to an increase in the extra-cellular osmotic pressure by 1.86 Osm, while Fig. 4.4 shows the response to a double stress of 0.5M NaCl at $t = 0$ and $t = 30$. The simulated data show how the turgor pressure and volume drop immediately upon the osmotic stress. While the volume returns to approximately the same value as before the stress, the turgor pressure, the controlled variable, doesn't reach

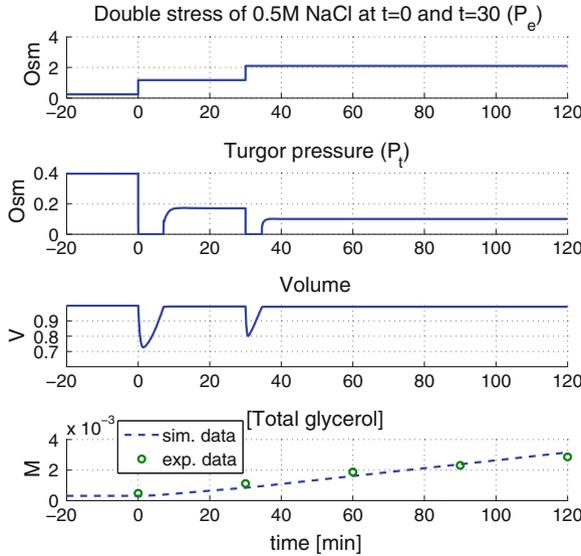


Fig. 4.4 Simulation of a double stress of 0.5M NaCl at $t = 0$ and $t = 30$. *Upper plot* external osmotic signal. *Second plot* the turgor pressure. *Third plot* the volume response. *Lower plot* the total glycerol concentration for the simulated (sim.) and experimental (exp.) data taken from [8]

its previous value. The main reason for incomplete recovery is that the model parameters are estimated by using the measured glycerol data, which are not sufficient for complete recovery of both volume and turgor pressure. The reason why volume and not turgor pressure is recovered is due to the high value of the estimated $V^{P_t=0}$, indicating a low elasticity of the cell wall. Therefore, turgor pressure is not recovered until the volume is almost completely recovered. For a lower value of $V^{P_t=0}$, the turgor pressure would be recovered faster and the volume slower. Figure 4.4 shows that the model can reproduce the regulatory behaviour of the system to a series of osmotic shocks. Moreover, it is able to predict the behaviour of modified strains. For example, Fig. 4.5 shows the simulation to an osmotic shock in a modified strain with constitutively open Fps1 (i.e. only one control mechanism via the HOG pathway). This test was experimentally demonstrated in [15]. To simulate this experiment we set $u_{Fps1} = k_{p2}$ (see Eq. (4.6)) and adjust the value of $Glyc(0)$ to obtain a realistic initial value of total glycerol. Figure 4.5 shows that the model correctly predicts the levels of total glycerol. Note, in particular, an over-production of the glycerol as experimentally measured (double production compared to wild type experiment) and a prolonged activation time of the HOG pathway (see Fig. 4d in [15]), that can not be explicitly observed using this model, since Hog1 is not a variable of the model, but implicitly deduced from the delay of volume recovery.

This model can therefore give us significant insight into the functioning of the system, and the results indicate that even such a simple model can predict the behaviour of different strains and the response to different input functions. It is also easier to

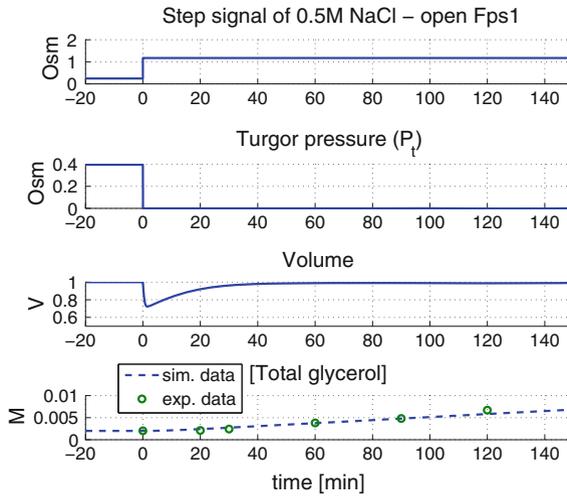


Fig. 4.5 Simulation of one osmotic stress of $0.5M$ NaCl at $t = 0$ in modified system (open Fps1). *Upper plot* external osmotic signal. *Second plot* the turgor pressure. *Third plot* the volume response. *Lower plot* the total glycerol concentration for the simulated (sim.) and experimental (exp.) data taken from [8]

understand and analyse than the detailed model developed in [15] (compare Fig. 4.1 with Fig. 4.2). However, detailed models are often important to completely understand a particular phenomena. For instance, in [15] the authors extracted novel information on the features of the system: the switch-like behaviour of the phosphorelay module consisting of three protein (Sln1, Ypd1 and Ssk1) that become more pronounced for higher number of components (see Fig. 2a in [15], where a comparison of the steady-state characteristics is performed for phosphorelay systems consisting of one, two and three proteins); and the main role of the phosphatases, that is to constantly counteract HOG pathway activation to set thresholds and reduce noise instead of providing a direct downregulation of the pathway.

4.4 Systems-Engineering Approaches

In this section we introduce some methods based on systems-engineering tools to better understand the dynamics of the osmo-adaptation response. In this area, important contributions have been produced by the group of van Oudenaarden. In a first work (see [24]), the authors analysed the dynamics of the system in the frequency domain, a feasible approach which allows the derivation of a concise model of the basic mechanisms of the osmoregulation, that emerge from an intricate network of interactions acting at very different time-scales, e.g. ligand binding or unbinding, phosphorylation, diffusion between compartments and transcription of genes. In [26], the authors later found that Hog1-dependent glycerol accumulation is cru-

cial for the perfect adaptation of yeast to simple step increases of osmotic change, suggesting that Hog1 may implement integral–feedback via an as yet–unknown role for protein–protein interactions that increase the internal osmolyte concentration.

4.4.1 A LTI System Identification

In this section we apply frequency domain analysis to derive a concise model of the HOG MAPK cascade in the budding yeast *S. cerevisiae*. Our treatment is mainly based on the results presented in [24]. After a hyper osmotic shock, membrane proteins trigger a signal transduction cascade that culminates in the activation of the MAPK Hog1. This activated protein, which is primarily cytoplasmic before the shock, is then imported into the nucleus, where it activates several transcription responses to osmotic stress. When the osmotic balance is restored, Hog1 is deactivated through dephosphorylation, thus allowing its export back in the cytoplasm. In order to identify the model, the input and the output of the system to be predicted have to be defined: in this case the input is the extra–cellular osmolyte concentration and the output is the concentration of active Hog1 protein. In [24] the input is manipulated by varying the salt concentration of the medium surrounding the cells, whereas the output is measured by estimating the localisation of Hog1 in the nucleus, $R(t)$, through fluorescence image analysis: the cellular localisation of Hog1–YFP, a yellow fluorescent protein fused to Hog1, and Nrd1–RFP, a red fluorescent protein fused to a strictly nuclear protein, are simultaneously monitored and $R(t)$ is measured as the nuclear to total Hog1 ratio in the cell ($R(t) = \langle YFP \rangle_{nucleus} / \langle YFP \rangle_{cell} population$, averaged over the 50–300 cells observed in the microscope’s field of view).

The experiments are performed by applying pulse wave signals to the cells with different values of the period T_0 , ranging from 2 to 128 min and they show that the steady–state response is approximately sinusoidal, with period T_0 (see Fig. 4.6). Using Fourier analysis, both the input and the output can be approximated as sine waves oscillating with a period $T_0 = 2\pi/\omega_0$. In particular, the experimental input, using a first harmonic approximation (see [5] pp. 26–30), can be written as

$$u(t) \approx 0.2 \left(\frac{1}{2} + \frac{2}{\pi} \sin(\omega_0 t) \right) \quad (4.9)$$

and the steady–state response $R_\infty(t)$ as

$$R_\infty(t) = R_0 + A(\omega_0) \sin(\omega_0 t + \phi(\omega_0)), \quad (4.10)$$

where R_0 is the offset term and A and ϕ are two parameters that characterise the oscillations. A and ϕ are represented through the absolute value and phase of the complex number $\tilde{R}(\omega_0)$, respectively. This complex number is calculated from the Fourier coefficient of the experimental data, $R_\infty(t)$, taken for stimuli with period T_0 using the following relation:

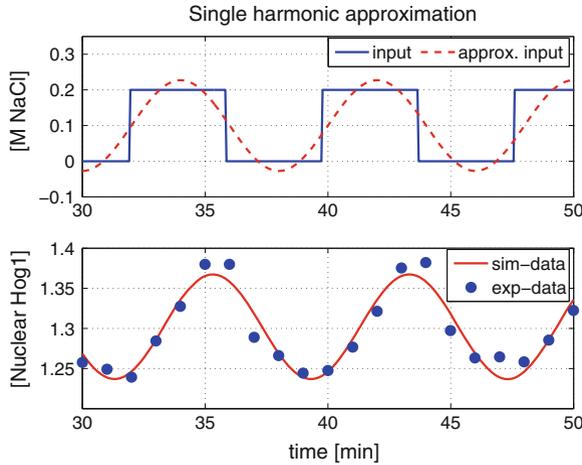


Fig. 4.6 *Upper plot* Pulse signal of 0.2 M NaCl with $T_0 = 8$ min and its approximation using the first harmonic. *Lower plot* the function $R_\infty(t)$ (sim.) defined by (4.11) and fitted to the experimental (exp.) measurements of nuclear Hog1 enrichment taken from [24]

$$\tilde{R}(\omega_0) = 2 \int_{nT_0}^{(n+m)T_0} \frac{\exp^{-i\omega_0 t} R_\infty(t)}{mT_0} dt, \quad (4.11)$$

The amplitude of the signal, defined as $A(\omega_0) = |\tilde{R}(\omega_0)|$, represents half the distance from the peak to the trough of the output sine wave. The phase parameter, $\phi(\omega_0)$, can be written implicitly as $\frac{\tilde{R}(\omega_0)}{|\tilde{R}(\omega_0)|} \exp^{i(\phi(\omega_0) - \pi/2)}$. The parameter n is chosen so that the system is allowed to approach steady state before computing $\tilde{R}(\omega_0)$. The parameter m , which represents the number of periods over which the Fourier transform is computed, is set to be at least two for periods less than 64 min. For periods greater than or equal to 64 min, it is found that the first period is a good representation of the steady state oscillations and thus $\tilde{R}(\omega_0)$ is computed over this period alone. However, the values $A(\omega_0)$ and $\phi(\omega_0)$ can be computed for different values of ω_0 by fitting the parameters of the Eq. (4.11) to the experimental time response as shown in the lower plot of Fig. 4.6 for $\omega_0 = 2\pi/8$ rad/min. The resulting frequency response is shown on the Bode plots in Fig. 4.7.

A predictive model can be identified from the available experimental data by using linear systems theory: a linear input–output relationship in Fourier space is defined by

$$\tilde{Y}(\omega) = A_0 \frac{\prod_{i=1}^n (z_n + i\omega)}{\prod_{i=1}^n (p_n + i\omega)} \tilde{U}(\omega), \quad (4.12)$$

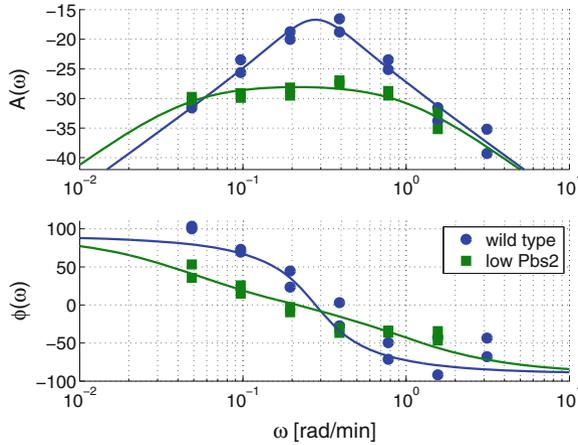


Fig. 4.7 The experimental data of the Fourier amplitude $A(\omega)$ and phase $\phi(\omega)$ (two measurements at each frequency), for wild type (*circles*) and underexpressed Pbs2 mutant (*squares*) strains, with the fitting models (*solid lines*). Experimental data taken from [24]

Table 4.3 Best-fitting parameters for the Eq. (4.13)

	p_1 [min^{-1}]	p_2 [min^{-1}]	A_0 [min^{-1}]
Wild type	$-0.1434 + 0.239i$	$-0.1434 - 0.239i$	0.3292
Low Pbs2	-0.0466	-0.9755	0.3169

where $\tilde{Y}(\omega)$ and $\tilde{U}(\omega)$ are the output and input Fourier spectra, respectively, z_n are the n roots of the numerator of Eq. (4.12), also called *zeros*, and p_n are the n roots of the denominator, also called *poles*. The simplest such model from this class, that describes the experimental points in Fig. 4.7, exhibits a zero at the origin ($z_1 = 0$) and a pair of poles p_1 and p_2 yielding

$$(p_1 + i\omega)(p_2 + i\omega)\tilde{Y}(\omega) = (i\omega)A_0\tilde{U}(\omega). \tag{4.13}$$

The best-fit parameters for the wild type and for the mutant (Pbs2 underexpression) strains are shown in Table 4.3. Applying the inverse Fourier transform (note a multiplication by $i\omega$ in Fourier domain corresponds to the derivative operator in time domain) the following relationship in the time-domain is given:

$$\ddot{y}(t) + (p_1 + p_2)\dot{y}(t) + (p_1 p_2)y(t) = A_0\dot{u}(t). \tag{4.14}$$

The identified second-order linear time-invariant (LTI) models, defined by the Eq. (4.13), are used to predict the response of the two strains to a step input of 0.2 M NaCl. Figure 4.8 shows the predicted responses of the two models and a comparison with the experimental measurements: the responses of the linear systems are offset by a constant value (1.23), which is the experimentally measured basal activity level of Hog1. The two models show a good qualitative match to the different sets of data

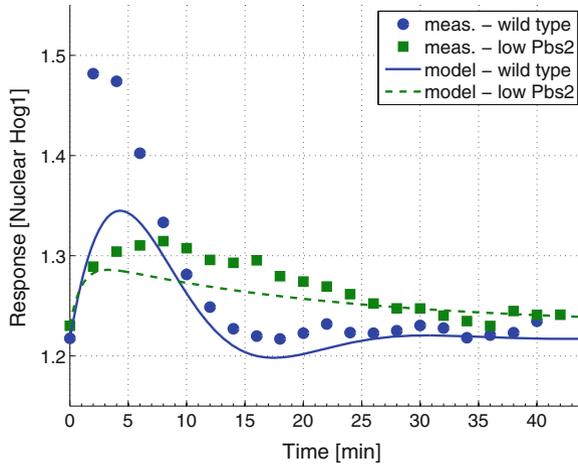


Fig. 4.8 Time domain response of the system to a step increase of 0.2 M NaCl: comparison of the responses predicted by the two linear models developed in the frequency domain vs the experimental measurements taken from [24]

for the two yeast strains (the match is not perfect, since these are linear models of a process that will clearly also involve some nonlinear dynamics). Note that the wild type model exhibits a pair of complex conjugated poles and therefore the response is oscillatory, with a larger overshoot and a faster response than the low Pbs2 model, as expected from the experimental data. Indeed, the latter has two real poles, and thus exhibits a limited initial overshoot, a fast initial rise (due to the pole with small time constant) and a slow decay (caused by the large time constant associated with the other real pole). The identified LTI model can be written as a pair of differential equations, that is more readily interpreted in terms of biological process:

$$\begin{pmatrix} \dot{z}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} u(t) \quad (4.15)$$

with rate constants a , b , c , d , e and f . The variable $y(t)$ is assumed to represent the observable output of the system (the level of Hog1 activity), whereas the variable $z(t)$ represents the hidden state and $u(t)$ the osmotic stimulus. When these equations are simplified to remove the hidden variable $z(t)$, a single second order differential equation in $y(t)$ is obtained:

$$\ddot{y}(t) = (a + d)\dot{y}(t) + (bc - ad)y(t) + (ec - af)u(t) + f\dot{u}(t) \quad (4.16)$$

This equation is equivalent to the one in (4.14), if $f = A_0$, $ce = aA_0$ and $c \neq 0$. Substituting these relations in (4.15), we obtain the following system:

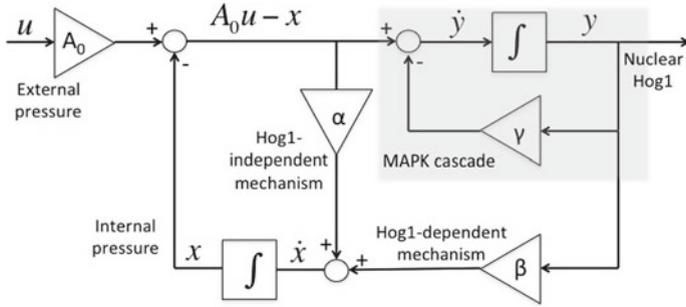


Fig. 4.9 Block diagram representation of the system (4.18): two negative feedback loops (Hog1-independent and Hog1-dependent mechanisms) act to reduce the difference between the stimulus, $A_0u(t)$, and the internal-state variable, $x(t)$, representing the internal pressure

$$\begin{pmatrix} \dot{z}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} \frac{ce}{A_0} & b \\ c & d \end{pmatrix} \begin{pmatrix} z(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} e \\ A_0 \end{pmatrix} u(t) = \begin{pmatrix} \frac{e}{A_0} & b \\ 1 & d \end{pmatrix} \begin{pmatrix} A_0u(t) + cz(t) \\ y(t) \end{pmatrix}. \quad (4.17)$$

Defining $x(t) = -cz(t)$, $\alpha = -\frac{ec}{A_0}$, $\beta = -bc$ and $\gamma = -d$ the system (4.17) is written as

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ 1 & -\gamma \end{pmatrix} \begin{pmatrix} A_0u(t) - x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \alpha(A_0u(t) - x(t)) + \beta y \\ A_0u(t) - x(t) - \gamma y \end{pmatrix}. \quad (4.18)$$

Comparing this relation with the LTI model, we can equate coefficients to obtain the relations:

$$\alpha + \gamma = p_1 + p_2, \quad p_1 = \frac{1}{2}((\alpha + \gamma) + \sqrt{(\alpha - \gamma)^2 - 4\beta}),$$

and

$$\alpha\gamma + \beta = p_1p_2, \quad p_2 = \frac{1}{2}((\alpha + \gamma) - \sqrt{(\alpha - \gamma)^2 - 4\beta}).$$

The identified model, described by the relation (4.18), contains two negative feedback loops, which act to reduce the difference, $(A_0u(t) - x(t))$, between the stimulus, $A_0u(t)$, and the internal-state variable $x(t)$ (see Fig. 4.9). This enables us to assign a physical meaning also to the variable, $x(t)$: since the input is the external pressure, x represents the internal pressure. Moreover the model tells us that one feedback mechanism is mediated by the Hog1 MAPK pathway (βy changes x through the activity of the observable output y), whereas a second one is mediated by a pathway which is independent of Hog1. Since Hog1 is activated by Pbs2, we can derive useful insight by comparing the responses of the wild type strain with the mutant strain, in which Pbs2 is underexpressed (see Fig. 4.8). This comparison suggests that the feedback action provided by the Hog1 pathway is stronger,

producing a faster response. As discussed in Sect. 4.3, the hyperosmotic-shock response in yeast is regulated by two parallel mechanisms: 1) the Hog1-independent pathway activating the membrane protein Fps1 that quickly (<2 min) responds by decreasing the glycerol-export rate (see [19, 36]); 2) the Hog1-dependent pathway increasing the expression of Gpd1 and Gpd2 which accelerate the production of the glycerol over a longer time scale (>30 min—see [2]). Although the topology of the model identified corresponds closely to that of the known biological system (see Fig. 4.9) the dynamic differences suggest that the MAPK Hog1 plays a role not only in the transcriptional regulation of glycerol producing proteins, but also in the control of the rapid accumulation of glycerol, consistent with previous studies (see [19, 36, 37]): from Fig. 4.8 the peak times of the responses of both wild type and mutant strains are less than 10 min and in both cases the response is much faster than the characteristic dynamics of gene expression. From this analysis the authors, in [24], have hypothesized that gene expression may be more important as a longer-time scale feedback in the hyperosmotic-shock response. To test this hypothesis, they stimulated cells with periodic pulses of NaCl (see Fig. S5 in [24]). The cells were shocked either in the absence or presence of cycloheximide, a small molecule that inhibits protein synthesis. They showed that cells respond very similarly to an initial pulse of osmolyte both in the absence or presence of cycloheximide. On the other hand, to adapt to subsequent pulses, cells need less time in the absence of cycloheximide and more in its presence. These results suggest that non transcriptional feedback mediates short-time scale osmolyte accumulation (see [8, 15, 28, 37]), whereas gene expression plays a role in osmolyte production only over longer time scales and for more intense shocks.

4.5 Perfect Adaptation in Yeast Osmoregulation

As shown in the last section the concise model developed by [24] is able to predict the Hog1 response by using only two differential equations. However, a detailed comparison of the LTI model's predictions with the experimental data sets shows that this model (only containing two negative feedback loops that control the rapid accumulation of glycerol) is too simple to fully reproduce the quantitative dynamics of the Hog1 nuclear enrichment when the cell are stimulated multiple times with periodic pulses of NaCl (see Figs. S8 and S9 in [40]). In particular, the experimental data sets and the model presented in [40] suggest that yeast can remember the first pulse of high osmolarity and needs less time to adapt to subsequent pulses of simulation. The LTI model developed by [24] fails to capture this dynamical property and, in [26], the same group proposed a revised concise LTI model by implementing an integral feedback mechanism which requires Hog1 kinase activity. They started with a minimalist model represented by the network diagram of Fig. 4.10, which aims to predict the dynamics of the osmoregulation system with only a few key parameters, starting from input-output data, and, using biological measurements and engineering principles, to better understand the relation of its dynamics with the

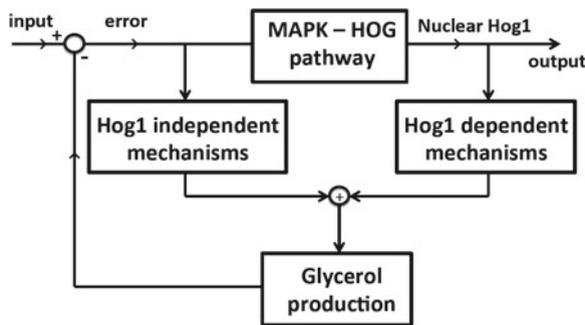


Fig. 4.10 Network diagram of the osmoregulation system presented in [26]

network topology. The authors represent in a subsystem all relevant reactions that determine the activation of the MAPK signalling pathway and the nuclear import of the activated protein Hog1 (MAPK signaling & nuclear import) and in another subsystem all the Hog1-dependent mechanisms that promote the glycerol production (Hog1-dependent mechanisms—such as the transcriptional activation of genes that encode enzymes involved in the glycerol production and potential protein-protein interaction initiated by Hog1 in the cytoplasm or nucleus that lead to glycerol accumulation). In contrast in the model of Fig. 4.2, developed by [8], the HOG pathway controller represents both the HOG signalling pathway, transcription/translation and the synthesis of enzymes involved in glycerol production.

4.5.1 Experimental Measurements for the Perfect Adaptation

In [26] the authors observed perfect adaptation of Hog1 nuclear enrichment in response to step increases of the extracellular osmolyte concentration (see Fig. 4.11 where step inputs of NaCl with different amplitude are applied and Fig. S3 in [26] where *KCl* and sorbitol are also used as osmolytes—in these and in the following figures of this section, Hog1 nuclear enrichment is defined as the relative change from the pre-shock level): this adaptation occurs with very low cell-to-cell variability and is robust to the signalling fidelity of the MAPK cascade. In particular, for different cells, the dynamics of Hog1 nuclear enrichment and cell volume are very similar in response to a step osmotic stress, with trends that closely follow the population average (see Fig. 2A, B in [26]). In fact, the cell-to-cell variability in unstressed cells is comparable to the one in osmo-stressed cells as shown in Fig. S2 in [26], further indicating that the intrinsic noise of signal propagation is low and suggesting that the experimental setup itself may be the predominant source of noise in the experimental data. Moreover, to demonstrate the robustness of this perfect adaptation, measurements of the Hog1 response have been performed in cells with compromised MAPK signalling, by controlling the expression of PBS2, which encodes the kinase of Hog1

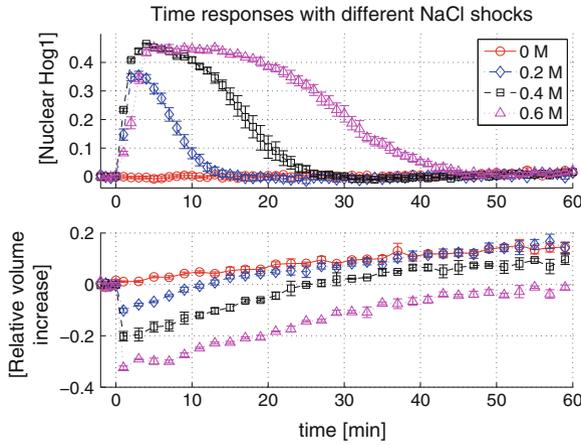


Fig. 4.11 Time measurements of Hog1 nuclear enrichment and volume to hyperosmotic shocks with indicated concentrations of salt. Data taken from [26]

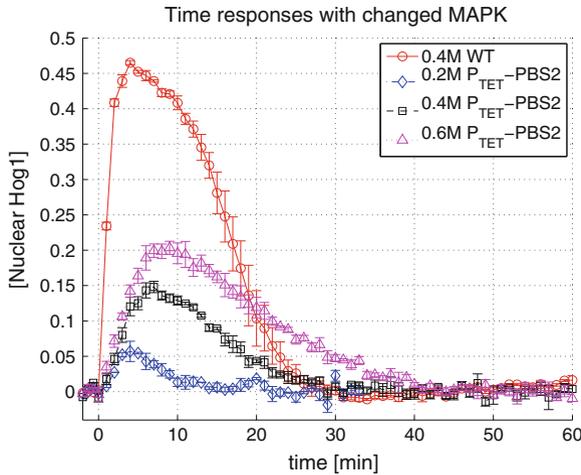


Fig. 4.12 Time measurements of Hog1 nuclear enrichment after changing the signaling fidelity of the MAPK cascade by controlling the expression of PBS2. Data taken from [26]

(see Fig. 4.12). Also in this case Hog1 nuclear enrichment still perfectly adapts and therefore we can say that the perfect adaptation is a robust property of the system and not a consequence of *ad hoc* parameter tuning. From these results, together with extensive theoretical analysis of adaptive systems in engineering, Muzzey et al hypothesised that this system implements integral feedback control in order to achieve robust perfect adaptation that does not require a precise tuning of system parameters such as protein levels or rate constant (see [14, 34, 39]).

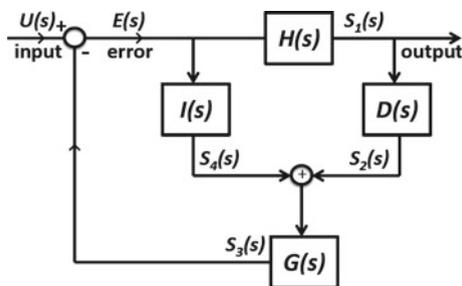


Fig. 4.13 Block diagram of the osmoregulation system presented in [26]. H represents all relevant reactions that link an osmotic disturbance at the membrane with Hog1 nuclear enrichment. D and I represent the Hog1-dependent and independent mechanisms for the glycerol accumulation, respectively. G represents the metabolic reactions involved in the glycerol synthesis

4.5.2 The Integral Feedback

The osmoregulation system is described by using the network diagram of Fig. 4.10, where the error indicates the deviation from the initial turgor pressure before applying the hyperosmotic stress. Figure 4.13 shows the corresponding block diagram of the osmosensing network of 4 subsystems denoted with G , D , H and I . H takes into account reactions that determine the activation of the MAPK signalling pathway and the nuclear import of the activated protein Hog1. D and I represent the Hog1-dependent and independent mechanisms that contribute to glycerol accumulation, respectively. Finally, G represents the metabolic reactions involved in the glycerol synthesis and any other reactions that promote glycerol accumulation. Approximating the network as being LTI, each subsystem can be described by a Laplace transform, or transfer function. In general a Laplace transform $F(s)$ of a function $f(t)$ is given by

$$F(s) = \int_0^{\infty} f(t) \exp^{-st} dt, \quad (4.19)$$

where s is a complex variable. The transfer function $S(s)$ of a LTI system is defined as $S(s) = Y(s)/U(s)$, where $U(s)$ and $Y(s)$ are the Laplace transform of the system input, $u(t)$, and output $y(t)$, respectively (see [5] pp. 30–33). The Laplace transform has the useful property that many relationships and operations in the time domain that require calculus can instead be performed using linear algebra in the s -domain (the differential equations in the time domain can be transformed into algebraic equations in the s -domain using the Laplace transform—these are then much easier to solve). By applying the final-value theorem (see [32] p. 43), the steady-state input and output are related via $y_{ss} = S(0)u_{ss}$, so perfect adaptation of the system output ($y_{ss} = 0$)

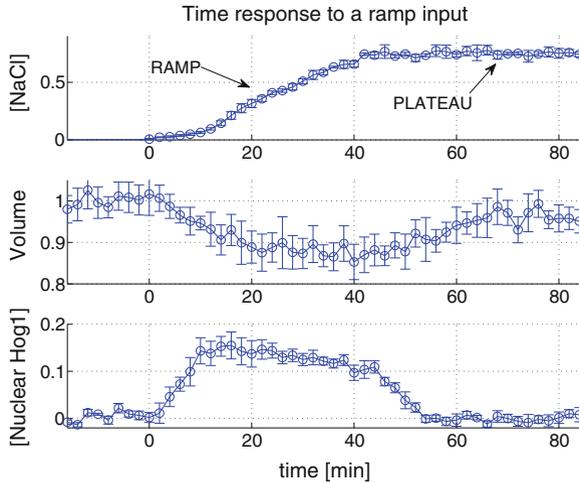


Fig. 4.14 Measurements of Hog1 nuclear enrichment and volume in an experiment where the salt concentration ramps upward over time, reaching a plateau after nearly 45 min. Experimental data taken from [26]

for an LTI system is equivalent to $S(0) = 0$, since the input is constant and nonzero. The input-error transfer function of the osmosensing system of Fig. 4.13 is given by

$$S_{ue}(s) = \frac{E(s)}{U(s)} = \frac{1}{1 + G(s)(D(s)H(s) + I(s))}, \tag{4.20}$$

where $G(s)$, $D(s)$, $H(s)$ and $I(s)$ are the transfer functions of the four subsystems in the network (see [5] pp. 42–46). We need that $S_{ue}(0) = 0$ to achieve perfect adaptation of the error to a step input. Therefore at least one of the four subsystems implements an integrator (its transfer function is given by $1/s$ —see [5] p. 31—thereby allowing S_{ue} to be zero at $s = 0$). In general, a system contains at least one feedback loop with at least $n + 1$ integrators connected in series in order to achieve perfect adaptation to an input corresponding to the n -th integral of a step function, where n is a positive integer. Perfect adaptation to a step input, where $n = 0$, requires at least one integrator, perfect adaptation to a ramp input, where $n = 1$ since the ramp is the integral of a step, requires at least two integrators in series, and so on. In [26] the authors showed that neither cell volume nor Hog1 perfectly adapt in response to a ramp input, confirming that there is exactly one integrator in the osmosensing network (see Fig. 4.14). Therefore, the perfect adaptation of the error to a step input requires that only one of the four subsystems contains one integrator. Similarly, the input-output transfer function is given by

$$S_{us_1}(s) = \frac{S_1(s)}{U(s)} = \frac{H(s)}{1 + G(s)(D(s)H(s) + I(s))}. \tag{4.21}$$

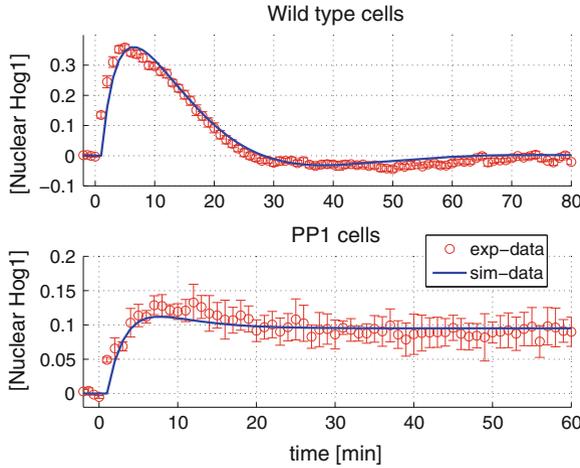


Fig. 4.15 Time domain response of the system (4.26) to a step increase of 0.4M NaCl: comparison of the responses (sim.) predicted by the two models (wild type and PP1) vs the experimental (exp.) measurements taken from Fig. 5A in [26]

If $H(s)$ were the only subsystem performing integration, then Hog1 would not perfectly adapt ($S_{us_1}(0) \neq 0$). Therefore one or more of the other subsystems must contain an integrator to achieve perfect adaptation of Hog1 but the system only contains one integrator. In [26] it is shown that the cells lose perfect adaptation with PP1, a treatment to completely eliminate the Hog1 kinase activity, as the steady-state Hog1 accumulation ($s_{1_{ss}}$ in Fig. 4.13) does not go back to the pre-stimulus level. Since the presence of PP1 disconnects the D subsystem from Hog1, the input-output transfer function of the system is modified as

$$S_{us_1}(s) = \frac{H(s)}{1 + G(s)I(s)}. \quad (4.22)$$

In this case Hog1 does not perfectly adapt, then the product $G(s)I(s)$ does not go to infinity at $s = 0$ ($S_{us_1}(0) \neq 0$), which implies that either the G and I subsystems both lack integrators, or one subsystem has an integrator but the other perfectly cancels the integrator (it is a differentiator with a transfer function equal to s —see [5] p. 31). If I contained the integrator, then the turgor pressure would perfectly adapt in the presence of PP1, and Hog1 likely would as well, but both properties are not observed in the data. If G were to act as an integrator, then cell volume and turgor pressure would continue to perfectly adapt for a nonzero input to the G subsystem. But, in the presence of PP1, the only input to subsystem G is the output from subsystem I , as subsystem D is disconnected. Thus, no volume recovery observed in PP1-treated cells would only occur if the output of subsystem I prematurely goes to zero (i.e. if it were a differentiator). As explained in [26], this observation would require that all Hog1-independent mechanisms completely desensitize within approximately 20 min (i.e. the time needed for Hog1 nuclear enrichment to reach steady state in PP1 cells—

see the lower plot of Fig. 4.15) despite persistence in their stimulus (i.e. the acute loss of turgor pressure). On the basis of this argument, it is extremely improbable that subsystem G acts an integrator. Therefore, the combination of all findings points to D as the subsystem with the only integrator in the feedback loop. Moreover, in PPI cells, levels of total glycerol and extracellular glycerol are measured over time in the presence and absence of osmotic shock (see [26]): in the absence of salt shock, glycerol synthesis is increased as well as glycerol leakage; in the presence of osmotic shock, glycerol leakage is rapidly and transiently diminished, as in wild type cells, whereas the absence of Hog1 kinase activity prevents an increase in glycerol synthesis, unlike in wild type cells. These data suggest that Hog1 kinase activity plays a critical role in rapidly regulating glycerol synthesis but not its leakage as in [38].

Note from Fig. 4.12 (see also Fig. 3D in [26]) that the time-integral of the Hog1 scales linearly with the shock strength. If the system were composed only of reactions modelled with linear dynamics, then the result that D subsystem is an integrator would be trivial. However, this result is valid also when the other subsystems are nonlinear stable systems without integrators (see Fig. 4.12 where the fact that the peak Hog1 amplitude saturates as a function of salt is an evidence of nonlinear dynamics in the H subsystem). If it is assumed that the error perfectly adapts and the steady-state output of the I subsystem is zero when its steady-state input is zero, then 1) the net change induced by the system in the steady-state input of the G subsystem simply equals the time-integral of Hog1, 2) the net change in the output of G must equal the net change in the system input in order for the error to go to zero. If the G subsystem were perfectly linear, then its output would be directly proportional to its input at steady state and so the time-integral of Hog1 would be directly proportional to the magnitude of the osmotic stress (despite potential nonlinearities in the H and I subsystems). This relationship is almost exactly what Fig. 3D in [26] shows, except that the line relating the integral of Hog1 nuclear enrichment to the magnitude of the osmotic stresses does not cross the origin. This difference may be due to nonlinearities in the input-output steady-state function of subsystem G that become evident for osmotic stresses of small magnitude (<0.2 M NaCl).

Finally, in order to validate these results, a LTI system can be used to implement the concise model represented by the block diagram of Fig. 4.13. The subsystems of the osmosensing network can be represented as follows: H and G as first-order systems where the corresponding transfer functions $H(s) = \frac{k_h}{s + \gamma_h}$, with gain k_h and time constant γ_h^{-1} , and $G(s) = \frac{1}{s + \gamma_g}$ with time constant γ_g^{-1} , I as a scalar α_i (i.e. $I(s) = \alpha_i$) and D as an integrator with gain α_d (i.e. $D(s) = \frac{\alpha_d}{s}$). Therefore the Laplace transform of the output, $S_1(s)$, of the H subsystem is defined as:

$$S_1(s) = \frac{k_h}{s + \gamma_h} E(s) = \frac{k_h}{s + \gamma_h} (U(s) - S_3(s)), \quad (4.23)$$

where the Laplace function error $E(s) = U(s) - S_3(s)$, with $U(s)$ and $S_3(s)$ the Laplace functions of the input $u(t)$ of the system and the output $s_3(t)$ of subsystem

Table 4.4 Best-fitting parameters for the system (4.26)

	k_h [min^{-1}]	γ_h [min^{-1}]	γ_g [min^{-1}]	α_d [min^{-1}]	α_i [min^{-1}]
Wild type	0.496	0.369	0.119	0.0106	0.0806
PP1	0.147	0.369	0.119	0	0.0806

G , respectively. We can obtain the rate equation for the output s_1 (corresponding to measured Hog1 nuclear enrichment) in the time domain applying the inverse Laplace transform of the following relation, by rewriting the Eq. (4.23):

$$sS_1(s) = -\gamma_h S_1(s) - k_h S_3(s) + k_h U(s). \quad (4.24)$$

By applying the property that the derivative operator with respect to time correspond to a multiplication by s in the s -domain (see [5] p. 31), the inverse Laplace transform of (4.24) follows as:

$$\dot{s}_1(t) = -\gamma_h s_1(t) - k_h s_3(t) + k_h u(t) \quad (4.25)$$

In the same way, we can obtain the rate equations for the outputs s_2 and s_3 of the corresponding subsystems D and G . Then the following system of differential equations is obtained:

$$\begin{pmatrix} \dot{s}_1(t) \\ \dot{s}_2(t) \\ \dot{s}_3(t) \end{pmatrix} = \begin{pmatrix} -\gamma_h & 0 & -k_h \\ \alpha_d & 0 & 0 \\ 0 & 1 & -(\alpha_i + \gamma_g) \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{pmatrix} + \begin{pmatrix} k_h \\ 0 \\ \alpha_i \end{pmatrix} u(t). \quad (4.26)$$

Figure 4.15 shows the response of two strains (wild type and PP1 cells) to a step input of 0.4 M NaCl. Table 4.4 reports the best set of parameters that fit the experimental data. For the PP1 experiment we set $\alpha_d = 0$ to break the connection between Hog1 and the D subsystem. The simulations show how the devised model is able to capture the dynamics of the system and produces an excellent match to the experimental data.

4.6 The Role of Ultrasensitivity

As shown above, systems and control theory provides a highly useful approach to abstract complex biological systems that seem to operate with similar goals as engineered control systems, and the osmoregulation system in yeast is a prime example of this. The models here presented, by combining proportional and integral feedback controllers capture the key dynamics of a homeostatic system like osmoregulation in yeast, but they do not shed light on how the evolution of such a biological control system can proceed to result in integral feedback control. In the following we explore the possible role of ultrasensitivity in osmoregulation. Indeed, it has been well-

documented that the upstream signalling pathways involved in this system implement high levels of ultrasensitivity, however, the role of such high gain in producing the observed perfect adaptation is not clear. Therefore, we extend the proportional controller model presented in [8] for this system with the implementation of ultrasensitivity, Montefusco [25].

4.6.1 Ultrasensitive Model and Parameters

The mathematical model used for our analysis is the same as that presented in Sect. 4.3, apart from allowing the Hog controller to be non-linear (see Fig. 4.2). Indeed, in this case, the output of the HOG branch, which corresponds to the HOG pathway dependent glycerol production, is expressed as

$$u_{HOG}(t) = \begin{cases} k_{HOG} \cdot f(e, t), & e(t) > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (4.27)$$

where the control function is given by

$$f(e, t) = \frac{e(t)^n}{\beta e(t)^n + K^n}, \quad (4.28)$$

with $\beta = 1$ and K and n being the nonlinear Hill function variables. We have thus modified the control law for the HOG pathway, compared to the model in [8], to allow for a non-linear controller response. This is inspired by the fact that MAPK systems, of which the HOG pathway is an example, often show Hill type responses, Huang and Ferrell [13]. The performance of the nonlinear controller is contrasted with the proportional controller given in [8], where $\beta = 0$ and $K = n = 1$. Our model contains 16 parameters as reported in Table 4.5. However, four of these are dependent parameters which do not need to be constrained. The other parameters are estimated by simulating the model with different osmotic shocks and minimising the error, defined by Eq. (4.5), and time adaptation corresponding to the time required by the cell to approximately return to its volume before the stress (see the definition in the next subsection). For the optimization, we use a hybrid Genetic Algorithm (GA) (see [18]), that combines the most well-known type of evolutionary algorithm with local gradient-based algorithms (see [7, 10]). We use the function *ga* from the MATLAB Global Optimization Toolbox, MATLAB [23], and *fmincon* from the MATLAB Optimization Toolbox, MATLAB [22], as the local algorithm. By the optimisation procedure some parameters do not significantly change their values, therefore, they are fixed equal to the values estimated in [8], except for $V^{P_t=0}$, which is set to 0.8, the value of the volume at zero P_t according to a recent study presented in [31].

The cost function used for the parameter estimation is given by

Table 4.5 Ultrasensitive model parameters: all volumes are scaled such that the initial volume of the cell is 1

Parameters		Bounds
k_{p1}	Water perm. coeff.	[0.0052 160] Osm ⁻¹
k_{p2}	Fps1 control const.	[0 10]
T_d	Time delay	[5 30] min
k_{HOG}	HOG control const.	[0 2] Osm ⁻¹
K	Hill const.	[0 0.01 2]
n	Hill exponent	[0 4]
Fixed parameters		Value
$Gly(0)$	Initial Gly	2×10^{-4}
$P_i(0)$	Initial P_i	0.636 Osm
$P_e(0)$	Initial P_e	0.24 Osm
V_b	Non osmotic volume	0.368
$V^{P_i=0}$	V when $P_i = 0$	0.8
V_e	External volume	4.79×10^3
Dependent parameters		Value
$V(0)$	Initial V - relative volume	1
$Gly_e(0)$	Initial Gly_e	$\frac{V_e Gly(0)}{(V(0) - V_b)}$
$P_i(0)$	Initial P_i	$P_i(0) - P_e(0)$
s	No. of osmolytes other than Gly	$P_i(0)(V(0) - V_b)$ $- Gly(0)$

Both Gly and Gly_e represent number of molecules (mol scaled by $V(0)$)

$$\min_x J, \quad (4.29)$$

where

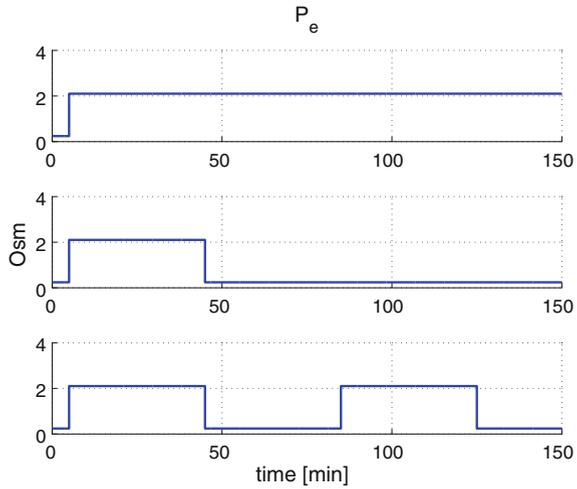
$$J = J_p + J_v + J_t \quad (4.30)$$

is a sum of three scalar functions: J_p is the turgor pressure error, J_v is the difference between the desired and the effective volume and J_t is the response time of the system after the perturbation.

4.6.2 Results: Ultrasensitive Versus Proportional Controller

In our adaptation of the model developed by [8], we particularly consider the observed ultrasensitivity in the HOG branch of the system. This branch was originally modeled as a proportional control in [8], which we have replaced here by a Hill-type function to model ultrasensitivity (see Eqs. (4.27) and (4.28)). We then compare the performance of this new model against the original model. In particular, we evaluate the two different controllers—proportional (Pr) and ultrasensitive (Us)—by simulat-

Fig. 4.16 Different osmotic stresses. *Upper plot* a constant step of 1M NaCl at $t = 5$ min corresponding to an increase of P_e equal to 1.96 Osm. *Middle plot* single pulse signal at $t=5$ min with duration of 40 min of 1M NaCl. *Lower plot* double pulse signal at $t_1 = 5$ and $t_2 = 85$ min, both with duration of 40 min and amplitude of 1M NaCl



ing their dynamics with different stress inputs (see Fig. 4.16) and optimizing their parameters for optimum response (i.e. minimal deviation of cell volume and turgor pressure in presence of an osmo-schock, see Sect. 4.6.1 for details). We repeat this procedure for different levels of overall sensitivity (i.e. gain) of the HOG branch and different types of osmo-shock sequences and evaluate the tests by using two different performance indices: adaptation precision and adaptation time. The adaptation precision is defined as

$$X_a = \prod_i X_{s,i}, \quad (4.31)$$

where $X_{s,i}$ is the steady state value of the variable X (volume V or turgor pressure P_t) after the i -th perturbation. Since the initial volume is set to unity, this measure gives 1 for perfect adaptation. Deviations from 1 indicate inability of the system to perfectly adapt volume to pre-perturbation levels. The time adaptation, T_a , defined as

$$T_a = \sum_i t_{a,i}, \quad (4.32)$$

where $t_{a,i}$ is the time required by the system to reach 85% of the volume V after the i -th osmotic stress. Figure 4.17 shows the results of the two controllers by applying three different osmotic stresses: constant step, single pulse and double pulse. For all different inputs the ultrasensitive controller achieves better and faster adaptation irrespective of the level of overall gain. The better performance is particularly significant when overall gain is limited to lower values, where the ultrasensitive controller achieves almost 2-fold faster responses. Indeed, using a Hill function within the HOG branch allows us to effectively achieve a steeper response from this branch compared to a linear function for any given error (see Eqs. (4.27) and (4.28)). Thus,

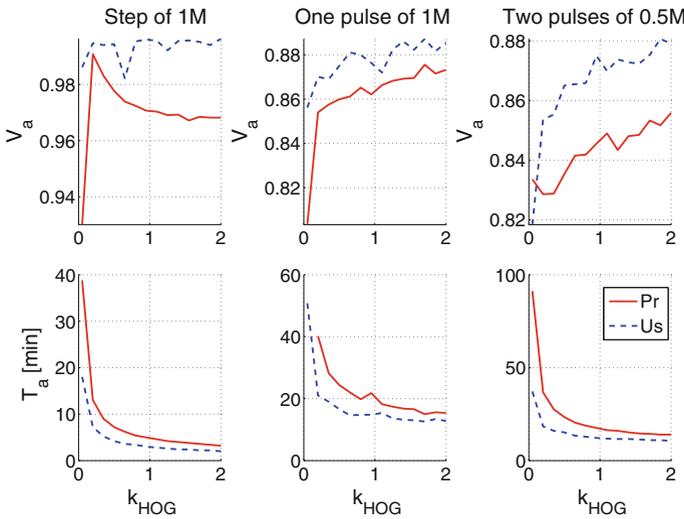


Fig. 4.17 Performance comparison between the proportional (Pr) and ultrasensitive (Us) controllers by applying different shocks: first column, a step of 1M of NaCl; second column, 1 pulse of 1M; third column, 2 pulses of 0.5M. V_a close to 1 indicates the capability of the system to adapt. T_a indicates the time adaptation

Table 4.6 Optimized parameters for a given k_{HOG} with a double pulse signal of 1M of NaCl

k_{HOG}	Optimized parameters—Pr/Us						Us	
	k_{p1}		T_d		k_{p2}		\bar{K}	n
	Pr	Us	Pr	Us	Pr	Us		
0.65	93	155	5	5	0.43	0.96	0.17	3.53
1.1	159	124	5	5	0.69	1.17	0.23	3.78
1.55	0.36	134	5	5	0.9	1.61	0.23	3.8
2	155	159	5	5	1	1.54	0.25	3

the controller acts faster and more strongly, allowing quicker and fuller recovery of the system. This insight is in line with the optimized parameters for both controllers as reported in Table 4.6: in most cases, the optimal parameters for the ultrasensitive controller result in a very steep Hill function that produces maximal outputs for even small error values. Of the other free parameters of the model, we note that certain parameters are optimized differently for the two controllers. For example, the permeability coefficient k_{p1} , which controls water flow in the model (see Eq. (4.2)) is usually optimized to higher values in the ultrasensitive controller compared to the proportional controller. This parameter affects the sensitivity of the system, as faster water movement can allow both a high volume reduction for a given osmo-shock and also fast recovery. Given its fast dynamics, the ultrasensitive controller can “afford” this parameter to become higher compared to the proportional controller.

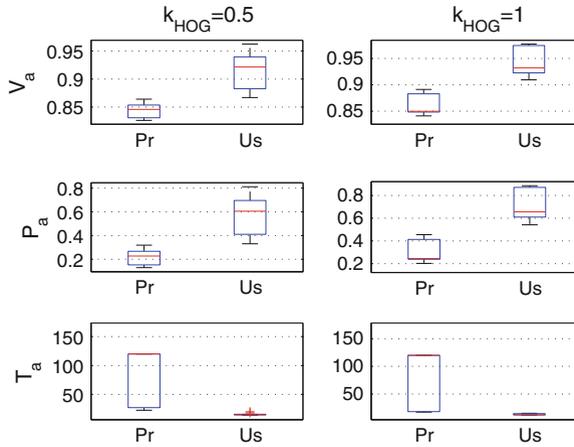


Fig. 4.18 Sensitivity analysis using the box-and-whisker representation (*median—middle line*, the 25th and 75th percentile—*lower and upper lines* of the box, and the outliers (crosses)) by fixing k_{HOG} and Hill function parameters ($K = 0.05$, $n = 3.5$) and applying one pulse signal of 1M of NaCl. The first column of each boxplot reports the results for the proportional (Pr) controller, the second column the results for the ultrasensitive (Us) controller. The system robustly adapts if V_a and P_a are close to 1. T_a indicates the time adaptation. Similar results are obtained with different Hill function parameters ($K = 0.2$, $n = 2$)—see Fig. 7 in [25]

Such differences between the optimal model parameters for the two controllers suggest that implementation of ultrasensitivity might allow more freedom in the other parameters of the model or allow them to be in a more favorable regime. To test the former possibility, we perform a simple sensitivity analysis for the two controllers. Given a certain gain, and Hill function parameters, we evaluate the adaptation precision and time of the two controllers for a set of 100 randomly generated parameters. Figure 4.18 shows that the ultrasensitive controller achieves much more robust adaptation performance than the proportional controller according to these two criteria.

As discussed above, the performance increase of the ultrasensitive controller over the proportional one stems from its high sensitivity to the error due to the Hill function. The incorporation of the Hill function, however, can also allow development of thresholds in the system. In particular, the ultrasensitive controller can be tuned as a filter allowing responses only to signals of certain magnitude or duration. To test this hypothesis, we devise an alternative cost function for the optimization procedure and optimize the system towards functioning as a filter. The new cost function is given by $J_n = J - J_{glyc}$, where J is defined by the Eq. (4.30) and J_{glyc} represents the glycerol production upon the signal of limited duration. Figure 4.19 shows the performance for a signal with a first short and then long duration pulse. The ultrasensitive controller ignores the first pulse and responds to the second by tuning the Hill parameters, whereas the proportional controller model is not able to respond to the second signal (the permeability coefficient k_{p1} , that affects the sensitivity of the system, is equal to the lower bound).

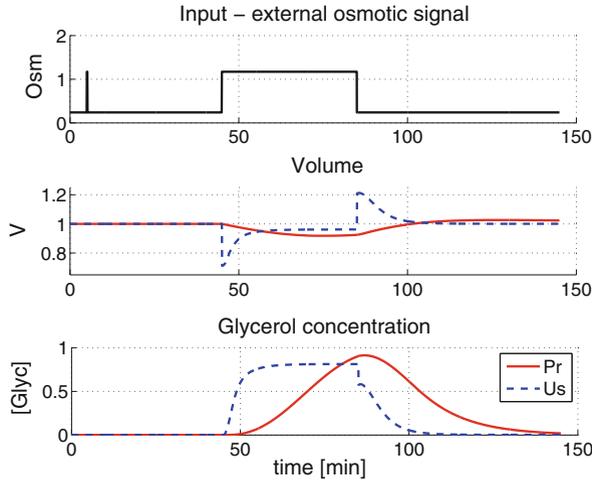


Fig. 4.19 Response to a signal with a short and long pulse duration, assuming $k_{HOG} = 0.25$. *Upper plot* external osmotic signal. *Second plot* the volume response for the proportional (Pr) controller and the ultrasensitive (Us) controller model. *Lower plot* the glycerol concentration for both the models

In conclusion, we show, using a previously developed proportional control model of the osmoeregulation incorporating two main and distinct branches (HOG and channel branches), that ultrasensitivity in the HOG branch allows better overall performance. We find that the primary effect of ultrasensitivity in the HOG branch is an increase in the response speed of the system and consequently in its adaptation precision. In addition to this, however, we find that ultrasensitivity provides also a non-trivial flexibility to the system parameters. By increasing the speed of overall system responses, ultrasensitivity in the HOG branch allows sensitivity to be increased in the other branch of the system. In the absence of ultrasensitivity, fast (i.e. highly sensitive) regulation of the glycerol exchange branch limits the cell's adaptability through the HOG branch (i.e. glycerol production). With ultrasensitivity in glycerol production, the other system parameters can be increased or varied more freely, without compromising performance. Moreover, by increasing the gain of the HOG branch, the system with a proportional HOG controller is able to improve the performance in terms of adaptation, but there is a presence of overshoot in the system response, whereas ultrasensitivity in the HOG branch allows to avoid this phenomena (we do not consider the overshoot to compute the performance). Note that for large values of the error ($e > 1$), a proportional branch may have a higher gain than an ultrasensitive one and, if $K > 1$, the gain of the proportional controller will always be higher, but this is not the case here because the error never goes above 1, given the system parameters (the absolute maximum value of the error is $P_f(0)$). The ultrasensitive response in the HOG branch also allows tuning of the overall system response towards certain signal regimes. In other words, the control system can be tuned to filter out signals below a threshold and respond only when volume decreases cross

this threshold. Considering that glycerol production is potentially highly costly for the cell, this ability of the system could give an evolutionary advantage by allowing cells to ignore short lived or low doses of osmo-shock.

4.7 Conclusions

The results illustrated in this chapter demonstrate the power of applying engineering principles to the analysis of the osmoregulation system in yeast. Gennemark et al. [8] proposed a simple model that describes the essential physics and biology of osmoregulation. This model has been abstracted from another more detailed model, developed by [15], by focusing on fewer components which allow the reproduction of the main dynamics of the system: the cell controls the biophysical system (in particular in terms of volume and turgor pressure) by using two proportional controllers, which act in parallel and regulate the glycerol production and the glycerol outflow (see Fig. 4.2). This simple model captures the main dynamical features of the osmoadaptive response by predicting the behaviour of different strains (wild type and modified) with different inputs and confirming the existence of two mechanisms of control (see Sect. 4.3). Note, however, that in general the volume adapts while the turgor pressure does not, because the model parameters are estimated using only glycerol concentration measurements which are not sufficient for complete recovery of both volume and turgor pressure. Therefore the model does not show robust adaptation, since the adaptation requires a careful tuning of the system parameters.

The group of van Oudenaarden, using frequency domain analysis, identified a minimal model represented by a LTI system with only two dynamics variables (see Sect. 4.4). Then, they estimated the biological quantities corresponding to the two relevant variables of the LTI model and, using these results, deduced the network diagram of Fig. 4.10. Using biological measurements and engineering principles, they showed that the robust perfect adaptation of Hog1 nuclear enrichment and cell volume (as turgor pressure) results from one integrating mechanism that requires Hog1 kinase activity and regulates the glycerol synthesis (see Sect. 4.5).

The models of Figs. 4.2 and 4.10 seem similar at a “formal” level but they are quite different from the system theoretical point of view. The model of the group of Van Oudenaarden is inferred by employing the measurements (the output of the model) and contains one branch of control modelled with exactly one integrator. Instead, in Gennemark’s model, the Hog protein cannot be observed (it is not a variable of the model and the Hog controller does not have a direct biological correspondence) and the two branches of control are modelled using simple proportional controllers. The model could be modified by adding measurable variables, for example Hog1, but this would obviously increase the complexity of the model.

Interesting additional results were recently presented in [20], where the authors investigated which network topologies in a generic signalling network are capable of robust adaptation. In particular, they used a network of three nodes as a minimal framework, where there is a first node that receives the input, a second that transmits

the output and a third that can play diverse regulatory roles. They found that all the networks containing one of the following two motifs achieve adaptation: negative feedback loop with a buffering node and incoherent feedforward loop with a “proportioner” node.

Despite the many striking insights that have been produced into the yeast osmoregulation system by the above analyses, it is still not clear how the evolution of biological control systems of this type can result in integral feedback, and in our recent work we investigated a heretofore largely unexplored alternative control system which also appears to be able to achieve perfect adaptation. In particular, we extended the proportional control model developed by [8] with the implementation of ultrasensitivity and found that a proportional controller implementing ultrasensitivity allows more precise and faster adaptation of cell volume following an osmo-shock. Further, the ultrasensitive controller can be tuned as a filter, where the proportional controller could not, and thereby allows responses to signals only above a certain threshold (see Sect. 4.6). These results provide new insights on the potential role of gain in biological systems and should be of interest to synthetic biologists attempting to design robust biomolecular control systems.

References

1. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) *Molecular biology of the cell*. Garland Publishing, Inc., New York
2. Albertyn J, Hohmann S, Thevelein JM, Prior BA (1994) Gpd1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. *Mol Cell Biol* 14(6):4135–4144
3. Blossy R, Bodart JF, Devys A, Goudon T, Lafitte P (2012) Signal propagation of the MAPK cascade in *Xenopus* oocytes: role of bistability and ultrasensitivity for a mixed problem. *J Math Biol* 64(1–2):1–39
4. Buchler NE, Cross FR (2009) Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol Syst Biol* 5:272
5. Cosentino C, Bates DG (2011) *Feedback control in systems biology*. CRC Press (Taylor & Francis), Boca Raton
6. El-Samad H, Goff JP, Khammash M (2002) Calcium homeostasis and parturient hypocalcemia: an integral feedback perspective. *J Theor Biol* 214(1):17–29
7. Fleming PJ, Purshouse RC (2002) *Evolutionary algorithms in control systems engineering: a survey*. *Control Eng Pract* 10:1223–1241
8. Gennemark P, Nordlander B, Hohmann S, Wedelin D (2006) A simple mathematical model of adaptation to high osmolarity in yeast. *Silico Biol* 6(3):193–214
9. Gervais P, Beney L (2001) Osmotic mass transfer in the yeast *saccharomyces cerevisiae*. *Cell Mol Biol (Noisy-le-grand)* 47(5):831–839
10. Goldberg DE (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Boston
11. Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. *PNAS* 78(11):6840–6844
12. Ho SN (2006) Intracellular water homeostasis and the mammalian cellular osmotic stress response. *J Cell Physiol* 206(1):9–15. doi:10.1002/jcp.20445, <http://dx.doi.org/10.1002/jcp.20445>

13. Huang CY, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. *PNAS* 93(19):10078–10083
14. Ingalls BP, Yi T, Iglesias P (2006) Using control theory to study biology. In: Knabe JF, Nehaniv CL, Schilstra MJ (eds) *System modeling in cellular biology*. MIT Press, Cambridge, pp 243–267
15. Klipp E, Nordlander B, Krger R, Gennemark P, Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nat Biotechnol* 23(8):975–982. doi:[10.1038/nbt1114](https://doi.org/10.1038/nbt1114), <http://dx.doi.org/10.1038/nbt1114>
16. Klitz D, Burg M (1998) Evolution of osmotic stress signaling via map kinase cascades. *J Exp Biol* 201(Pt 22):3015–3021
17. Levin RL, Ushiyama M, Cravalho EG (1979) Water permeability of yeast cells at sub-zero temperatures. *J Membr Biol* 46(2):91–124
18. Lobo FG, Goldberg D (1996) Decision making in a hybrid genetic algorithm. *IlligAL Report No 96009*
19. Luyten K, Albertyn J, Skibbe WF, Prior BA, Ramos J, Thevelein JM, Hohmann S (1995) Fps1, a yeast member of the mip family of channel proteins, is a facilitator for glycerol uptake and efflux and is inactive under osmotic stress. *EMBO J* 14(7):1360–1371
20. Ma W, Trusina A, El-Samad H, Lim WA, Tang C (2009) Defining network topologies that can achieve biochemical adaptation. *Cell* 138(4):760–773. doi:[10.1016/j.cell.2009.06.013](https://doi.org/10.1016/j.cell.2009.06.013), <http://dx.doi.org/10.1016/j.cell.2009.06.013>
21. Markevich NI, Hoek JB, Kholodenko BN (2004) Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol* 164(3):353–359
22. MATLAB (1990) *Optimization Toolbox User's Guide*. Natick, MA
23. MATLAB (2004) *Global Optimization Toolbox User's Guide*. Natick, MA
24. Mettetal JT, Muzzey D, Gómez-Uribe C, van Oudenaarden A (2008) The frequency dependence of osmo-adaptation in *saccharomyces cerevisiae*. *Science* 319(5862):482–484. doi:[10.1126/science.1151582](https://doi.org/10.1126/science.1151582), <http://dx.doi.org/10.1126/science.1151582>
25. Montefusco F, Steinacher A, Akman OE, Bates DG, Soyer OS (2012) On the role of ultrasensitivity in biomolecular control systems. In: *Proceedings of the 51st IEEE CDC conference, Maui, Hawaii*.
26. Muzzey D, Gómez-Uribe CA, Mettetal JT, van Oudenaarden A (2009) A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell* 138(1):160–171
27. Ni XY, Drenstig T, Ruoff P (2009) The control of the controller: molecular mechanisms for robust perfect adaptation and temperature compensation. *Biophys J* 97(5):1244–1253
28. Profit M, Struhl K (2004) Map kinase-mediated stress relief that precedes and regulates the timing of transcriptional induction. *Cell* 118(3):351–361. doi:[10.1016/j.cell.2004.07.016](https://doi.org/10.1016/j.cell.2004.07.016), <http://dx.doi.org/10.1016/j.cell.2004.07.016>
29. Reed RH, Chudek JA, Foster R, Gadd GM (1987) Osmotic significance of glycerol accumulation in exponentially growing yeasts. *Appl Environ Microbiol* 53(9):2119–2123
30. Rep M, Krantz M, Thevelein JM, Hohmann S (2000) The transcriptional response of *saccharomyces cerevisiae* to osmotic shock. *hot1p* and *msn2p/msn4p* are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes. *J Biol Chem* 275(12):8290–8300
31. Schaber J, Adrover MA, Eriksson E, Pelet S, Petelenz-Kurdziel E, Klein D, Posas F, Goksör M, Peter M, Hohmann S, Klipp E (2010) Biophysical properties of *Saccharomyces cerevisiae* and their relationship with HOG pathway activation. *Eur Biophys J* 39(11):1547–1556
32. Skogestad S, Postlethwaite I (1996) *Multivariable feedback control: analysis and design*. Wiley, Chichester
33. Smith AE, Zhang Z, Thomas CR (2000) Wall materials properties of yeast cells: part 1. Cell measurements and compression experiments. *Chem Eng Sci* 55:2031–2041
34. Sontag E (2003) Adaptation and regulation with signal detection implies internal model. *Syst Contr Lett* 50:119–126
35. Sunder S, Singh AJ, Gill S, Singh B (1996) Regulation of intracellular level of na^+ , k^+ and glycerol in *saccharomyces cerevisiae* under osmotic stress. *Mol Cell Biochem* 158(2):121–124

36. Tams MJ, Luyten K, Sutherland FC, Hernandez A, Albertyn J, Valadi H, Li H, Prior BA, Kilian SG, Ramos J, Gustafsson L, Thevelein JM, Hohmann S (1999) Fps1p controls the accumulation and release of the compatible solute glycerol in yeast osmoregulation. *Mol Microbiol* 31(4):1087–1104
37. Thorsen M, Di Y, Tngemo C, Morillas M, Ahmadpour D, Van der Does C, Wagner A, Johansson E, Boman J, Posas F, Wysocki R, Tams MJ (2006) The mapk hog1p modulates fps1p-dependent arsenite uptake and tolerance in yeast. *Mol Biol Cell* 17(10):4400–4410. doi:10.1091/mbc.E06-04-0315, <http://dx.doi.org/10.1091/mbc.E06-04-0315>
38. Westfall PJ, Patterson JC, Chen RE, Thorner J (2008) Stress resistance and signal fidelity independent of nuclear mapk function. *Proc Natl Acad Sci U S A* 105(34):12212–12217. doi:10.1073/pnas.0805797105, <http://dx.doi.org/10.1073/pnas.0805797105>
39. Yi TM, Huang Y, Simon MI, Doyle J (2000) Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *PNAS* 97(9):4649–4653
40. Zi Z, Liebermeister W, Klipp E (2010) A quantitative study of the Hog1 MAPK response to fluctuating osmotic stress in *Saccharomyces cerevisiae*. *PLoS One* 5(3):e9522

Chapter 5

Analysis of Synchronizing Biochemical Networks via Incremental Dissipativity

Abdullah Hamadeh, Jorge Gonçalves and Guy-Bart Stan

Keywords Dissipativity · Incremental dissipativity · Passive · Passivity · Incremental signal · Zero-state detectability · Storage function · Supply rate · Incremental supply rate · Incrementally passive · Incrementally output feedback passive (iOFP)

5.1 Introduction

Synchronization, defined in a broad sense, is the phenomenon in which communicating agents coordinate outputs. The abundance of examples of this process in nature and engineering has led to its becoming an active sub-area of research in networks theory, as evidenced by the multitude of publications on the subject [4].

The aim of this chapter is to re-visit, generalize and extend earlier work in [11, 12, 23], on the synchronization of interconnected control systems, in which a dissipativity approach is employed to arrive at the coupling conditions necessary to ensure the convergence of nodal outputs to a common value. The motivation for the development of these tools comes from a systems biology example, namely the modeling and analysis of synchrony in the neuronal networks that control circadian rhythms in the

A. Hamadeh
Department of Mechanical Engineering,
Massachusetts Institute of Technology, Cambridge, MA, USA
e-mail: ahamadeh@mit.edu

J. Gonçalves
Department of Engineering, University of Cambridge, Cambridge, UK
e-mail: jmg77@cam.ac.uk

G.-B. Stan (✉)
Department of Bioengineering Imperial College, Centre for Synthetic Biology
and Innovation, London, UK
e-mail: g.stan@imperial.ac.uk

mammalian hypothalamus [10]. In this context, the ‘network’ is composed of a set of cells (the network nodes). Each cell communicates with its neighbors by sending a biochemical output signal that reflects its internal state, and by taking, as an external input, a measure of outputs of its neighboring cells.

An assumption we make is that the internal dynamics of each cell can be modularized into a set of interconnected compartments. In [11, 12, 23] these compartments were connected in a ring that is structurally similar to the Goodwin oscillator. Such a structure represents a simple yet common genetic circuit whereby DNA is transcribed to mRNA, which is then translated into a protein, which feeds back to inhibit transcription. As shown in [19], in many cases the dynamics of such modules can be characterized, from an input-output perspective, as being passive passive in the sense of [25]. A biochemical circuit that is formed by the interconnection of such modules thus lends itself to stability analysis by the dissipativity theory tools developed in [2, 20, 24]. A particular advantage of this approach with regards to biological systems is that the internal dynamics of each module need not be known precisely.

To analyze synchronization in networks of such interconnected cells, references [12, 23] regard synchronization as the stability of signals that represent the differences in output between two nodes. In parallel with the use of passivity theory [2] to analyze the stability of circuits composed of the interconnection of passive passive subsystems, the work in [12, 23] employs the concept of *incremental passivity*, first introduced in [21, 22], to study the synchronization of cells composed of incrementally passive passive subsystems. Given two identical copies of a system that has an input-state-output description, the system is said to be *incrementally passive* if it is passive passive with respect to the *difference* between its inputs, states and outputs (termed the *incremental signals* of the system). The class of network agents we will study in this chapter is such that the compartments of each individual node are subsystems that are individually incrementally output feedback passive passive [18]. We will use measures of their incremental passivity in order to quantify the degree of shortage of incremental passivity of each node with respect to its coupling inputs and outputs. Then, in analogy with the use of strong negative feedback for purposes of stabilizing output feedback passive passive systems, we will show that linear static coupling that is strongly connected can similarly be used to incrementally stabilize the network nodes, thus leading to asymptotic output synchrony and asymptotic state synchrony under a zero-state detectability assumption on the *differences* between the corresponding states and outputs of network nodes. Following [23], an alternative input–output approach was developed in [17] to analyze the synchrony of network agents that have structures more general than the cyclic nodes studied in [23]. An aim of this chapter is to show that the incrementally passifying role of coupling, which is analogous to the stabilizing role of feedback, can be used for the analysis of synchrony in networks of nodes such as those studied in [17].

With respect to other synchrony analysis tools in the literature such as contraction theory [6, 13, 15] or incremental input-to-state stability [1], the methodology we present here takes an input–output, modular approach. This results in a natural framework with which to analyze synchrony of agents composed of the interconnection of subsystems, and little knowledge of the subsystem dynamics is required.

To illustrate our results, we apply them towards the analysis of synchrony in networks of the repressilator genetic (synthetic) circuit [8, 9]. Genetic circuits can generally be posed in a modular form similar to that of the repressilator. For this reason we envision that the tools we present here will prove to be especially useful for the analysis of networks of such systems.

5.2 Synchronization and Incremental Dissipativity

As networked systems are generally connected through their inputs and outputs, it is natural to characterize them through their input–output properties to identify sufficient synchronization conditions. This chapter considers an incremental dissipativity characterization of the network nodes that will be termed incremental output-feedback passivity (iOFP). The following section will give a brief introduction to the concepts of incremental dissipativity, first introduced in [21, 22].

5.2.1 Incremental Dissipativity

Consider a system \mathcal{Y} represented by a state-space model of the form

$$\mathcal{Y} \begin{cases} \dot{x} = f(x, e), & x \in \mathbb{R}^r, e \in \mathbb{R} \\ y = g(x), & y \in \mathbb{R} \end{cases} \quad (5.1)$$

where $e(t)$, $y(t)$, and $x(t)$ denote its input, output and state respectively and the functions $f(x, e) : \mathbb{R}^r \times \mathbb{R} \rightarrow \mathbb{R}^r$ and $g(x) : \mathbb{R}^r \rightarrow \mathbb{R}$ are Lipschitz continuous. Let $x_a(t)$ and $x_b(t)$ be two solutions of \mathcal{Y} , with the corresponding input–output pairs $(e_a(t), y_a(t))$, and $(e_b(t), y_b(t))$. Denote by $\Delta x = x_a - x_b$, $\Delta e = e_a - e_b$, and $\Delta y = y_a - y_b$ the corresponding incremental variables. System (5.1) is incrementally dissipative if there exists a radially unbounded incremental storage function

$$S_\Delta : \mathbb{R}^r \rightarrow \mathbb{R}, S_\Delta(\Delta x) > 0 : \forall \Delta x \neq 0, S_\Delta(0) = 0, S_\Delta \in \mathcal{C}^1 \quad (5.2)$$

and an incremental supply rate $w(\Delta e, \Delta y)$ such that, if $S_\Delta(\Delta x)$ is at least once differentiable (i.e. $S_\Delta \in \mathcal{C}^1$)

$$\dot{S}_\Delta(\Delta x) \leq \mathcal{W}(\Delta e, \Delta y) \quad (5.3)$$

is satisfied for all time t and along any pair of trajectories $(x_a(t), x_b(t))$ (see [25] for a definition of dissipativity).

Definition 1 [21, 22] System \mathcal{Y} in (5.1) is said to be

- *incrementally passive* when it is incrementally dissipative with incremental supply rate $\mathcal{W}(\Delta e, \Delta y) = \Delta y \Delta e$.

- *incrementally output feedback passive* (iOFP $\left(\frac{1}{\gamma}\right)$) when it is incrementally dissipative with the incremental supply rate $\mathcal{W}(\Delta e, \Delta y) = -\frac{1}{\gamma}(\Delta y)^2 + \Delta y \Delta e$ with $\gamma \in (-\infty, \infty)$.
- *incrementally output strictly passive* (iOSP) when it is incrementally dissipative with the incremental supply rate $\mathcal{W}(\Delta e, \Delta y) = -\frac{1}{\gamma}(\Delta y)^2 + \Delta y \Delta e$ and $\gamma > 0$.

When $\gamma > 0$ the system possesses an *excess of incremental passivity* of $\frac{1}{\gamma}$. On the other hand, when $\gamma < 0$ the system possesses a *shortage of incremental passivity* and $-\frac{1}{\gamma}$ quantifies the minimum gain of proportional negative incremental output feedback required to make the system incrementally passive.

Definition 2 (*Incremental secant gain*) Following the concept of the ‘secant gain’ in [2, 19], the smallest $\gamma > 0$ such that the iOSP dissipation inequality in Definition 1 is satisfied will be termed the *incremental secant gain* of the system.

Remark 1 [21, 22] Passivity implies incremental passivity for linear systems, that is, if the quadratic storage function $S(x) = \frac{1}{2}x^*Px \geq 0$ satisfies the dissipation inequality $\dot{S} \leq yu$ then the incremental storage function $S_\Delta(\Delta x) = \frac{1}{2}(\Delta x)^*P\Delta x \geq 0$ satisfies the incremental dissipation inequality $\dot{S}_\Delta \leq \Delta y \Delta e$. Passivity also implies incremental passivity for a monotone increasing, static nonlinearity: if $\phi(\cdot)$ is monotone increasing, then $(e_a - e_b)(\phi(e_a) - \phi(e_b)) = \Delta e \Delta \phi(e) \geq 0$, $\forall \Delta e = e_a - e_b$, $\Delta \phi(e) = \phi(e_a) - \phi(e_b)$. Similarly, it is easy to show that for linear systems, output strict passivity implies incremental output strict passivity with the incremental secant gain equal to the secant gain.

5.2.2 Incremental Output-Feedback Passivity and Synchronization

Thus far we have seen that a system is incrementally dissipative if, given any two sets of initial conditions, input trajectories and corresponding outputs, the inequality (5.3) is satisfied. For this reason, the incremental dissipativity property, which is a property of each individual node, can be used as an analysis tool for an entire network composed of interconnected copies of such a node. The main result that will link incremental output feedback passivity of nodes of a network to (output) synchronization states that if each subsystem is iOFP and the coupling strength between nodes is large enough then all the nodes will asymptotically synchronize.

5.3 Notation

In the following sections, we consider networks composed of N coupled identical nodes, each composed of n interconnected SISO subsystems. As a general convention $j = 1, \dots, N$ will denote the index associated to a particular node of the network

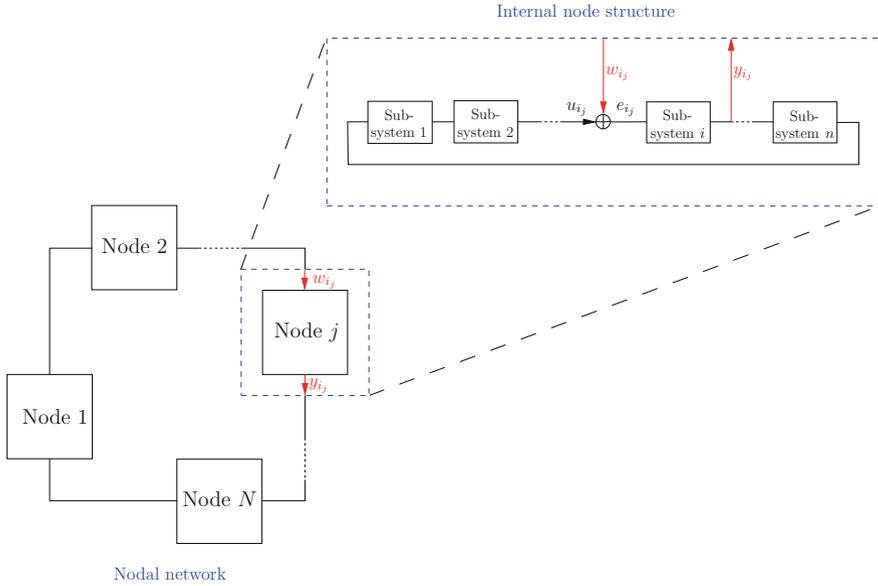


Fig. 5.1 Network of nodes and illustration of internal nodal structure. The subsystem interconnection structure can be arbitrary

whilst $i = 1, \dots, n$ will denote the index associated to a particular subsystem in a given node. The signals to be introduced in Assumption 1 below carry the following notations and are illustrated in Fig. 5.1.

- The subsystem i of node j has a state vector $x_{i_j} \in \mathbb{R}^s$, an input $e_{i_j} = u_{i_j} + w_{i_j}$ and output y_{i_j} , with $e_{i_j}, u_{i_j}, w_{i_j}, y_{i_j} \in \mathbb{R}$. The *internal input* u_{i_j} is a function of the outputs of different subsystems from the same node j . The *external input* w_{i_j} is a function of the outputs of corresponding subsystems i from the different nodes.
- The vectors of the states, inputs, internal inputs, external inputs and outputs of the j th node are respectively denoted by $\mathbf{x}_j, \mathbf{e}_j, \mathbf{u}_j, \mathbf{w}_j, \mathbf{y}_j$, where $\mathbf{x}_j = [x_{j_1}^* \cdots x_{j_n}^*]^*$ and $\mathbf{e}_j, \mathbf{u}_j, \mathbf{w}_j, \mathbf{y}_j$ are similarly defined.
- The vectors of the i th states, inputs, external inputs and outputs of each node are respectively denoted by X_i, E_i, U_i, W_i, Y_i , where $X_i = [x_{i_1}^* \cdots x_{i_n}^*]^* \in \mathbb{R}^N$ and E_i, U_i, W_i, Y_i are similarly defined.
- The vectors of all the states, inputs, internal inputs, external inputs and outputs are respectively denoted by X, E, U, W, Y , where $X = [X_1^* \cdots X_n^*]^*$, and the vectors E, U, W, Y are similarly defined.
- The incremental states, inputs, internal inputs, external inputs and outputs are respectively denoted by $\Delta x_{i_j,m}, \Delta e_{i_j,m}, \Delta u_{i_j,m}, \Delta w_{i_j,m}, \Delta y_{i_j,m}$, where $\Delta x_{i_j,m} \triangleq x_{i_j} - x_{i_m}$, and the signals $\Delta e_{i_j,m}, \Delta u_{i_j,m}, \Delta w_{i_j,m}, \Delta y_{i_j,m}$ are similarly defined.
- The vectors of incremental states, inputs, internal inputs, external inputs and outputs for two nodes j, m are respectively denoted by $\Delta \mathbf{x}_{j,m}, \Delta \mathbf{e}_{j,m}, \Delta \mathbf{u}_{j,m}, \Delta \mathbf{w}_{j,m},$

$\Delta \mathbf{y}_{j,m}$, where $\Delta \mathbf{x}_{j,m} \triangleq \mathbf{x}_j - \mathbf{x}_m$, and the signals $\Delta \mathbf{e}_{j,m}$, $\Delta \mathbf{u}_{j,m}$, $\Delta \mathbf{w}_{j,m}$, $\Delta \mathbf{y}_{j,m}$ are similarly defined.

- The vector in $\mathbb{R}^{N N_{ns}}$ of all incremental state vectors $\Delta x_{i,j,m}$ is denoted by X_Δ .

5.4 Characterization of Network Nodes and Their Dissipativity Properties

The following assumption gives a formal description of the networks and nodes that we shall consider in this chapter.

Assumption 1 Consider a network of N identical nodes. It is assumed that:

- Each node j is composed of n interconnected SISO subsystems of the form (5.1), and each such subsystem i has state vector, input and output x_{ij} , e_{ij} , y_{ij} respectively.
- Each subsystem i is iOFP $\left(\frac{1}{\gamma_i}\right)$ and therefore, for any two nodes j, m there is associated with each subsystem i a function $S_{i,j,m}(\Delta x_{i,j,m})$ that satisfies (5.2) and an incremental dissipation inequality of the form (5.3), with

$$\dot{S}_{i,j,m} \leq \mathcal{W}_i(\Delta e_{i,j,m}, \Delta y_{i,j,m}) = -\frac{1}{\gamma_i}(\Delta y_{i,j,m})^2 + \Delta e_{i,j,m} \Delta y_{i,j,m}, \quad \gamma_i \in \mathbb{R} \quad (5.4)$$

- The input to subsystem i of node j is given by $e_{ij} = u_{ij} + w_{ij}$ where u_{ij} are inputs from within the same node j and $u_{ij} = \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} y_{\ell j}$, $\alpha_{i,\ell} \in \mathbb{R}$ and where w_{ij} is an exogenous input.

Under Assumption 1, and by linearity we have

$$\Delta e_{i,j,m} = \Delta w_{i,j,m} + \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} \Delta y_{\ell j,m} \quad (5.5)$$

Combining this relation with the incremental dissipation inequality of the i th subsystem yields

$$\dot{S}_{i,j,m} \leq -\frac{1}{\gamma_i}(\Delta y_{i,j,m})^2 + \Delta w_{i,j,m} \Delta y_{i,j,m} + \Delta y_{i,j,m} \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} \Delta y_{\ell j,m} \quad (5.6)$$

Definition 3 (*Interconnection matrix*) For the vector of elements $\{\gamma\} \triangleq [\gamma_1 \cdots \gamma_n]^*$, define the *interconnection matrix* $A(\gamma)$ as

$$A(\gamma) \triangleq \begin{bmatrix} -\frac{1}{\gamma_1} & \alpha_{1,2} & \cdots & \alpha_{1,n-1} & \alpha_{1,n} \\ \alpha_{2,1} & -\frac{1}{\gamma_2} & \alpha_{2,3} & \cdots & \alpha_{2,n} \\ \alpha_{3,1} & \alpha_{3,2} & -\frac{1}{\gamma_3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \alpha_{4,n} \\ \alpha_{n,1} & \cdots & \alpha_{n,2} & \alpha_{n,n-1} & -\frac{1}{\gamma_n} \end{bmatrix}$$

which is such that

$$\sum_{i=1}^n \left[-\frac{1}{\gamma_i} (\Delta y_{i,j,m})^2 + \Delta y_{i,j,m} \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} \Delta y_{\ell,j,m} \right] = \frac{1}{2} (\Delta \mathbf{y}_{j,m})^* (A(\gamma)^* + A(\gamma)) \Delta \mathbf{y}_{j,m}$$

Lemma 1 For an interconnection matrix $A(\gamma)$ as defined in Definition 3, there exist diagonal matrices $D > 0$, $D \in \mathbb{R}^{n \times n}$, $D = \text{diag}\{d_1, \dots, d_n\}$ and $K \geq 0$, $K \in \mathbb{R}^{n \times n}$, $K = \text{diag}\{k_1, \dots, k_n\}$ so that for all diagonal matrices $K' = \text{diag}\{k'_1, \dots, k'_n\} \in \mathbb{R}^{n \times n}$ which satisfy $K' \geq K$, there exists $\epsilon_{D,K} > 0$ which is such that

$$\frac{1}{2} (A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) \leq -\epsilon_{D,K} I_n, \quad \tilde{\gamma} = \{\tilde{\gamma}_i\}, \quad \tilde{\gamma}_i \triangleq \frac{\gamma_i}{1 + k'_i \gamma_i} \quad (5.7)$$

Proof Since $\tilde{\gamma}_i = \frac{\gamma_i}{1 + k'_i \gamma_i}$ and $\tilde{\gamma} = \{\tilde{\gamma}_i\}$, it follows that $A(\tilde{\gamma}) = A(\gamma) - K'$. To prove the existence of a pair of matrices $D > 0$, $K \geq 0$, that satisfy (5.7), note that if $D = I_n$ then there always exists a set of elements k'_i which are individually sufficiently large in magnitude to make the diagonal elements of $A(\tilde{\gamma})$ negative and also sufficiently large in magnitude to ensure that $A(\tilde{\gamma})^* D + DA(\tilde{\gamma}) = A(\tilde{\gamma})^* + A(\tilde{\gamma}) < 0$ by diagonal dominance.¹ Taking any such pair D , K which are such that

$$\frac{1}{2} ((A(\gamma) - K)^* D + D(A(\gamma) - K)) \leq -\epsilon_{D,K} I_n$$

is satisfied with $\epsilon_{D,K} > 0$, then since $D(K' - K) \geq 0$, it necessarily follows that for any $K' \geq K$

$$\frac{1}{2} (A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) \leq -\epsilon_{D,K} I_n$$

This completes the proof.

Lemma 1 proves that the diagonal stability of interconnection matrices can always be achieved by making their diagonal elements large in magnitude and negative in size. The following theorem makes use of this result to quantify the shortage of

¹ Note that the choice of matrix $D = I_n$ is not unique and in most cases a matrix D can be constructed to reduce the sizes of elements k_i required to achieve negative definiteness of $A(\tilde{\gamma})^* D + DA(\tilde{\gamma})$.

passivity of the network nodes from the degree of passivity of the individual nodal subsystems.

Theorem 1 *For a network of identical nodes that satisfy Assumption 1, there exist diagonal matrices $D > 0$, $D \in \mathbb{R}^{n \times n}$, $D = \text{diag}\{d_1, \dots, d_n\}$ and $K \geq 0$, $K \in \mathbb{R}^{n \times n}$, $K = \text{diag}\{k_1, \dots, k_n\}$ so that for the storage function $S_{j,m} = \sum_{i=1}^n d_i S_{i,j,m}$ each network node is iOFP($-K$) and satisfies the incremental dissipation inequality*

$$\dot{S}_{j,m} \leq -\epsilon_{D,K} (\Delta \mathbf{y}_{j,m})^* (\Delta \mathbf{y}_{j,m}) + (\Delta \mathbf{y}_{j,m})^* D (K \Delta \mathbf{y}_{j,m} + \Delta \mathbf{w}_{j,m}) \quad (5.8)$$

where $\epsilon_{D,K} > 0$ is such that

$$\frac{1}{2} (A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) \leq -\epsilon_{D,K} I_n, \quad \tilde{\gamma} = \{\tilde{\gamma}_i\}, \quad \tilde{\gamma}_i \triangleq \frac{\gamma_i}{1 + k_i \gamma_i} \quad (5.9)$$

and where $A(\cdot)$ is as defined in Definition 3.

Proof The first step of the proof is to add and subtract to each dissipation inequality (5.6) the term $k_i (\Delta y_{i,j,m})^2$, with $k_i \geq 0$, to obtain

$$\dot{S}_{i,j,m} \leq -\frac{1}{\tilde{\gamma}_i} (\Delta y_{i,j,m})^2 + k_i (\Delta y_{i,j,m})^2 + \Delta y_{i,j,m} \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} \Delta y_{\ell,j,m} + \Delta w_{i,j,m} \Delta y_{i,j,m} \quad (5.10)$$

where

$$\tilde{\gamma}_i = \frac{\gamma_i}{1 + k_i \gamma_i}$$

Now, defining the incremental storage function $S_{j,m}(\Delta \mathbf{x}_{j,m})$ as the linear sum $S_{j,m} = \sum_{i=1}^n d_i S_{i,j,m}$, its time derivative becomes

$$\begin{aligned} \dot{S}_{j,m} &\leq \sum_{i=1}^n d_i \left(-\frac{1}{\tilde{\gamma}_i} (\Delta y_{i,j,m})^2 + \Delta y_{i,j,m} \sum_{\substack{\ell=1 \\ \ell \neq i}}^n \alpha_{i,\ell} \Delta y_{\ell,j,m} + \Delta w_{i,j,m} \Delta y_{i,j,m} + k_i (\Delta y_{i,j,m})^2 \right) \\ &= \frac{1}{2} (\Delta \mathbf{y}_{j,m})^* (A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) (\Delta \mathbf{y}_{j,m}) + \sum_{i=1}^n d_i \left(\Delta w_{i,j,m} \Delta y_{i,j,m} + k_i (\Delta y_{i,j,m})^2 \right) \\ &= \frac{1}{2} (\Delta \mathbf{y}_{j,m})^* (A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) (\Delta \mathbf{y}_{j,m}) + (\Delta \mathbf{y}_{j,m})^* D (K \Delta \mathbf{y}_{j,m} + \Delta \mathbf{w}_{j,m}) \end{aligned}$$

where $d_i > 0$, $\forall i$, $D = \text{diag}(d_1, \dots, d_n)$, and $A(\tilde{\gamma})$ is as defined in Definition 3 but with the vector of elements $\tilde{\gamma} = \{\tilde{\gamma}_i\}$. The key step at this point is that, following the definition of $\tilde{\gamma}_i$, increasing k_i sufficiently can make the quantity $\tilde{\gamma}_i$ positive if γ_i is negative. In this way the subsystem i , which satisfies the incremental dissipation inequality (5.10) becomes iOFP($-k_i$).

Moreover, if each k_i is made sufficiently large, the values of $\tilde{\gamma}_i$ can be made small enough so that a diagonal matrix D can be constructed which makes

$$\frac{1}{2} (A(\tilde{\gamma})^* D + D A(\tilde{\gamma})) \leq -\epsilon_{D,K} I_n \quad (5.11)$$

for some $\epsilon_{D,K} > 0$, as shown in Lemma 1. Therefore if quantities k_i and d_i are chosen so that (5.11) is satisfied then

$$\dot{\mathbf{S}}_{j,m} \leq -\epsilon_{D,K} (\Delta \mathbf{y}_{j,m})^* (\Delta \mathbf{y}_{j,m}) + (\Delta \mathbf{y}_{j,m})^* D (K \Delta \mathbf{y}_{j,m} + \Delta \mathbf{w}_{j,m})$$

and the node is therefore iOFP($-K$) since a negative feedback of $\Delta \mathbf{w}_{j,m} = -K \Delta \mathbf{y}_{j,m}$ would render the node iOSP. This completes the proof.

5.4.1 Network Coupling Topology

Now consider a network composed of N identical nodes, where each node is iOFP ($-K$) as shown in Theorem 1. Assume that the nodes are connected using their i th subsystems through a weighted directed graph \mathcal{G}^i (the graph can be different for each i) and assume that the coupling structure is restricted to a linear, static input–output interconnection, so that the i th subsystem on the j th node is coupled to the i th subsystem on other nodes in the network through its inputs w_{ij} and outputs y_{ij} using the Laplacian coupling matrix $\Gamma_i \in \mathbb{R}^{N \times N}$, so that $W_i = -\Gamma_i Y_i$. The graph $\mathcal{G}^i = \{\mathcal{A}^i, \mathcal{D}^i\}$ has the following definitions.

Definition 4 (*Weighted Adjacency Matrix*) A weighted adjacency matrix $\mathcal{A}^i = \{\rho_{i,j,l}\}$, $j, l = 1, \dots, N$, $\mathcal{A}^i \in \mathbb{R}^{N \times N}$, is a positive matrix where $\rho_{i,j,l}$ represents the weight of the edge from node l to node j . It is assumed that the graph is simple, i.e. $\rho_{i,j,l} \geq 0$, $\forall j \neq l$ and $\rho_{i,j,j} = 0$, $\forall j, l$.

Definition 5 (*Degree Matrix*) The degree matrix \mathcal{D}^i associated with the adjacency matrix \mathcal{A}^i is a diagonal matrix $\mathcal{D}^i = \text{diag}\{\delta_j^i\}$, $j = 1, \dots, N$, $\mathcal{D}^i \in \mathbb{R}^{N \times N}$ with $\delta_j^i = \sum_{\substack{l=1 \\ l \neq j}}^N \rho_{i,j,l}$.

Definition 6 (*Laplacian Matrix*) The weighted Laplacian matrix $\Gamma_i \in \mathbb{R}^{N \times N}$ associated with the adjacency matrix \mathcal{A}^i is defined as $\Gamma_i = \mathcal{D}^i - \mathcal{A}^i = \{\Gamma_{i,j,l}\}$ for $j, l = 1, \dots, N$ and $\Gamma_{i,j,j} = \delta_j^i$, $\forall j = 1, \dots, N$ and $\Gamma_{i,j,l} = -\rho_{i,j,l}$, $\forall j \neq l$. The matrix $\tilde{\Gamma}$ is defined as

$$\tilde{\Gamma} \triangleq \text{diag}\{\Gamma_1, \dots, \Gamma_n\}$$

The interconnection rule $W_i = -\Gamma_i Y_i$ then corresponds to the linear consensus protocol $w_{ij} = -\sum_{l=1}^N \rho_{i,j,l} (y_{ij} - y_{il})$ (see [14]). The following assumptions are made on Γ_i :

- (A1) $\text{rank}(\Gamma_i) = N - 1$
- (A2) $\Gamma_i + \Gamma_i^T \geq 0$
- (A3) $\Gamma_i \mathbf{1}_N = \Gamma_i^T \mathbf{1}_N = \mathbf{0}_N$

The conditions (A1)–(A3) characterize the coupling structure we consider here as *diffusive coupling* [16]. Assumption (A1) holds provided that the graph is strongly connected (see [14]). Assumption (A3) holds if the graph is balanced, i.e. if $\mathcal{A}^i \mathbf{1}_N = \mathcal{A}^{i*} \mathbf{1}_N$ (see [5]). Furthermore, this latter property implies (A2) (see [5], which uses Gershgorin’s disk theorem to prove this fact). Note that these assumptions do not imply that Γ_i is symmetric which would be equivalent to assuming an undirected graph. We denote by λ_{k_i} the k th eigenvalue of the symmetric part of the Laplacian Γ_i , which is given by $\frac{1}{2}(\Gamma_i + \Gamma_i^*)$.

The eigenvalues λ_{k_i} are such that $\lambda_{1_i} < \lambda_{2_i} \leq \dots \leq \lambda_{N_i}$. From (A2) it follows that $\lambda_{k_i} \geq 0$ whilst from (A1) $\lambda_{1_i} = 0$. From (A3) $\lambda_{1_i} = 0$ corresponds to the eigenvector $\mathbf{1}_N$. The quantity λ_{2_i} has a special significance in graph theory and is known as the algebraic connectivity. As will be shown in Theorem 2 this quantity is a measure of the coupling strength of the network Laplacian Γ_i .

To compare each nodal output with its average over all the N nodes outputs, the projector matrix $\Pi \in \mathbb{R}^{N \times N}$, which is first defined in [21, 22]

$$\Pi \triangleq I_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^* \quad (5.12)$$

is employed. This projector measures the instantaneous difference between a signal and its average over all nodes in the network, e.g. the j th element of $\Pi Y_1(t)$ measures the difference between output $y_{1_j}(t)$, $j = 1, \dots, N$ and the average output $\frac{1}{N} \sum_{j=1}^N y_{1_j}(t)$. Note that the projector has the following properties from [21, 22]

- $\Pi^* \Pi = \Pi$
- $\Pi = \Pi^*$
- $\Pi \mathbf{1}_N = \mathbf{0}_N$

We also define the matrix

$$\tilde{\Pi}_r \triangleq I_r \otimes \Pi \quad (5.13)$$

which will be used to measure consensus in the concatenated signal vectors (for example the concatenated output vector Y).

5.4.2 Main Result on Network Synchronization

Because the nodes are identical, the incremental storage function $\mathbf{S}_{j,m}$ is such that, given any two sets of initial conditions, inputs, states and outputs for any two nodes $j, m \in \{1, \dots, N\}$, their corresponding trajectories satisfy an incremental dissipation inequality of the form (5.8). Due to Assumption 1, the incremental storage function

$\mathbf{S}_{j,m}$ has the properties $\mathbf{S}_{j,m}(\Delta \mathbf{x}_{j,m}) > 0 \forall \Delta \mathbf{x}_{j,m} \neq \mathbf{0}_n$, $\mathbf{S}_{j,m}(\mathbf{0}_n) = 0$. As we have seen in Theorem 1, Eq. (5.8) can be arrived at by a suitable choice of matrices D, K .

In Theorem 2 below, the following property will be used to deduce state synchronization from output synchronization.

Definition 7 (*Incremental zero-state detectability*) A system of the form (5.1) is *incrementally zero-state detectable* if, $\Delta u(t) = 0$ and $\Delta y(t) = 0, \forall t$, implies $\lim_{t \rightarrow \infty} \Delta x = \mathbf{0}_r$.

In the following theorem, the result on global asymptotic state synchronization of network nodes is given.

Theorem 2 (*Asymptotic State Synchronization*) Consider a network of N identical nodes satisfying Assumption 1, linearly coupled through the interconnection matrices Γ_i so that $W_i = -\Gamma_i Y_i$ where matrices Γ_i satisfy the assumptions (A1), (A2), and (A3). Assume that each node is incrementally zero-state detectable as in Definition 7 and is *iOFP*($-K$) as shown in Theorem 1, so that for every pair of nodes $j, m \in \{1, \dots, N\}$ there exists a radially unbounded incremental storage function $\mathbf{S}_{j,m}$ satisfying (5.8). Assume also that the network satisfies the strong coupling assumption $L \geq K$ where $L = \text{diag}\{\lambda_{2_1}, \dots, \lambda_{2_n}\}$. Then, each bounded network solution that exists for all $t \geq 0$ is such that $\forall i = 1, \dots, n, \forall j, l = 1, \dots, N$: $\lim_{t \rightarrow +\infty} (x_{ij}(t) - x_{il}(t)) = 0$ (global asymptotic synchronization). In addition to global asymptotic synchronization, any bounded network solution is such that the state solution of each node converges to the omega-limit set of the isolated node.

Proof Summing the storage functions $\mathbf{S}_{j,m}$ given in (5.8) for all node pairs j, m and then scaling by $\frac{1}{2N}$ gives the incremental storage function $S(X_\Delta) = \frac{1}{2N} \sum_{j=1}^N \sum_{m=1}^N \mathbf{S}_{j,m}$ for the network. From (5.8), S obeys the dissipation inequality

$$\dot{S} \leq -\epsilon_{D,K} (\tilde{\Pi}_n Y)^* (\tilde{\Pi}_n Y) + (\tilde{\Pi}_n Y)^* (D \otimes I_N)^* ((K \otimes I_N) \tilde{\Pi}_n Y + \tilde{\Pi}_n W) \quad (5.14)$$

Using (A3) and the relation $W_i = -\Gamma_i Y_i$, we have $\Pi W_i = -\Pi \Gamma_i Y_i = -\Gamma_i \Pi Y_i$ and therefore $\tilde{\Pi}_n W = -\tilde{\Gamma} \tilde{\Pi}_n Y$, where $\tilde{\Gamma} = \text{diag}\{\Gamma_1, \dots, \Gamma_n\}$. From this, (5.14) becomes

$$\dot{S} \leq -\epsilon_{D,K} (\tilde{\Pi}_n Y)^* (\tilde{\Pi}_n Y) + (\tilde{\Pi}_n Y)^* (D \otimes I_N)^* ((K \otimes I_N) \tilde{\Pi}_n Y - \tilde{\Gamma} \tilde{\Pi}_n Y) \quad (5.15)$$

From (A1)–(A3), $\Pi Y_i = Y_i - (\frac{1}{N} \mathbf{1}_N^* Y_i) \mathbf{1}_N = 0$ iff $Y_i \in \ker(\Gamma_i)$. Since $\ker(\Gamma_i)$ is of dimension one, it follows that

$$(\Pi Y_i)^* \Gamma_i \Pi Y_i \geq \lambda_{2_i} (\Pi Y_i)^* \Pi Y_i \quad (5.16)$$

Letting $L = \text{diag}\{\lambda_{2_1}, \dots, \lambda_{2_n}\}$ and substituting (5.16) in (5.15) yields

$$\dot{S} \leq -\epsilon_{D,K} (\tilde{\Pi}_n Y)^* (\tilde{\Pi}_n Y) + (\tilde{\Pi}_n Y)^* (D(K - L) \otimes I_N)^* \tilde{\Pi}_n Y$$

Noting that the diagonal matrix $D > 0$, then if, $\forall i \lambda_{2i}(\Gamma_i) > k_i$ (strong coupling) then $K - L < 0$. This gives the Lyapunov inequality

$$\dot{S} \leq -\epsilon_{D,K}(\tilde{\Pi}_n Y)^*(\tilde{\Pi}_n Y) \quad (5.17)$$

If this inequality holds, then letting $S_0 = S(X_\Delta(0))$, the initial value of the incremental storage function for the whole network, we note that, since $S \geq 0$ and $\dot{S} \leq 0$, the set $\mathcal{M} = \{X_\Delta | S(X_\Delta) \leq S_0\}$ is an invariant set with respect to the signal X_Δ . Note that \mathcal{M} also contains the incremental origin $X_\Delta = \mathbf{0}_{NNns}$ which is a strict minimum of $S(X_\Delta)$ since $S(X_\Delta) > 0$ for $X_\Delta \neq \mathbf{0}_{NNns}$ and $S(\mathbf{0}_{NNns}) = 0$.

From (5.17), and using the LaSalle invariance principle, the incremental signal X_Δ will converge to the largest invariant subset of $\{X_\Delta \in \mathbb{R}^{NNns} | \dot{S}(X_\Delta) = 0\}$ as $t \rightarrow \infty$. Due to (5.17), $\dot{S}(X_\Delta) = 0$ only if $\tilde{\Pi}_n Y = \mathbf{0}_{Nn}$. This implies asymptotic output synchronization since $\forall i$ and for any pair $j, m \in \{1, \dots, N\}$, $\lim_{t \rightarrow \infty} (y_{ij}(t) - y_{im}(t)) = 0$. Furthermore, from the incremental zero-state detectability assumption, the condition $\tilde{\Pi}_n Y = \mathbf{0}_{Nn}$ and the fact that there is no external input to the network means that $\lim_{t \rightarrow \infty} X_\Delta = \mathbf{0}_{NNns}$. This proves that each network solution that exists for all $t \geq 0$ is, regardless of initial conditions, such that $\forall i = 1, \dots, n, \forall j, m = 1, \dots, N: \lim_{t \rightarrow +\infty} (x_{ij}(t) - x_{im}(t)) = 0$.

Since $\Gamma \mathbf{1}_N = 0$, the effect of the coupling disappears when output synchrony is reached and each node in the network is then effectively isolated. Therefore, in addition to global asymptotic state synchronization, for any bounded network solution, the solution of each node converges to the omega-limit set of an isolated node. This completes the proof.

The preceding discussion has presented two main ideas, formalized in Theorems 1 and 2. In the Theorem 1 it was shown that *any* node of the form specified in (1) (that is, any node composed of an interconnection of subsystems that obey an iOFP property) is iOFP($-K$), where $K \geq 0$ is a diagonal matrix. In Theorem 2, it is then shown that an iOFP($-K$) node can be made to synchronize by sufficiently strong coupling, where the coupling strength is quantified by the eigenvalues of the diagonal matrix L (which are the algebraic connectivities of Laplacians Γ_i). In effect, the coupling acts as an incrementally stabilizing negative feedback that compensates for any shortage of incremental passivity by the nodes.

Lemma 1 and Theorem 1 show that the network nodes are iOFP($-K$), $K = \text{diag}\{k_1, \dots, k_n\}$, by demonstrating that, associated with any given K is another matrix $D > 0$ such that $A(\tilde{\gamma})^* D + DA(\tilde{\gamma}) < 0$ where

$$\tilde{\gamma} = \{\tilde{\gamma}_i\}, \quad \tilde{\gamma}_i = \frac{\gamma_i}{1 + k_i \gamma_i} \quad (5.18)$$

If $D = I_n$ the diagonal elements of $A(\tilde{\gamma})^* + A(\tilde{\gamma})$ equal $-\frac{2}{\tilde{\gamma}_i}$. By sufficiently increasing each k_i , the diagonal elements of $A(\tilde{\gamma})^* + A(\tilde{\gamma})$ can therefore be made negative and large enough for $A(\tilde{\gamma})^* + A(\tilde{\gamma})$ to become diagonally dominant and

hence negative definite. However, the size of the eigenvalues of K required to achieve diagonal dominance is usually conservatively high.

Moreover, it is important to note that, for a given interconnection matrix $A(\gamma)$ the diagonal matrices D and K that make each node iOFP ($-K$) are not unique, and there are, in fact, some choices that are ‘better’ than others in that some matrices D can be used to make $A^*(\tilde{\gamma})D + DA(\tilde{\gamma}) = (A(\gamma) - K)^*D + D(A(\gamma) - K) < 0$ using matrices K which are more sparse than others (by the sparsity of K we mean the number of zeros on its diagonals). The sparser the matrix K , the fewer the coupling connections that need to be made between the nodes since for every positive k_i , λ_{2_i} needs to also be positive to meet the synchronization condition of Theorem 2. If K is a positive definite diagonal matrix then all subsystems need to be coupled in order to meet the synchronization condition. This results in a conservative coupling structure.

For certain nodal structures it is possible to find conditions on the values of the elements $\tilde{\gamma}_i$ such that $A(\tilde{\gamma})^*D + DA(\tilde{\gamma}) < 0$ using matrices K that are only positive semi-definite. For example the work in [12, 23] uses the results of [2] to show that for nodes with a cyclic feedback structure there exists $D > 0$ such that $A(\tilde{\gamma})^*D + DA(\tilde{\gamma}) < 0$ if and only if the *secant condition*

$$\tilde{\gamma}_1 \cdots \tilde{\gamma}_n < \sec^n\left(\frac{\pi}{n}\right)$$

is satisfied. The value of each $\tilde{\gamma}_i$ can be made arbitrarily small (and positive) by increasing k_i . Therefore to construct a matrix $D > 0$ such that $A(\tilde{\gamma})^*D + DA(\tilde{\gamma}) < 0$, it is sufficient to increase the values of the elements k_i to the point where the secant condition is met. In fact, as shown in [12, 23], if all elements γ_i are positive, the secant condition can be met by making only a single element k_i positive and sufficiently large in magnitude.

5.4.3 Network of Repressilator Circuits

As an example demonstrating the methods presented in this chapter, we consider the synchronization of a network of repressilator circuits, [8]. The repressilator is a synthetic oscillating genetic circuit that was developed in *Escherichia coli* (*E. coli*), and is composed of a cyclic network of three genes and their protein products, wherein each protein inhibits the transcription of the next gene in the cycle. The circuit is illustrated schematically in Fig. 5.2 and works in the following cyclic manner:

- The *E. coli* gene *lacI* expresses the protein LacI which inhibits transcription of the gene *tetR*.
- The gene *tetR* expresses the protein TetR which inhibits transcription of the gene *cI*.
- The gene *cI* expresses the protein CI which inhibits transcription of the gene *lacI*.

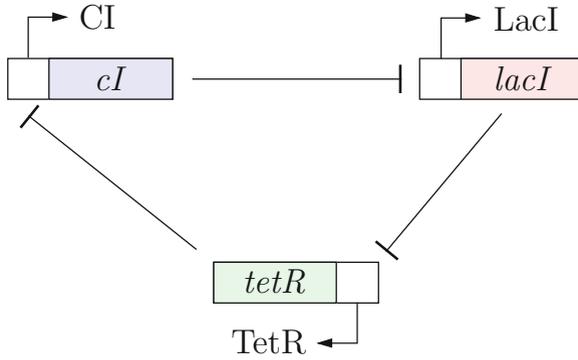


Fig. 5.2 The repressilator genetic network

A dimensionless dynamical model of the repressilator is given in (5.19)–(5.24), where states x_{1j}, x_{3j}, x_{5j} respectively represent concentrations of the mRNA transcribed from *lacI*, *tetR* and *cI* and states x_{2j}, x_{4j}, x_{6j} respectively represent concentrations of the proteins LacI, TetR and CI (here, the subscript j is an index denoting the particular repressilator circuit for the network analysis which is to follow).

$$\dot{x}_{1j} = -x_{1j} + g(x_{6j}) \tag{5.19}$$

$$\dot{x}_{2j} = -x_{2j} + x_{1j} \tag{5.20}$$

$$\dot{x}_{3j} = -x_{3j} + g(x_{2j}) \tag{5.21}$$

$$\dot{x}_{4j} = -x_{4j} + x_{3j} \tag{5.22}$$

$$\dot{x}_{5j} = -x_{5j} + g(x_{4j}) \tag{5.23}$$

$$\dot{x}_{6j} = -x_{6j} + x_{5j} \tag{5.24}$$

where $g(x_{ij}) = \begin{cases} \frac{5}{1+x_{ij}^2} & x_{ij} \geq 0 \\ 5 & x_{ij} \leq 0 \end{cases}$

In [9] a modification to the repressilator circuit is proposed that enables the coupling of the multiple such circuits for the purpose of building a synchronized genetic clock. The authors propose the inclusion in the repressilator of an inter-cellular communication mechanism found in the bacterium *Vibrio fischeri*. In this mechanism, the protein LuxI is used to synthesize an autoinducer (AI) molecule which diffuses through the cell membrane. The AI forms a complex with the protein LuxR, which in turn activates the transcription of certain genes. The authors suggest that this coupling mechanism be added to the repressilator circuit in *E. coli* in addition to an extra copy of the *lacI* gene so that the coupling functions as a feedback loop in the following way

$$\dot{x}_{1j} = -x_{1j} + g(x_{6j}) + f(x_{9j}) \quad (5.27)$$

$$\dot{x}_{2j} = -x_{2j} + x_{1j} \quad (5.28)$$

$$\dot{x}_{3j} = -x_{3j} + g(x_{2j}) \quad (5.29)$$

$$\dot{x}_{4j} = -x_{4j} + x_{3j} \quad (5.30)$$

$$\dot{x}_{5j} = -x_{5j} + g(x_{4j}) \quad (5.31)$$

$$\dot{x}_{6j} = -x_{6j} + x_{5j} \quad (5.32)$$

$$\dot{x}_{7j} = -x_{7j} + g(x_{2j}) \quad (5.33)$$

$$\dot{x}_{8j} = -x_{8j} + x_{7j} \quad (5.34)$$

$$\dot{x}_{9j} = -x_{9j} + x_{8j} - \frac{1}{N} \rho_9 \sum_{k=1}^N (x_{9j} - x_{9k}) \quad (5.35)$$

Here, ρ_9 is a measure of coupling strength. The uncoupled ($\rho_9 = 0$) model (5.27)–(5.35) is illustrated in Fig. 5.4, where each block represents an incrementally passive subsystem. For $i = 1, \dots, 9$ each subsystem H_i represents the dynamic block

$$\begin{aligned} \dot{x}_{ij} &= -x_{ij} + e_{ij}, \quad e_{ij} = u_{ij} + w_{ij} \\ y_{ij} &= x_{ij} \end{aligned}$$

Each H_i^s represents the monotonically increasing static map

$$y_{ij}^s = \begin{cases} -g(u_{ij}^s), & i = 2, 4, 6 \\ f(u_{ij}^s), & i = 9 \end{cases}$$

Inputs u_{ij} to each dynamic block H_i are such that

$$u_{ij} = \begin{cases} y_{i-1j}, & i = 2, 4, 6, 8 \\ -y_{6j}^s + y_{9j}^s, & i = 1 \\ -y_{i-1j}^s, & i = 3, 5 \\ -y_{2j}^s, & i = 7 \\ y_{8j}, & i = 9 \end{cases}$$

Inputs u_{ij}^s to each static block H_i^s are such that

$$u_{ij}^s = \begin{cases} y_{i-1j}, & i = 2, 4, 6 \\ -y_{9j}, & i = 9 \end{cases}$$

For blocks H_i the incremental storage function $S_i = \frac{1}{2} \Delta x_{ij,m}^2$ satisfies the incremental dissipation inequality

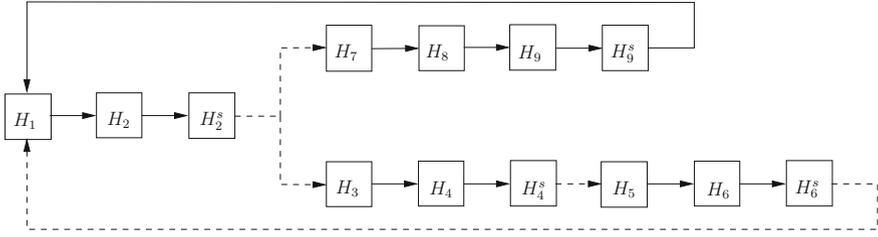


Fig. 5.4 Block diagram representation of the uncoupled repressilator genetic circuit of the model (5.27)–(5.35). Dashed lines represent inhibitory reactions

$$\dot{S}_i = -\frac{1}{\gamma_i} (\Delta y_{i,j,m})^2 + \Delta y_{i,j,m} \Delta e_{i,j,m}$$

where $\gamma_i = 1$ for $i = 1, \dots, 9$. For blocks H_i^s the incremental storage function $S_i^s = 0$ satisfies the incremental dissipation inequality

$$\dot{S}_i^s \leq -\frac{1}{\gamma_i^s} (\Delta y_{i,j,m})^2 + \Delta y_{i,j,m}^s \Delta y_{i,j,m}$$

where

$$\gamma_i^s = \begin{cases} \sup_{x_{i_j} \in \mathbb{R}} -g'(x_{i_j}) & i = 2, 4, 6 \\ \sup_{x_{i_j} \in \mathbb{R}} f'(x_{i_j}) & i = 9 \end{cases}$$

Note that from the definitions of outputs y_{i_j} for this example, each node is incrementally zero-state observable since for all i and any j, m $|y_{i_j} - y_{i_m}| = 0 \Leftrightarrow |x_{i_j} - x_{i_m}| = 0$. This implies the iZSD property required in Theorem 2 to deduce asymptotic state synchronization from asymptotic output synchronization.

With the above definitions of the input–output relations between the different blocks of Fig. 5.4, it is now possible to construct an incremental storage function $S(X_\Delta)$ for the incremental state vector X_Δ . For the two sets $\mathcal{I} = \{1, \dots, 9\}$ and $\mathcal{I}^s = \{2, 4, 6, 9\}$ let

$$S(X_\Delta) = \frac{1}{2N} \sum_{j=1}^N \sum_{j=1}^N \left(\sum_{i \in \mathcal{I}} d_i S_{i,j,m} + \sum_{i \in \mathcal{I}^s} d_i^s S_{i,j,m}^s \right)$$

where $d_i, d_i^s > 0$. Defining

$$Y = [Y_1^* \ Y_2^* \ Y_2^{s*} \ Y_3^* \ Y_4^* \ Y_4^{s*} \ Y_5^* \ Y_6^* \ Y_6^{s*} \ Y_7^* \ Y_8^* \ Y_9^* \ Y_9^{s*}]^*$$

where $Y_i^s = [y_{i_1}^s \ \dots \ y_{i_N}^s]^*$ and defining

$$K = \text{diag}\{k_1, k_2, k_2^s, k_3, k_4, k_4^s, k_5, k_6, k_6^s, k_7, k_8, k_9, k_9^s\}$$

where

$$\begin{cases} k_i = 0 & i = 1, \dots, 8 \\ k_i > 0 & i = 9 \end{cases} \quad \text{and} \quad k_i^s = 0, \quad i \in \mathcal{I}^s \quad (5.36)$$

we then have the incremental dissipation inequality

$$\begin{aligned} \dot{S} &\leq Y^* \tilde{\Pi}_n^* ((A(\gamma)^* D + DA(\gamma)) \otimes I_N) \tilde{\Pi}_n Y + W_9^* \Pi^* \Pi Y_9 \\ &\leq Y^* \tilde{\Pi}_n^* ((A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) \otimes I_N) \tilde{\Pi}_n Y + k_9 Y_9^* \Pi^* \Pi Y_9 + W_9^* \Pi^* \Pi Y_9 \end{aligned} \quad (5.37)$$

with $D = \text{diag}\{d_1, d_2, d_2^s, d_3, d_4, d_4^s, d_5, d_6, d_6^s, d_7, d_8, d_9, d_9^s\}$ and the interconnection matrix

$$A(\tilde{\gamma}) = \begin{bmatrix} -\frac{1}{\tilde{\gamma}_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\ 1 & -\frac{1}{\tilde{\gamma}_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{\tilde{\gamma}_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -\frac{1}{\tilde{\gamma}_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -\frac{1}{\tilde{\gamma}_5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_6} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_6} & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tilde{\gamma}_7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_9} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_9} \end{bmatrix} \quad (5.38)$$

where

$$\tilde{\gamma}_i = \frac{\gamma_i}{1 + k_i \gamma_i} \quad \text{and} \quad \tilde{\gamma}_i^s = \frac{\gamma_i^s}{1 + k_i^s \gamma_i^s}$$

and where the interconnection matrix $A(\gamma)$ is of the same structure as $A(\tilde{\gamma})$ but with $\tilde{\gamma}_i, \tilde{\gamma}_i^s$ replaced with γ_i, γ_i^s .

Theorem 1 shows that there always exists a matrix $K \geq 0$ such that (5.37) is iOSP($-K$). However in the case of this network the only state directly coupled to others in the network is the state x_{9_j} , and therefore the coupling can only compensate for a shortage of incremental passivity as in Theorem 2 if K is limited to the form (5.36) since the matrix L in Theorem 2 is constrained by the coupling to the structure

$$\begin{cases} \lambda_{2_i} = 0 & i = 1, \dots, 8 \\ \lambda_{2_i} > 0 & i = 9 \end{cases} \quad \text{and} \quad \lambda_{2_i}^s = 0, \quad i \in \mathcal{I}^s$$

In other words, with this incremental output passification method, it will only be possible to prove asymptotic global state synchronization under strong coupling if decreasing $\tilde{\gamma}_9$ by increasing k_9 is sufficient to guarantee that there exists $D > 0$ such that $A(\tilde{\gamma})^*D + DA(\tilde{\gamma}_i) < -\epsilon_{D,K}I_n$ for some $\epsilon_{D,K} > 0$.

To see if this is possible, first note the branched structure of the block diagram shown in Fig. 5.4. A similar branched structure was analyzed in [20], which derived a necessary and sufficient condition on the quantities $\tilde{\gamma}_i$ and $\tilde{\gamma}_i^s$ for the diagonal stability of the interconnection matrix associated with the branched structure in that reference. The nodal structure in Fig. 5.4 is different to that in [20]. However, similar arguments to those in [20] can be used to derive at least a necessary condition for the diagonal stability of $A(\tilde{\gamma})$ in (5.38). The main idea in [20] concerning such structures is that a necessary condition for the diagonal stability of $A(\tilde{\gamma})$ is that all its principal submatrices are also diagonally stable [3]. For $A(\tilde{\gamma})$, consider the principal submatrix obtained by deleting the 10th–13th rows and columns and the principal submatrix obtained by deleting the 4th–9th rows and columns. These are

$$A(\tilde{\gamma})_{(10-13)} = \begin{bmatrix} -\frac{1}{\tilde{\gamma}_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & -\frac{1}{\tilde{\gamma}_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{\tilde{\gamma}_3^s} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -\frac{1}{\tilde{\gamma}_3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_4^s} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -\frac{1}{\tilde{\gamma}_5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_6^s} & 0 \end{bmatrix}$$

and

$$A(\tilde{\gamma})_{(4-9)} = \begin{bmatrix} -\frac{1}{\tilde{\gamma}_1} & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -\frac{1}{\tilde{\gamma}_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{\tilde{\gamma}_2^s} & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -\frac{1}{\tilde{\gamma}_7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_9} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -\frac{1}{\tilde{\gamma}_9^s} \end{bmatrix}$$

These two principal submatrices exhibit a cyclic feedback structure (with negative feedback). As discussed above, the stability of systems with this structure was studied in [2], where it was shown that the matrix such as $A(\tilde{\gamma})_{(10-13)}$ is diagonally stable if and only if

$$\tilde{\gamma}_1\tilde{\gamma}_2\tilde{\gamma}_2^s\tilde{\gamma}_3\tilde{\gamma}_4\tilde{\gamma}_4^s\tilde{\gamma}_5\tilde{\gamma}_6\tilde{\gamma}_6^s < \sec^9\left(\frac{\pi}{9}\right) \quad (5.39)$$

whilst $A(\tilde{\gamma})_{(4-9)}$ is diagonally stable if and only if

$$\tilde{\gamma}_1 \tilde{\gamma}_2 \tilde{\gamma}_2^s \tilde{\gamma}_7 \tilde{\gamma}_8 \tilde{\gamma}_9 \tilde{\gamma}_9^s < \sec^7 \left(\frac{\pi}{7} \right) \quad (5.40)$$

Since $\tilde{\gamma}_9$ appears in (5.40) only, these two necessary conditions can only be met by strengthening the coupling if (5.39) is satisfied a priori.

If it were possible to modify the repressilator circuit so that the coupling state is x_{1_j} or x_{2_j} it could then be possible to diagonally stabilize $A(\tilde{\gamma})$ by reducing the incremental secant gains $\tilde{\gamma}_1$, $\tilde{\gamma}_2$ or $\tilde{\gamma}_2^s$ by increasing k_1 , k_2 , k_2^s and then compensating for the shortage of incremental passivity with strong coupling. This is because the quantities $\tilde{\gamma}_1$, $\tilde{\gamma}_2$, $\tilde{\gamma}_2^s$ appear in both (5.39) and (5.40), and these two necessary conditions can therefore be satisfied under such a change. To see this, we propose a modification to the repressilator model wherein *lacI* is replaced with a different gene, the protein product of which behaves as LacI in inhibiting the transcription of *luxI* and *tetR*, but with the difference that the new protein product is also a coupling signal in the same manner as AI. Equation (5.28) is then modified to

$$\dot{x}_{2_j} = -x_{2_j} + x_{1_j} - \frac{1}{N} \rho_2 \sum_{k=1}^N (x_{2_j} - x_{2_k}) \quad (5.41)$$

For this example, we consider a network of $N = 4$ nodes of the form (5.27)–(5.35) but with (5.28) replaced with (5.41). The incremental secant gains γ_i can be calculated from the model to be as follows:

$$\gamma_i = 1, \text{ for } i = 1, \dots, 9 \quad \gamma_i^s = 3.25, \text{ for } i = 2, 4, 6 \quad \gamma_9^s = 1$$

Since the coupling is only through the states x_{2_j} and x_{9_j} , strong coupling can incrementally passify the network nodes only if $K = \text{diag}\{0, k_2, 0, \dots, 0, k_9\}$ as L would have the same structure as K .

We set $k_2 = 12$ and $k_9 = 0$, which makes $\tilde{\gamma}_2 = 0.04$ and $\tilde{\gamma}_9 = \gamma_9 = 1$. Otherwise, $k_i = k_i^s = 0$, which leaves $\tilde{\gamma}_i = \gamma_i$, $\forall i \neq 2$ and $\tilde{\gamma}_i^s = \gamma_i^s$ for $i = 2, 4, 6, 9$. Without a sufficient condition on the gains $\tilde{\gamma}_i$ that guarantees the diagonal stability of $A(\tilde{\gamma})$, it is nevertheless possible to use an LMI solver to find a matrix $D > 0$ such that $A(\tilde{\gamma})^* D + DA(\tilde{\gamma}) < 0$. One possible matrix D is given by

$$D = \text{diag}\{1, 17, 208, 53, 57, 19, 7, 7, 3, 17, 18, 19, 26\} \quad (5.42)$$

which is such that $A(\tilde{\gamma})^* D + DA(\tilde{\gamma}) < -0.0014I_n$ and therefore $S(X_\Delta)$ satisfies the incremental dissipation inequality

$$\begin{aligned} \dot{S} &\leq (\tilde{\Pi}_n Y)^* ((A(\tilde{\gamma})^* D + DA(\tilde{\gamma})) \otimes I_N) \tilde{\Pi}_n Y + k_2 Y_2^* \Pi^* \Pi Y_2 + W_2^* \Pi^* \Pi Y_2 \\ &\leq -0.0072 (\tilde{\Pi}_n Y)^* \tilde{\Pi}_n Y + k_2 Y_2^* \Pi^* \Pi Y_2 + W_2^* \Pi^* \Pi Y_2 \end{aligned} \quad (5.43)$$

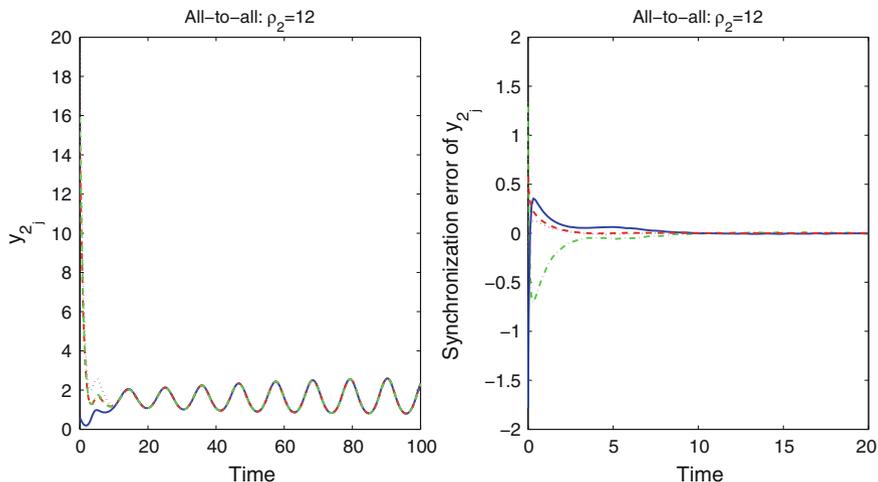


Fig. 5.5 Synchronization of four repressilator circuits. *Left* Synchronization of output y_{2_j} . *Right* Synchronization error of output y_{2_j}

The coupling is given by the relation $W_2 = -\Gamma_2 Y_2$ where, from (5.41), Γ_2 has the all-to-all structure

$$\Gamma_2 = \rho_2 \left(I_N - \frac{1}{N} \mathbf{1}_{N \times N} \right)$$

and $\lambda_{2_2} = \rho_2$. From Theorem 2, we require $\lambda_{2_2} \geq k_2$ to guarantee synchronization.

This condition therefore requires $\rho_2 > 12$. The simulation in Fig. 5.5 shows the synchronization of the output y_{2_j} across the network nodes as well as the asymptotic stability of the synchronization error in the output y_{2_j} under this coupling. Note that this figure also demonstrates the synchronization of state x_{2_j} since $y_{2_j} = x_{2_j}$.

5.5 Discussion

This chapter has presented a constructive approach to finding sufficient conditions for global asymptotic state synchronization in networks of identical nodes. The principal assumptions are that each node is composed of iOFP subsystems, and that these subsystems are directly coupled to their corresponding subsystems on other nodes in the network using linear static coupling.

By taking advantage of the iOFP property, it was possible to quantify the degree of the shortage of incremental passivity of each node (Theorem 1). In Theorem 2 it was shown that the nodal coupling can act as a passifying feedback and the degree of shortage of passivity was used to determine a lower bound on the minimum coupling strength required to render each node iOSP and hence guarantee asymptotic output

synchronization. With an additional iZSD assumption, this also implied asymptotic state synchronization. These two theorems demonstrated that for arbitrary nodal structures satisfying Assumption 1, it is always possible to characterize the shortage of passivity and it is always possible to find a strong enough coupling topology that can eliminate this shortage and thus achieve asymptotic state synchronization.

Inequalities such as (5.4) is an incremental dissipation inequality that can be used to represent general (iOFP $\left(\frac{1}{\gamma_i}\right)$) subsystems. Therefore if a given network satisfies our sufficient conditions for synchrony and if we were to replace the i th subsystem of each node in the network with another subsystem that has a shortage of passivity that is equal to or less than that of the original subsystem, the network with the new subsystem would also synchronize. This ability to modify the parameters, and indeed the structure, of the network subsystems and yet maintain synchrony lends a significant degree of robustness to the results we have presented. In the applied setting of synchronizing biochemical reaction networks such as [10], where biological parameters typically vary significantly, placing a biologically plausible upper bound on the quantity γ_i would allow us to analyze synchrony in such a system in a way that is robust to such parametric variations. Furthermore, the proposed methodology can also have implications for the design of synthetic circuits that synchronize upon interconnection because it can yield insight into the what system outputs could serve as network coupling signals that lead to incremental stability.

References

1. Angeli D (2002) A Lyapunov approach to incremental stability properties. *IEEE Trans Autom Control* 47:410–422
2. Arcak M, Sontag E (2006) Diagonal stability for a class of cyclic systems and applications. *Automatica* 42:1531–1537
3. Barker GP, Berman A, Plemmons RJ (1978) Positive diagonal solutions to the Lyapunov equations. *Linear and Multilinear Algebra* 5(4):249–256
4. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: structure and dynamics. *Phys Rep* 424(4–5):175–308
5. Cremean LB, Murray RM (2003) Stability analysis of interconnected nonlinear systems under matrix feedback. In: *Proceedings of the 42nd Conference on Decision and Control, Maui, Hawaii, USA, vol 4, pp 3078–3083*
6. Demidovich BP (1967) *Lectures on stability theory*. Nauka, Moscow
7. Dockery JD, Keener JP (2001) A mathematical model for quorum sensing in *Pseudomonas aeruginosa*. *Bull Math Biol* 63(1):95–116
8. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338
9. Garcia-Ojalvo J, Elowitz MB (2004) Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proc Nat Acad Sci* 101(30):10955–10960
10. Gonze D, Bernard S, Waltermann C, Kramer A, Herzel H (2005) Spontaneous synchronization of coupled circadian oscillators. *Biophys J* 89:120–189
11. Hamadeh A, Stan G, Goncalves J (2008) Robust synchronization in networks of cyclic feedback systems. In: *Proceedings of the 47th IEEE Conference on Decision and Control (IEEE-CDC)*
12. Hamadeh AO, Stan GB, Sepulchre R, Goncalves JM (2012) Global state synchronization in networks of cyclic feedback systems. *IEEE Trans Autom Control* 57:478–483

13. Lohmiller W, Slotine JJE (1998) On contraction analysis for nonlinear systems. *Automatica* 34(6):683–696
14. Olfati-Saber R, Murray RM (2004) Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans Autom Control* 49(9):1520–1533
15. Pavlov A, Pogromsky A, van de Wouw N, Nijmeijer H (2004) Convergent dynamics, attribute to Boris Pavlovich Demidovich. *Syst Control Lett* 52:257–261
16. Pogromsky A, Nijmeijer H (2001) Cooperative oscillatory behavior of mutually coupled dynamical systems. *IEEE Trans Circuits Syst I: Fundam Theory Appl* 48(2):152–162
17. Scardovi L, Arcak M, Sontag E (2009) Synchronization of interconnected systems with an input–output approach. Part I: main results. In: *Proceedings of the 48th IEEE Conference on Decision and Control*, pp 609–614
18. Sepulchre R, Jankovic M, Kokotovic P (1997) *Constructive nonlinear control*. Springer, New York
19. Sontag E (2005) Passivity gains and the “secant condition” for stability. *Automatica* 55:177–183
20. Sontag ED, Arcak M (2008) *Lecture notes in control and information sciences*. Chapter passivity-based stability of interconnection structures. Springer, Berlin/Heidelberg, pp 195–204
21. Stan GB (March 2005) *Global analysis and synthesis of oscillations: a dissipativity approach*. PhD thesis, University of Liege
22. Stan GB, Sepulchre R (2007) Analysis of interconnected oscillators by dissipativity theory. *IEEE Trans Autom Control* 52:256–270
23. Stan GB, Hamadeh A, Sepulchre R, Goncalves J (2007) Output synchronization in networks of cyclic biochemical oscillators. In: *Proceedings of the 26th IEEE American Control Conference (IEEE-ACC)*
24. Vidyasagar M (1981) *Input–output analysis of large scale interconnected systems*. Springer-Verlag, Berlin
25. Willems J (1972) Dissipative dynamical systems: parts I and II. *Arch Ration Mech Anal* 45:321–393

Chapter 6

Robustness Model Validation of Bistability in Biomolecular Systems

Luca Salerno, Carlo Cosentino, Alessio Merola, Declan G. Bates and Francesco Amato

Abstract Bistability is a key system-level dynamical property to understand the basic mechanisms underpinning some cellular functions, like persistent memory, switch-like biochemical responses and irreversible cell differentiation. These processes are guaranteed by evolved molecular modules, involving genes, proteins and metabolites, which implement transitions between distinct operative conditions in response to exogenous and endogenous signals. In many cases, such a coordinated control action leads to a change in the dynamic behaviour of the cell, which persists even after the activating signal (e.g., the concentration of a certain molecular species) has returned to the initial concentration. A propaedeutical step to the construction of biomodels for this class of systems, is the analysis of the structure of the underlying biochemical reaction network; in particular, a necessary requirement is that the topology of this network is compatible with the assumed bistable behaviour. Subsequently, one can face the question of whether the same performance is guaranteed even in the presence of endogenous and exogenous perturbations, i.e., whether the model is *robustly* bistable in the face of, e.g., parametric uncertainty (deriving from interindividual variability) or fluctuating environmental conditions (due to the intrinsic stochastic nature of cellular processes). The present chapter focusses on the

L. Salerno · C. Cosentino (✉) · A. Merola · F. Amato
Dipartimento di Medicina Sperimentale e Clinica, Università degli Studi Magna Græcia di Catanzaro, Catanzaro, Italy
e-mail: carlo.cosentino@unicz.it

F. Amato
e-mail: amato@unicz.it

L. Salerno
e-mail: l.salerno@unicz.it

A. Merola
e-mail: merola@unicz.it

D. G. Bates
College of Engineering, Mathematics and Physical Science, University of Exeter, Exeter, UK
e-mail: D.G.Bates@exeter.ac.uk

presentation of methodological approaches for the characterization of bistability in biological systems; such methods are important as tools to assess the plausibility of mathematical models of biosystems that exhibit bimodal experimental behaviour, without recurring to large-scale computational simulations.

Keywords Bistability · Robustness analysis · Domain of attraction · Local sensitivity · Bifurcation · Global sensitivity · Multistable dynamics · Pathway · Regulatory network · Molecular switches · *Escherichia coli* (*E. coli*)

6.1 Introduction

Many biological systems have been experimentally verified to be robust with respect to significant variations of the environmental conditions. Accordingly, one would expect such robustness to be exhibited by the corresponding mathematical models in order for the latter to be considered plausible. Model robustness entails that the qualitative behaviour of the system is insensitive, to a certain extent, to parametric variations. In the biological context, robustness is essential to guarantee that the specific signaling, metabolic and regulatory tasks are accomplished, despite the significant inter-individual variability of kinetic parameters and molecular species concentrations. In this chapter, we outline a procedure, based on control theoretical results and convex optimization tools, to characterize bistable biomolecular circuits in terms of robustness and domains of attraction (DA) of the equilibrium points. The proposed procedure can then be applied to test the plausibility of biological models in terms of their capability to robustly reproduce bistable dynamics. Therefore, this approach provides a suitable and effective tool for tackling the problem of validation/invalidation of models of biomolecular regulatory networks.

6.2 Bistable Dynamics in Biological Systems

The mechanisms through which cells commit to a certain fate, alternate between different physiological modes or implement persistent cellular memory are, essentially, guaranteed by simple key modules of interactions, which form the core of multistability and are widely present in genetic, metabolic and signaling networks, both in prokaryotic, [5, 25, 51, 60, 68] and eukaryotic organisms [6, 7, 12, 57, 72].

The term bistability appears as early as 1948, when Delbrück proposed the idea of bistability as a general principle to explain how discontinuous transitions between two stable states occur in biochemical reaction systems [23]. A few years later, Waddington developed the idea of *epigenetic landscape* as a metaphor to illustrate how gene regulatory mechanisms modulate cellular development [69]. During the same period, Novick and Weiner showed the occurrence of all-or-none transitions between different cellular states in *Escherichia coli*, depending on the lactose metabolism, and

that such states can be preserved across different generations even after removal of the chemical inducer and genomic mutation [56]. In 1961, Monod and Jacob proposed a bistable gene regulatory circuit in order to explain cell differentiation [53], originating the idea that control of intracellular enzyme levels occurs through transcriptional feedback. Another key contribution to the topic of bistability in biology is provided by Thomas [64]: he showed that a necessary condition for the existence of multistable dynamics and, consequently, a switch-like behavior is the presence of a positive feedback loop. Notwithstanding these early contributions on the topic, only recently the concept of bistability as a key property of mathematical biomodels is receiving a significant attention in the field of molecular developmental biology.

In what follows, we review some biological examples of molecular switches and report the necessary conditions that must be satisfied by a regulatory network to be compatible with bistability. These conditions provide a simple and effective method for testing the structural bistability of a biological network, based on the analysis of its topology.

6.2.1 Cellular Switches

In all living organisms, the key components of life, i.e. genes, proteins and metabolites, are highly interconnected in evolved coordinated molecular modules. These control modules govern the changes in the dynamical behaviour of the cell, like irreversible differentiation or reversible phenotypic alteration. When the outcome of these changes persists even after the activation signal has returned to pre-stimulus levels, the system implements a molecular switch. The control mechanisms implementing molecular switching are inherently non-linear, indeed a key feature of this class of systems is the presence of multiple equilibrium points.

Some examples of regulatory molecular switches concern prokaryotic and eukaryotic organisms. In prokaryotic cells, genes involved in a specific biological process are expressed in a coordinated manner from a single promoter in a natural logic module, called operon. The *lac* operon in *E. coli*, a well-investigated prokaryotic system [41, 52], consists of three genes (*lacZ*, *lacY*, and *lacA*) involved in the metabolism and absorption of the disaccharide lactose. In particular, in the absence of lactose, a repressor protein binds to the operator, inhibiting the transcription of the *lac* operon genes (in this case the operon is in “off” state). When the inducer, lactose, is added, it binds to the repressor and changes the structure of the repressor, unbinding to the operator. While the operator remains free of the repressor, RNA polymerase recognizes the promoter and promotes the transcription of the operon structural genes into mRNA, (the operon is “on”). This makes the *lac* operon a classic example of a genetic switch relying on autocatalytic feedback [61].

In eukaryotic organisms, cellular switches characterize the mechanisms by which a cell determines own cell fate [31, 32, 48, 73]. A typical cellular switch is involved in *Xenopus*, where fully grown oocytes may persist in an immature state, or, in

response to the steroid hormone progesterone, can be induced to differentiate into fertilizable oocytes, (this process is called oocyte maturation). The oocyte maturation is triggered by a protein kinase cascade (defined the Mos-Mek-MAPK cascade), which is activated just by the hormone progesterone. This biological system provides an interesting example of how signaling cascades leading the cell into the cell cycle, depending on the environmental changes, showing a bistable behavior. More precisely, for low concentrations of the hormone progesterone, the oocytes did not mature, whereas for high concentrations, the majority of oocytes can mature [31]. It is interesting to note also that, after the removal of progesterone, mature oocytes did not de-mature, demonstrating the irreversible nature of this process [73]. In *Saccharomyces cerevisiae*, one of the most extensively studied genetic switch is that implemented by the regulatory network of the galactose metabolic pathway: *S. cerevisiae* normally uses glucose as the carbon and energy source. However, in the absence of glucose, it can metabolize galactose as an alternative source, through Leloir metabolic pathway [14]. Such pathway is regulated by a set of enzymes, commonly referred to as the *GAL* system, whose molecular interactions have been thoroughly dissected in the last years through the classical reductionist molecular biology experimental approach.

6.2.2 Modeling of Bistability

The identification of the components for a regulatory network, as transcription factors and other molecular regulators, plays a very important role in the analysis and understanding of the mechanisms controlling biological state transitions. While many of the biological regulatory mechanisms have already been investigated, many other systems remain still to be identified. To analyze theoretically the transitions from one cellular state to another, it needs to have a specific mathematical framework that describes, in a formal way, the structural relationships between the components involved in the control mechanism [9, 55], and that allows us to quantify the dynamical behavior of the system, given a certain structure of the network [3, 34, 39]. In this context, the theory of dynamical systems and non-linear dynamics [63] offers a powerful tool for the description and analysis of the phenomenon of switching in biological processes. Mathematically, the basic idea of this method is to represent each different cellular state as an attractor of dynamical systems. The existence of these attractors, and accordingly, the possibility that cells can toggle between an attractor (a specified functional state of the cell, e.g. undifferentiated cell) and the other (e.g., differentiated cell) and, potentially, proceed backwards (in this case may occur hysteresis phenomenon, often associated with the exhibition of the bistability) depends on the network structure. More precisely, the existence of attractors depends on several factors, such as the structure of the network (as the form of the interactions between the connected elements), the dynamical characterization of the network (such as the parameter related to these interactions), the particular state at which the cell is located (e.g., the actual expression profile of their genes, or the con-

centrations of proteins and metabolites) and, finally, any disturbances on the system (such as perturbations in the cellular environment).

The ability to ensure the efficient operative conditions of the biological mechanisms, even in the face of continued internal or external uncertainty (e.g., molecular noise, fluctuations, changes in the cellular environment), leads to the hypothesis that in biological systems there is a considerable presence of feedback mechanisms that govern regulatory and decision-making processes, [21]. Indeed, in the last decade, the literature has decisively pointed to highlight the ubiquity of feedback control in biomolecular systems, [13, 24, 33].

Thus, the switching phenomena can arise, in principle, in the presence of a positive feedback in the cascade of signals that involve. The positive feedback loop can be implemented in a double positive form with two components mutually activation. A positive feedback may also be implemented in the form of a double negative feedback with two mutually inhibitory components, which acts as a switch, fostering a cell to alternate between two alternative states [10, 32, 34]. The positive feedback loop is one common module in establishing bistable motifs, focusing on the behavior of the component parts (genes or proteins). Over the past decade, the study on the transcriptional regulatory networks of the bacterium *E. coli* [59] and the simple organism *S. cerevisiae* [49] revealed a core of recurring regulatory modules, each with a characteristic structure and the ability to perform specific functions of biological processing [3, 12, 65].

To characterize the dynamics of these structures, a parameter evaluation must be performed, to check bistability-exhibiting parameter space more thoroughly. This leads to one of the main problems in modeling the biological systems, namely, determining the parameter set for which a given structure, in its interaction collectivity, produces a specific physiological function or a particular phenotype. Theoretically, one would to measure each parameter individually, but this task becomes very difficult in the analysis of complex networks whose dynamics, subjected often to different operating conditions, are performed by many parameters. Supporting this issue, there are no general methods, but different mathematical approaches have been proposed that allow the determination of the occurrence of bistability in a network of biochemical reactions.

6.2.2.1 The Chemical Reaction Network Theory

Given a certain reaction network, the capability of the system to admits two or more attractors is tied by the mathematical form of the related reaction rate system and by the values of the kinetic parameters. In general, it is very hard to infer any conclusion without restricting the scope of the analysis to a more specific class of system. The characterization of steady state attractors of mass-action systems can be dealt with a powerful analysis tool, Chemical Reaction Network Theory (CRNT) [28, 29]. CRNT joins the structure of a biochemical network, endowed by applying the principle of mass-action, to the capability of the network to engender multiple positive steady-states. The huge potential of this method is that it provides

an immediate way to analyze the type of dynamical behavior that one can expect from an arbitrarily complex reaction network, just by inspection of the topology of the associated graph. More specifically, this theory allows us to establish the conditions for the existence, multiplicity and stability of attractors for the associated ordinary differential equations (ODE) system, without the need to know a priori the reaction rate equations nor to choose precise values for the kinetic parameters. A given network is associated with a non-negative integer called *deficiency* [27, 38], that is not dependent on the values assigned to the parameters, but only on the underlying network of reactions. The development of the theory, in its completeness, is based on the assumption that behind the nonlinearity of systems endowed with mass-action kinetics, there exists considerable degree of hidden linearity. This feature allows to bring the dynamical properties in more simple qualitative behaviors, making the CRNT specifically suitable in the multistability analysis of the biomolecular systems, whose parameters are often unknown or subject to significant variability among different individuals.

However, for more accurate treatment of the CRNT, the interested reader is referred to the original articles [28, 29], or to [21] for an introductory overview of the main results. Despite the complexity of its theoretical foundations, the algorithm is coded in the CRNT Toolbox.¹ The CRNT toolbox indicates either that the associated dynamical system admits multiple positive steady states for some values of the kinetic parameters, or else that no such combination of the parameter values exists. In the affirmative case, the algorithm will also provide a set of values of the kinetic parameters for which the system is multistable.

6.3 Robustness Analysis of Bistable Biomolecular Systems

The idea of robustness refers to a characteristic of complex biological systems that evolve in an environment constantly changing and unpredictable. This characteristic is a property of the system, which cannot be understood looking at the level of individual components, and is ubiquitous in living organisms. It should be pointed out that robustness does not coincide with the stability or homeostasis concepts; robustness is associated with the preservation of a certain function of the system, rather than the state of the system. More precisely, the robustness, intended as the capability of a system to maintain its functionality through flexible changes at the structural level or in their operating conditions, despite environmental or internal perturbations (tolerance to stochastic fluctuations, noise, mutations), is one of the fundamental aspects in the analysis of cellular systems [62]. Following the occurrence of a perturbation, the system can return in the current steady state (attractor), this case is defined robust adaptation, or move to a new steady state which maintains the same functionality [44]. In the analysis of biological bistable systems, the property

¹ The CRNT algorithm is implemented in the CRNT toolbox, available at <http://www.che.eng.ohio-state.edu/~feinberg/crnt/>.

that must be preserved corresponds to the bistable qualitative dynamics, which, in this case, can be defined also robust if both the number and the type of attractors and both the connectivity and the shape of the trajectories remain the same, even in the presence of internal/external perturbations.

The robustness in biological systems is often ensured by the high degree of complexity that involves control mechanisms (feedback), modularity, and fail-safe mechanisms (redundancy and diversity). The control mechanisms, by means of feedback loops promote the adaptation of the system by modulating the intensity of the stimulus; the modularity is the mechanism to contain the perturbations and damages to a local level, globally minimizing their effects by means of a hierarchical organization; finally, the fail-safe (or “backup”) mechanisms provide an increase of the robustness due to the presence of different mechanisms coupled to control systems to perform the same function. In particular, the redundant mechanisms correspond to numerous identical or similar elements (modules) that are replaced each other when one fails, while the diversity ensures that the same function is achieved in different ways, through a set of heterogeneous elements (e.g., phenotypic plasticity), [44].

It is possible to specify the robustness measure in different types, depending on variations related to the state variables, or to the parameter space. These perturbations cause consequences that have implications both by a mathematical that biological point of view. To get an estimation of the robustness associated with the state variables, one considers the robustness closely related to the initial condition from which a trajectory starts in the state space. Moreover, even further state constraints, for example related to the species conservation, can occur. Namely, an enzyme does not undergo net changes in its concentration in any reaction that catalyses it. If it is not synthesized or degraded, then its total concentration remains constant over time. Similarly, if a substrate is bound in different complexes or modifications, for example, multi-site phosphorylation, and it is not further synthesized or degraded, its total concentration remains constant. In the case k linear independent state constraints are present (note that, with respect to the states, these constraints can be also nonlinear), the dynamics is limited in a subspace of dimension $n' = n - k$, where n is the size of the state space. Then, the state space is partitioned in different slices of dimension n' , corresponding to a given set of constraint conditions, Fig. 6.1. Within each section, the dynamics keeps certain properties, with attractors, DA and stability in function of a state space of dimension n' (not n). However, its characteristic can be qualitatively modified if different constraints are considered. Therefore, also the space of these constraints is divided into regions, within each of which the dynamics of the slices corresponding remain qualitatively similar. Mathematically, the parameters and constraints with respect to the initial conditions have different behaviour. The parameters, that identify the dynamics, can be chosen independently from the initial conditions, while the state constraints, that limit the dynamics of the system, no. In summary, the robustness with respect to variations in the constrained states can take two forms. If the variations concern only the initial conditions and the constraint conditions are considered fixed, then it can confirm the stability of the attractor and the size and shape of its DA in the state space (with respect to the state space of actual dimension d) provide a measure of it. Instead, if the constraint conditions are varied,

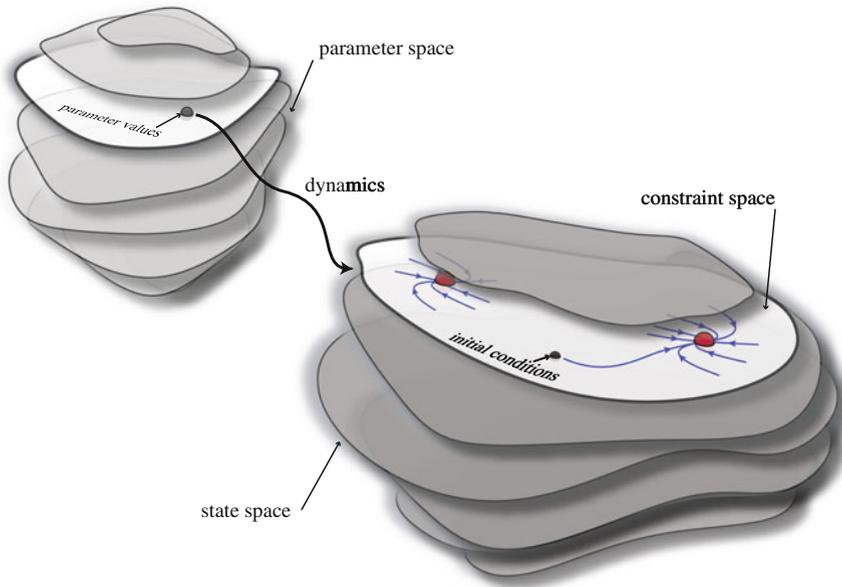


Fig. 6.1 Abstract space representation of biological system suspected for bistability. The dynamical behavior of the system is characterized by a preassigned set of values in the parameter space, for which the system admits two stable attractors in the state space. For ODE models, these space can be opportunely represented in high-dimensional Euclidean space but, often, biological models are also defined by infinite dimensional or nonlinear space structure. Moreover, the presence of state constraints divides the state space in different slices, each of which, is related to a particular choice of these constraints. Given a particular point in the parametric space, starting from a given initial condition, within a slice in the state space, the system trajectories remain in the same slice during the evolution. For a bistable systems, this slice is characterized by the presence of two stable attractors, to one of which the trajectory can converge. The figure is readapted from [36]

then the robustness goes beyond the concept of stability and the size and shape of the appropriate region in constraint space become very relevant, [36].

Looking at the the robustness with respect to parametric variations, the biological interpretation of the changes in the parameter values depends both on the model and the nature of its parameters. The real parameters may be different from those used in the model (nominal parameters) due to slight variations over time, difficulties in the identification of the real values, changes in operating conditions (also defined as structured uncertainty), or dynamics (typically too fast) can often have been neglected in the derivation of the model, which are not considered (also referred to as unstructured uncertainty). To clarify the meaning of the parametric uncertainty in terms of robustness, consider an ODE model derived from the description of a network of biochemical reactions by means of the mass-action principle. In this case, the parameters correspond to rate constants of various nature: association and dissociation rates, constant degradation, and so on. However, the biological process can be described

according to different kinetics. The processing of information through cellular signalling cascades may occur over a wide range of time scales (e.g., fast ligand-receptor dynamics versus the much slower response of gene expression changes), thus, the separation of time scales is often very convenient to reduce the complexity. Similarly, models of allosteric enzymes [54] or the rate functions of gene expression in terms of association constants of the transcription factor [2] are based on Hill-functions. On a physiological time scale, the concentration is derived from equilibrium between synthesis and degradation and could easily vary from cell to cell within a single organism, tissue, or through clonal population of differences related to cell volume, of intrinsic noise transcription/translation, and stochastic subdivisions of molecules during cell division. These factors may also play a remarkable role during the evolution. Different methodologies to measure the robustness related to parametric variations have been proposed, such as single parameter robustness for dynamical attractors [50], robustness with respect to random multiparametric variations [15, 50], and robustness measures based on Monte-Carlo techniques, to estimate the robustness of bistability in the process of apoptosis [26]. However, the problem to consider high-dimensional systems makes very difficult to establish a theory for the measure of robustness in biological systems [45]. In this section we propose an overview of the existing methods for the analysis of the robustness in the systems biology field, compared to uncertainties of various nature, related to the states and parameters, which allow us to sketch a suitable framework for the robustness analysis of bistable systems.

6.3.1 *Robustness Measure to Initial Conditions*

In order to characterize the behavior of a bistable system (either in the engineering or biological field), it is possible to study the robustness of this property by estimating the DA of the two asymptotically stable equilibrium points. In general, the DA can be defined as the set of points, including the equilibrium point, such that every state trajectory starting in this set converges to that equilibrium point [43]. An accurate estimation of the DA extent and shape provides valuable information in terms of robustness about the ability of a bistable system to adapt to exogenous perturbations without switching to the different operative condition (i.e., a different biological state).

A Simple Bistable Switch. A useful example to show what information can be inferred by estimating the DA in a bistable biomolecular system is the gene autoregulation motif. A mathematical representation of gene expression with autoregulation [42] is described by the following

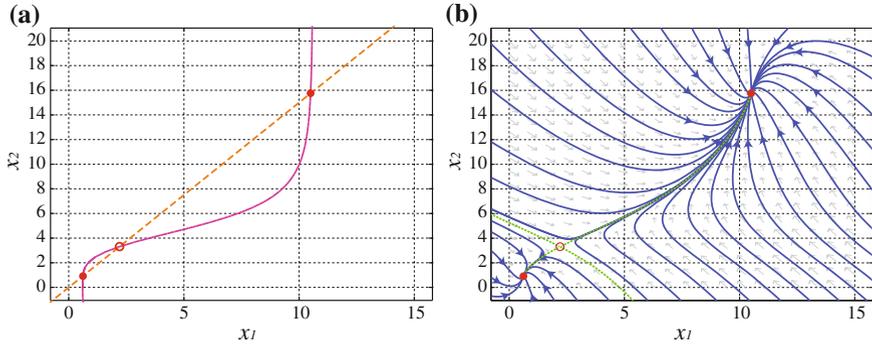


Fig. 6.2 **a** Points of intersection of the x_1 - and x_2 -nullclines (pink and orange line, respectively) represent the steady states of the system (6.1a, 6.1b). **b** The system trajectories on the $(x_1; x_2)$ phase plane converge to two attractors; the filled circles are the stable attractors, the empty one is the unstable equilibrium point. Arrows and blue lines show convergence of trajectories towards the stable steady states and away from the unstable steady state. The green dashed line, crossing the unstable steady state, represents locus of the unstable points, dividing the phase plane into the domains of attraction of the two stable equilibrium points

$$\dot{x}_1 = k_0 + k_1 \frac{x_2^n}{k_{d1}^n + x_2^n} - \mu_1 x_1 \quad (6.1a)$$

$$\dot{x}_2 = k_4 x_1 - \mu_2 x_2 \quad (6.1b)$$

where, \dot{x}_1 and \dot{x}_2 are the rate functions for the concentration of mRNA and protein, respectively, μ_1 and μ_2 are the decay rate of the mRNA and protein, k_0 , the basal transcription rate constant, k_1 and k_4 , the synthesis rate of mRNA and protein, respectively, k_{d1} , the effective affinity constant for protein x_2 , and n , the Hill coefficient for gene activation. In this bistable system there are two stable and an unstable attractor; each trajectory starting from different initial conditions may converge to one of the two stable equilibria, each of which is characterized by its DA. Thus, the state space is divided into these particular regions (see Fig. 6.2).

6.3.1.1 Estimation of the DA for Nonlinear Quadratic Systems

The exact estimation of dimensions and shape of the DA for a generic nonlinear system is usually a very hard problem to solve, especially for systems of medium/high order. The problem of determining the DA of the equilibrium points for uncertain polynomial systems has now been studied for a long time and several methods, based on Lyapunov functions, have been proposed, [35, 67]. More recently, Chesi has directed his attention to a novel convex optimization method, through polynomial relaxations based on Sum of Squares (SOS programming). In [17, 18], the author deals with the problem by means of Lyapunov and convex optimization methods: polynomial Lyapunov functions whose sublevel sets approximate the DA are numerically computed via Semi-Definite Programming (SDP).

Another method has been proposed by our group to compute invariant polytopic subsets of DA. The solution to this problem is greatly simplified when dealing with the class of nonlinear quadratic systems; indeed, by applying the method proposed in [4], it can be cast as a Linear Matrix Inequalities (LMI) feasibility problem [16]. The main result of Amato et al. is shown below. Since the proposed method yields a polytopic estimate of the DA, let us first recall the following definition:

Definition 1 A polytope (or box) $\mathcal{P} \subset \mathbb{R}^n$ can be defined as follows:

$$\begin{aligned} \mathcal{P} &= \text{conv}\{x_{(1)}, x_{(2)}, \dots, x_{(p)}\} \\ &= \{x \in \mathbb{R}^n : a_k^T x \leq 1, k = 1, 2, \dots, q\} \end{aligned} \quad (6.2)$$

Considering a generic dynamical system written in the following matrix form as

$$\dot{x}(t) = Ax(t) + B(x) + Nu(t), \quad (6.3)$$

where $x \in \mathbb{R}^n$ is the system state and

$$Bx = \begin{pmatrix} x^T B_1 \\ x^T B_2 \\ \vdots \\ x^T B_n \end{pmatrix} x, \quad (6.4)$$

with $B_i \in \mathbb{R}^{n \times n}$, the following problem can be solved by the proposed method. Note that for the sake of brevity, the statement of the problem refers to the equilibrium point in the origin. In this case, it is easy to generalize the definition and the entire procedure to equilibrium points different than zero, by adopting a change of the state variables, as shown in the origin article [4].

Problem 1 Assume that every eigenvalue of A in (6.3) has strictly negative real part; then, given a polytope \mathcal{P} , with the origin of the state space lying in the interior of \mathcal{P} , establish whether \mathcal{P} belongs to the DA of system (6.3).

At this point, we can recall the following sufficient conditions, under which Problem 1 admits a solution.

Theorem 1 *The polytope \mathcal{P} defined in (6.2) belongs to the DA of system (6.3) if there exist scalars γ , c and a symmetric matrix P such that*

$$0 < \gamma < 1, \quad (6.5a)$$

$$c > 0, \quad (6.5b)$$

$$P > 0, \quad (6.5c)$$

$$\begin{pmatrix} 1 & \gamma a_k^T \\ \gamma a_k & P/c \end{pmatrix} \geq 0, k = 1, 2, \dots, 2n \quad (6.5d)$$

$$x_{(i)}^T (P/c) x_{(i)} \leq 1, i = 1, 2, \dots, 2^n, \quad (6.5e)$$

$$\begin{aligned} & \gamma(A^T P + P A) + \begin{pmatrix} x_{(i)}^T B_1 \\ x_{(i)}^T B_2 \\ \vdots \\ x_{(i)}^T B_n \end{pmatrix} P \quad (6.5f) \\ & + (B_1^T x_i \ B_2^T x_i \ \dots \ B_n^T x_i) P < 0, i = 1, 2, \dots, 2^n \end{aligned}$$

The inequality conditions given in Theorem 1 can be easily solved through off-the-shelf efficient numerical algorithms (e.g., the LMILAB provided in the MATLAB Robust Control Toolbox [8]). In the Sect. 6.4, an application of this method will be provided.

6.3.2 Local Methods in Parametric Sensitivity Analysis

The sensitivity analysis provides a quantitative measure of the dependence of the behavior of the system, at some fixed points in the state space, with respect to the parameters. In general, the quantification of the individual contribution of the model parameters is the purpose to which is aimed a parametric sensitivity analysis, in order to understand the way how the nominal behavior of the biochemical reaction network changes with respect to parameter variations. Robustness and sensitivity are closely related; a relatively small sensitivity with respect to a particular property reflects on the robustness of the same property, so there must be a compromise between robustness and sensitivity. Some mechanisms invoking this compromise has been revealed, e.g., the robustness analysis in the system of heat shock response [46] and the steady-state analysis of polarized cells in yeast [20].

In addition to some measures of robustness, sensitivity analysis has been widely used to quantify the robustness with respect to perturbations of various parameters. In a complex system, such as a biological system, to describe which is required a high number of parameters, the behavior of the system can be robust to variations of some parameters, but sensitive to changes in others. Each parameter can affect differently

the dynamics of the system, with an intensity that can be measured by means of suitable techniques. Several methods are aimed at the definition and implementation of numerical local methods for the sensitivity determination of the parameters in various mathematical models, including those with stochastic characteristics and not constant parameters.

Let consider the system in Eq. (6.4) expressed in the following general ODE form

$$\dot{x}(t) = f(x(t); p), \quad \forall t \geq 0 \quad (6.6)$$

where $x \in \mathbb{R}^n$ is the system state, $p \in \mathbb{R}^m$ represents the parameter space, and f consists of the rate functions. If the system (6.6) admits a solution, the sensitivity of the system, defined by the matrix $S(t)$, which measures, locally to a point in the parameter space, p_{nom} , the influence of parameter variations on the system trajectories, can be defined as follows

$$S(t) = \left. \frac{\partial x_i}{\partial p_j} \right|_{x=x(t; p_{nom}), p=p_{nom}} = s_{i,j} \quad (6.7)$$

where $s_{i,j}$ are the sensitivity coefficients of each state x_i with respects to the parameter p_j , [40]. These sensitivity coefficients describe the change of the system output (state trajectory) at time t with respect to (an infinitesimal) perturbation of the parameter values. The sensitivity matrix $S(t)$, for each parameter p_j can be determined by means of finite differences

$$S(t) = \left. \frac{\partial x_i(t)}{\partial p_j} \right|_{x=x(t; p_{nom})} \simeq \frac{x(t; p_j + \Delta p_j) - x(t; p_j)}{\Delta p_j} \quad (6.8)$$

Since the state and parameter values can be subjected to variations in very large range of magnitude, in order to perform an appropriate comparison among states and parameters and to generate parameter ranking, normalized sensitivity value, $\bar{S}(t)$, is often used, which is calculated as follows:

$$\bar{S}(t) = \left. \frac{p_j}{x_i(t)} \frac{\partial x_i(t)}{\partial p_j} \right|_{x=x(t; p_{nom})} \quad (6.9)$$

Nevertheless, a parametric sensitivity analysis performed with traditional local methods are only valid with respect to a particular point in the parameter space, that is, in a neighborhoods of a certain parameter set. Thus, these coefficients only provide information on the robustness of a particular choice in the parameter space. Moreover, another significant limitation is due to the possibility to compute only the sensitivity with respect to the variations in a single parameter at a time; a model could display low sensitivity to such variations while being extremely sensitive to simultaneous variations in multiple parameters.

6.3.3 Bifurcation Analysis

Also by characterizing as rigorously as mathematical models for biological systems, it is not possible neither express nor measure the parameters model with exact precision. Nonlinear systems can exhibit multiple equilibria, each of which may be stable (either simply or asymptotically) or unstable, and the position of the equilibrium points, their stability properties and regions of attraction, can vary depending on the variations in the parametric space.

Definition 2 Given the system in (6.6), a bifurcation is represented as the interruption of its dynamical state determined by a structural modification, interpretable as a variation (as infinitesimal) of the parameter p_j , (bifurcation parameter), [47, 71].

Generally, two types of bifurcations can distinguished: local, which occur in the neighborhood of a single fixed point or limit cycle, and global, which change the whole structure of the phase space.

When the model has been defined, the behavior of the system is showed in respect to the parameter variations, identifying particular partitions of the parameter space in which dynamical discontinuities are found (for example the sets of all the limit points). The study of the parametric continuation deals with the continuation of the solution curves of a given dynamical system with respect to variations in one system parameter. The relative variations in the map of equilibrium points can be effectively visualized by using a ‘bifurcation diagram’, in which the equilibrium values of some state variable are plotted as a function of the bifurcation parameter. Consider the system (6.6), varying the parameter p_j in a particular range of interest and calculating the equilibrium points by solving the $f(x(t); p) = 0$, the diagram is constructed by evaluating the first solution for p_j , therefore continuing with the subsequently assignment of a parameter variation in the form $p_{j+1} = p_j + \Delta p_j$, with Δp_j properly settled. To see this, consider the bifurcation diagram for the nonlinear system in (6.1a, 6.1b) showed in Fig. 6.3. As aforementioned, this systems can alternates between two stable steady states, note that the location and the number of equilibrium points changes for different values of the parameters. The diagram shows the two bifurcation or limit points, LP, where the number of equilibrium points changes from one to three and then back to one, with respect to variations in k_{d1} . The solid lines represent the asymptotically stable equilibrium values, whereas the dashed line represents the unstable one. For intermediate values of k_{d1} , comprised between the two LP’s, the system preserves the bistability property, and it can reach the upper or lower branch of the diagram. Many bifurcation tools have been developed to perform various interesting dynamics such as switching and oscillatory dynamics in biomolecular networks, e.g., XPPAUT,² and the MATLAB package MATCONT.³

The bifurcation diagrams are powerful tools to investigate the qualitative changes in the behavior of nonlinear biomolecular systems due to parametric uncertainty.

² Available at <http://www.math.pitt.edu/~bard/xpp/xpp.html>.

³ Available at <http://www.matcont.ugent.be/matcont.html>.

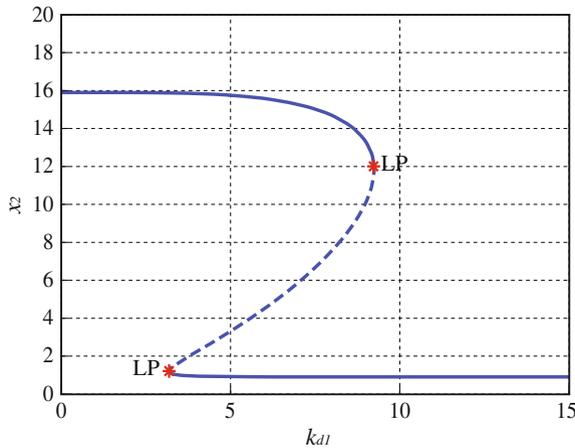


Fig. 6.3 Bifurcation diagram for the bistable system in (6.1a, 6.1b). The *solid lines* represent the asymptotically stable equilibrium values, whereas the *dashed line* represents the unstable one. For intermediate value of the bifurcation parameter k_{d1} the system is bistable and it can evolve to the upper or lower branch of the diagram, depending on whether the initial condition is above or below the *dashed* branch, respectively

However, it is necessary to take into account that (i) analytical solutions for bifurcations are only available for low-dimensional models, and that (ii) bifurcation diagrams are applicable only to study the effect of one or two parameters variation at a time.

6.3.4 Global Sensitivity Analysis

The previous approaches can be employed to obtain information on the robustness with respect to a particular region of the parameter space, and particular care must be taken in drawing any conclusions about global properties of the system under investigation. Therefore, it is more appropriate to investigate, the effects of simultaneous parameter variations of arbitrary magnitudes considering a global sensitivity analysis. Global parametric sensitivity analysis methods, as differential analysis, sampling-based methods using Monte-Carlo analysis, and variance decomposition methods, only recently, have been combined to biochemical network models [11, 19]. These methods must perform variations of all possible combinations of parameters for a preassigned range of possible parameter values, to analyze how the equilibrium points it can be affected. The global sensitivity analysis can also provides a guide not only in the design of synthetic biomolecular networks with specified functions highly robust, but also in the experimental realization of desired system behavior. For example, in [30] is shown an optimization technique of genetic networks based on the global analysis of sensitivity.

In [70], global sensitivity analysis for biochemical networks is approached via SDP to obtain reliable estimates of the boundaries of the regions for the given equilibria under parametric uncertainties. Given the admissible parameters variation range, the approach proposed by Waldherr et al. allows one to compute an outer approximation, $\tilde{\mathcal{S}}$, of the region of the state space that contains all the equilibrium points for the system, denoted by \mathcal{X}_e . The problem can be formalized by means of infeasibility certificates as follows.

Problem 2 Given system (6.6) and a box \mathcal{B}_p in the parameter space, compute a box $\tilde{\mathcal{S}}$ such that $\tilde{\mathcal{S}} \supseteq \mathcal{X}_e$, where

$$\mathcal{X}_e = \{x \in \mathbb{R}^n \mid \exists p \in \mathcal{B}_p : f(x(t); p) = 0\}. \quad (6.10)$$

The analytical calculation of \mathcal{X}_e for nontrivial system is practically impossible. Moreover, computational approaches are applicable only to very low-order systems. Monte-Carlo techniques can be applied in the other cases, although they may require a large computational effort and guarantee only probabilistic results. The method proposed by Waldherr et al. allows to effectively solve Problem 2 recasting it in the form of the following feasibility problem

$$\begin{cases} \text{find } x \in \mathbb{R}^n, p \in \mathbb{R}^m \\ \text{s.t. } f(x_i(t); p_j) = 0 \\ p_{j,\min} \leq p_j \leq p_{j,\max}, \quad j = 1, \dots, m \\ x_{i,\min} \leq x_i \leq x_{i,\max}, \quad i = 1, \dots, n. \end{cases} \quad (6.11)$$

The previous feasibility problem (6.11) requires a remarkable computational burden. However, it can be dealt with solving its dual version, thus by computing regions of the state space that are guaranteed not to contain any steady state for any parameter value in \mathcal{B}_p . Thus, it can be relaxed to a SDP problem [66], and solved by means of computationally efficient convex optimization tools. For a detailed description of this procedure, the reader is referred to [37, 70], in which the computation of a solution to Problem 2 constitutes the main goal of an iterative procedure, implemented by the bioSDP algorithm⁴: starting from an initial large region of the state space, the algorithm tries to compute one or more partitions containing \mathcal{X}_e . However, this procedure is very effective for low-order systems ($n \leq 3$), for which a bisection algorithm can be used. For systems of higher order, a box shrinkage procedure is employed, which can only return one partition $\tilde{\mathcal{S}}$ and, therefore, is not useful in the analysis of the bistability persistence. In order to solve this problem, we have proposed an alternative strategy, which combines the results of the DAs analysis with the bioSDP algorithm.

Based on the bioSDP algorithm, our approach provides to separately compute distinct robust steady state subsets (instead of computing one set containing all the equilibrium points), $\tilde{\mathcal{S}}_1$ and $\tilde{\mathcal{S}}_2$, which define the boundaries for the variation of the

⁴ Implemented in bioSDP toolbox available at <http://biosdp.sourceforge.net/>.

two attractors in the bistable system, x_{e_1} and x_{e_2} , respectively. Thus, setting $\tilde{\mathcal{F}}_i^0$ as the initial search space for the bioSDP algorithm, the analysis is focused on those equilibrium points that belong to a neighborhood of x_{e_i} , instead of searching for all the equilibrium points. Thus, we need two initial outer approximation subsets, let us denote them by $\tilde{\mathcal{F}}_1^0$ and $\tilde{\mathcal{F}}_2^0$, respectively. In general, for high-dimensional system, an appropriate estimation of two initial outer approximations is a daunting task. However, we surmise that, for small-enough variations of the parameter values, the DA represent good initial guesses. More precisely, considering $\tilde{\mathcal{F}}_i^0 = \varrho \mathcal{D}_i$, with $i = 1, 2$, where $\varrho > 0$ is a scaling factor, the initial guesses are enlarged versions of the measured DA. Performing this analysis separately on the two attractors, enables us to ascertain whether the bistability is preserved against parametric perturbations: the answer is affirmative if we are able to compute two disjoint robust steady state subsets, i.e., $\tilde{\mathcal{F}}_1 \cap \tilde{\mathcal{F}}_2$. If this problem is feasible for an assigned parameter box \mathcal{B}_p , then it is possible to suppose that the bistability is preserved for all the values of p_j belonging to \mathcal{B}_p .

In the following Sect. 6.4, we present a detailed analysis of the dynamical properties of the bistability in the *GAL* regulatory network in yeast, in order to validate it as a plausible quantitative characterization of the preservation of the bistable dynamics in biomolecular system.

6.4 A Procedure for Robustness Analysis of Bistability in the *GAL* Network

An interesting bistable mechanism has been identified in the *GAL* regulatory network in yeast *S. Cerevisiae*, that allows to regulate the metabolism of another sugar, galactose, when the organism perceives the presence of glucose in its environment. The bistable dynamics of the regulatory network of the galactose metabolic pathway, in the model organism *S. cerevisiae*, has been experimentally characterized [1]. History-dependent experiments on the galactose regulatory network, revealed the capability of this system to exhibit persistent memory of previous exposure to different carbon sources: under the same experimental conditions, cells previously grown with different nutrients (i.e., glucose or galactose) generate different responses and stabilize into two distinct steady states. From a system-theoretical point of view, this property can be explained by the switching between two stable steady states, based on galactose concentration.

The GAL regulatory system. The network of galactose metabolism in *S. Cerevisiae*, Fig. 6.4, is regulated by the following factors: a transcriptional activator protein Gal4p, a signal transducer protein Gal3p and an inhibitor protein Gal80p. In the presence of galactose, Gal4p activates transcription of genes *GAL1*, *GAL2*, *GAL3*, and *GAL80*. The protein encoded by gene *GAL2* acts as a carrier of external galactose into the yeast cell. In the absence of external galactose, Gal80p binds to the activation domain of Gal4p, thus inhibiting the expression of the *GAL* genes. In the presence of galactose, however, the inducer Gal3p binds to Gal80p, promoting the shuttling of Gal80p from the nucleus to the cytoplasm. This decreases the fraction of Gal80p-bound Gal4p in the nucleus. Thus, galactose relieves the inactivation of Gal4p and promotes transcription of the *GAL* genes [14].

6.4.1 Evidence of Bistability in the GAL System

The mathematical model of this regulatory network considered here is based on mass action kinetics and represents an extended version of the model proposed in [22].

$$\dot{G}_3 = k_8 G_4 - k_2 G_3 G_i + k_{r2} G_{3a} - \mu_{13} G_3 \quad (6.12a)$$

$$\dot{G}_i = k_1 G_{ex} G_2 - k_2 G_3 G_i + k_{r2} G_{3a} - \mu_{16} G_i \quad (6.12b)$$

$$\begin{aligned} \dot{G}_{3a} = & k_2 G_3 G_i - k_{r2} G_{3a} - k_4 G_{4,80} G_{3a} + k_{r4} G_{80,3a} G_4 - \mu_3 G_{3a} \\ & - k_{r19} G_{80} G_{3a} + k_{19} G_{80,3a} \end{aligned} \quad (6.12c)$$

$$\dot{G}_4 = k_5 - k_{11} G_4 G_{80} + k_{r11} G_{4,80} + k_4 G_{4,80} G_{3a} - k_{r4} G_{80,3a} G_4 - \mu_6 G_4 \quad (6.12d)$$

$$\dot{G}_{80} = -k_{11} G_4 G_{80} + k_{r11} G_{4,80} + k_7 G_4 - k_{r19} G_{80} G_{3a} + k_{19} G_{80,3a} - \mu_{14} G_{80} \quad (6.12e)$$

$$\dot{G}_{4,80} = k_{11} G_4 G_{80} - k_{r11} G_{4,80} - k_4 G_{4,80} G_{3a} + k_{r4} G_{80,3a} G_4 - \mu_{12} G_{4,80} \quad (6.12f)$$

$$\dot{G}_{80,3a} = k_4 G_{4,80} G_{3a} - k_{r4} G_{80,3a} G_4 - \mu_{15} G_{80,3a} - k_{19} G_{80,3a} + k_{r19} G_{80} G_{3a} \quad (6.12g)$$

$$\dot{G}_2 = k_9 G_4 - \mu_{17} G_2 \quad (6.12h)$$

$$\dot{G}_1 = k_{10} G_4 - \mu_{18} G_1. \quad (6.12i)$$

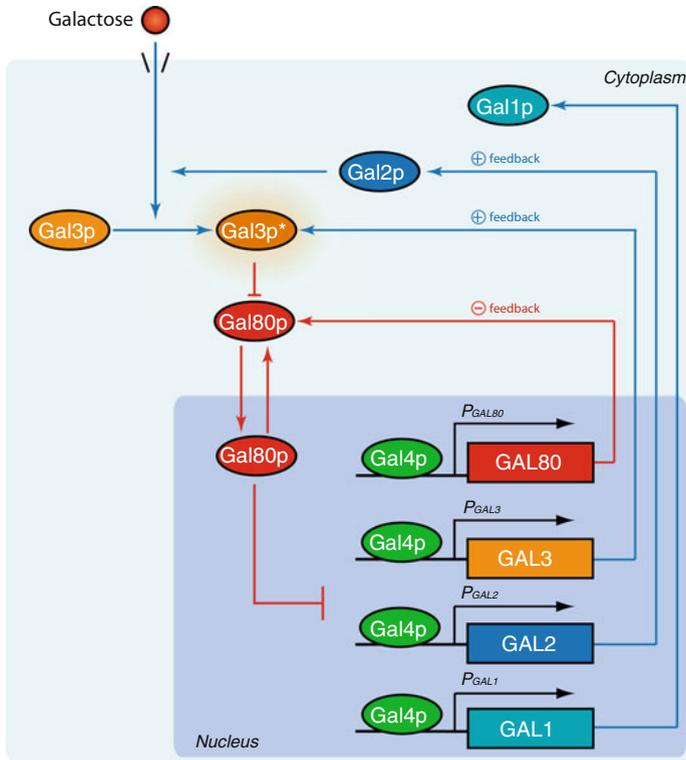


Fig. 6.4 Schematic diagram of the *GAL* regulatory network in *S. Cerevisia*. In absence of glucose, a molecule of galactose is imported from the external environment in the cytoplasm by *Gal2p*; the internalized galactose, in turn, activates *Gal3p* and, subsequently, sequesters *Gal80p* in the cytoplasm, shuttling *Gal80p* from the nucleus. The transcriptional activator *Gal4p*, which is constitutively bound to promoters of the *GAL* genes in the nucleus, is then released from the inhibitory action of *Gal80p* and activates expression of the *GAL1*, *GAL2*, *GAL3* and *GAL80* genes

See Table 6.1, for a detailed declaration of the state variables. The model (6.12a–6.12i) is able to reproduce bistability if there exists a particular set of parameters, $p \in \mathbb{R}_+^{23}$, for which the system admits two finite distinct equilibrium points $x_{e_1}, x_{e_2} \in \mathbb{R}^9$. By applying the CRNT algorithm (see 6.2.2.1), a possible choice of the parameter vector p and the associated equilibrium points x_{e_1} , and x_{e_2} can be computed. This analysis confirms that system (6.12a–6.12i) admits two asymptotically steady states, (reported in Table 6.1), when the kinetic parameters assume the values reported in Table 6.2. Note that, all the values have been opportunely scaled to achieve a good agreement with realistic dimensions [1, 58].

Table 6.1 Steady states of the mass-action model (6.12a–6.12i), with the parameters values given in Table 6.2

x_{e1}	Species	Description	x_{e2}
172.8212	G_3	<i>Gal3p protein</i>	2711.1839
172.8208	G_i	<i>Internalized galactose</i>	2711.2003
1.0	G_{3a}	<i>Active Gal3p protein</i>	318.5443
1.0	G_4	<i>Gal4p protein</i>	19.9479
1.0	G_{80}	<i>Gal80p protein</i>	0.3061
21.0604	$G_{4,80}$	<i>Gal4p:Gal80p complex</i>	2.1126
7.5945	$G_{80,3a}$	<i>Gal80p:Gal3p active complex</i>	589.1342
1.0	G_2	<i>Gal2p protein</i>	19.9479
1.0	G_1	<i>Gal1p protein</i>	19.9479
1000.0	G_{ex}	<i>External galactose</i>	1000.0

Values are given as (μM)

Table 6.2 A set of parameters values for which the system (6.12a–6.12i) is bistable

Parameter	Value	Parameter	Value
k_1	0.1814 ($\mu M \cdot h$)	k_{11}	85.8185 (1/h)
k_2	8.4586E-4 ($1/\mu M \cdot h$)	k_{r11}	4.7482E-2 ($1/\mu M \cdot h$)
k_{r2}	16.6691 (1/h)	μ_{12}	1.0 (1/h)
μ_3	1.0 (1/h)	μ_{13}	1.0 (1/h)
k_4	3.0749 ($1/\mu M \cdot h$)	μ_{14}	1.0 (1/h)
k_{r4}	0.1317 (1/h)	μ_{15}	1.0 (1/h)
k_5	22.0604 ($\mu M/h$)	μ_{16}	1.0 (1/h)
μ_6	1.0 (1/h)	μ_{17}	1.0 (1/h)
k_7	29.6549 (1/h)	μ_{18}	1.0 (1/h)
k_8	181.4157 (1/h)	k_{19}	36.6342 ($1/\mu M \cdot h$)
k_9	1.0 (1/h)	k_{r19}	222.0536 (1/h)
k_{10}	1.0 (1/h)		

6.4.2 Characterization of the DA

In general, a complete characterization of the dynamical properties of an attractor requires an estimation of its DA, to get valuable information about the ability of the system to adapt to exogenous perturbations without switching to a different operative condition (i.e., a different biological state). Applying the method proposed in [4] to find the best possible estimates of the DA, we have adopted an iterative approach, which starts from a small polytopic region, surrounding the equilibrium point; thus, this set is enlarged along one dimension at each iteration, stopping when the feasibility condition is no longer verified. The process is repeated many times, modifying at

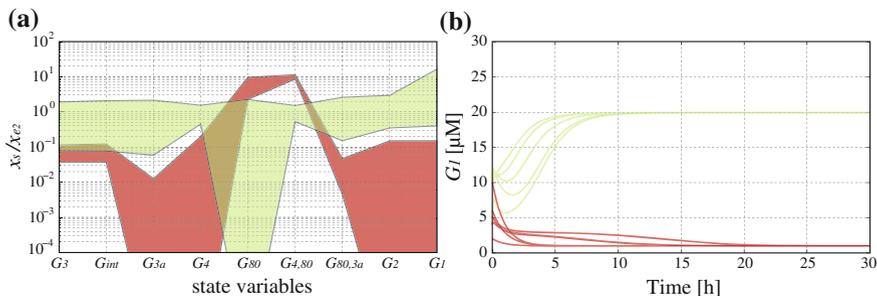


Fig. 6.5 **a** Hyperbox estimates for the DA of the system (6.12a–6.12i) plotted in normalized parallel coordinates. The *green* (*red*) hyperbox represents the DA for the high (low) equilibrium point, x_{e2} (x_{e1}). **b** Free evolutions, for different initial conditions, of the concentrations of Gal1p protein with the set of parameter given by application of CRNT toolbox. The curves funnel into low (*red*) or high (*green*), depending on initial conditions, confirming the bistable nature of the system

each run the sequential order of the dimensions employed in the enlargement steps. Two sets of admissible variations for the DA of x_{s1} , x_{s2} , \mathcal{D}_1 and \mathcal{D}_2 , respectively, have been found. The validity of estimated DA can be confirmed by numerical simulations, as reported in Fig. 6.5.

6.4.3 Local and Global Analysis Confirms the Robustness of Bistability

Parametric sensitivity analysis enables us to evaluate the influence of each parameter on the system. The relative local sensitivity of each species has been calculated with respect to the system parameters. Employing the method presented in [40], the differential equations describing the dynamics of our system are coupled to the differential equations of the sensitivity coefficients. The normalized sensitivity coefficients for model (6.12a–6.12i) are reported in Fig. 6.6: larger values are exhibited by the parameters involved in the feedback terms, k_7 , k_8 , k_9 , the basal expression of Gal4p, k_5 , those involved in the internalization of external galactose and in the activation of Gal3p protein, k_1 , k_2 , and the parameters that describe the consumption of Gal3p, internalized galactose and Gal2p proteins, μ_{13} , μ_{16} and μ_{17} , respectively. Note that the indications of robustness provided by the sensitivity coefficients is only valid locally, i.e., in the neighborhood of the nominal values reported in Table 6.2.

The parameters with high sensitivities are usually those leading to major qualitative changes in the dynamical behavior of the system. To investigate this point, we have performed a bifurcation analysis with respect to each of the above identified parameters. In the parameter space, the region of bistability for system (6.12a–6.12i) is delimited by a pair of limit points forming the classical S-shaped bifurcation curve, as shown in Fig. 6.7a: here one can readily identify the admissible range of variation

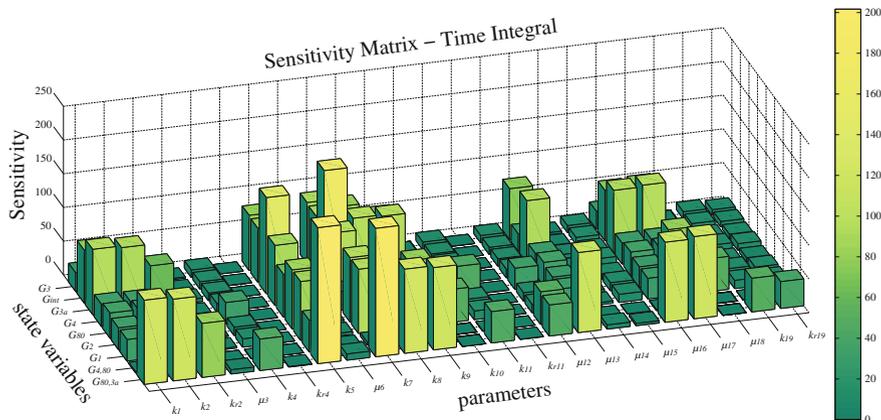


Fig. 6.6 Local sensitivity analysis at the steady states of the species involved with respect to all model parameters. The sensitivity matrix denotes the normalized sensitivity coefficients of all species in the model across 20h

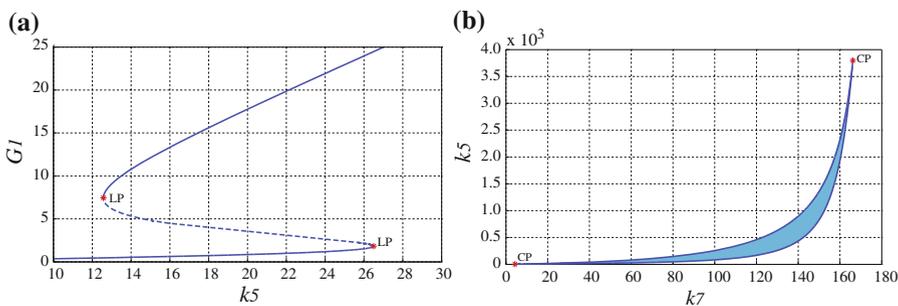


Fig. 6.7 **a** Continuation in one parameter. A branch of equilibria in the (k_5, G_1) -plane displaying bistability. *LP* Limit Point. **b** Continuation in two parameters. Bifurcation curves in the (k_7, k_5) -plane with codimension 2 points. At $(k_7, k_5) = (4.363, 3.835)$ and $(k_7, k_5) = (166.2, 3797.0)$ two cusp points are detected

of k_5 such that bistability is preserved. The bifurcation analysis can also be performed considering two-parameters variations: in this case, the bifurcation boundaries, corresponding to the limit point bifurcations, are curves in the parameters plane. For example, in Fig. 6.7b two bifurcation cusp points (CP) are shown. Any set of parameter values of the pair (k_7, k_5) included in the shadowed region guarantees bistable behavior.

Bifurcation analysis is an efficient tool to locally evaluate the changes in steady state concentrations with respect to one or two simultaneous parameters variations. In order to evaluate the preservation of bistability with respect to simultaneous variation of multiple parameters, we have applied the method proposed in [70]. By partitioning the state space into two subsets, each containing one of the nominal

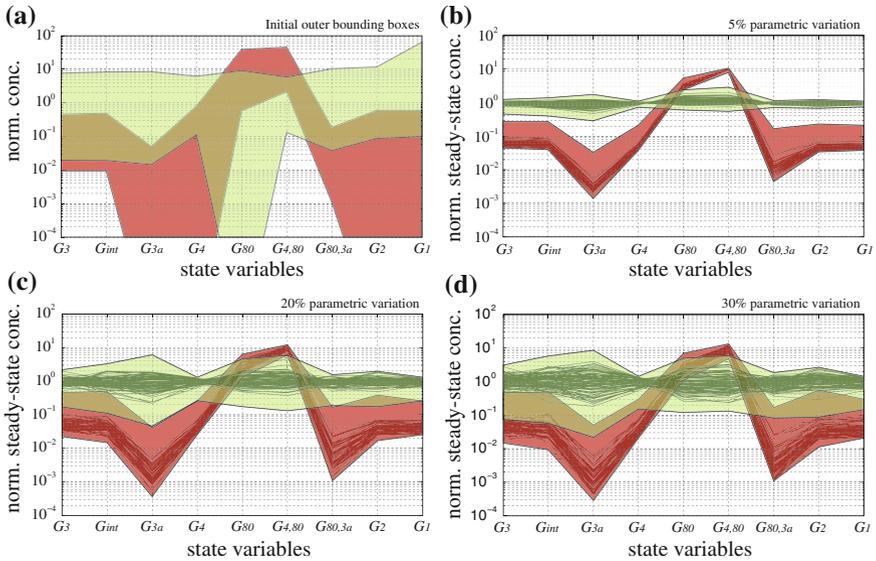


Fig. 6.8 **a** Initial approximated partitions, computed by scaling the \mathcal{D}_1 and \mathcal{D}_2 and reported in parallel coordinates normalized to the high-equilibrium, x_{e_2} . **b–d** Steady state uncertainty sets for *GAL* system plotted in parallel coordinates, in $\mathcal{P}_i \subset \mathbb{R}^9$ regions. Each line represents sampled steady states, under the parametric uncertainty for the lower (in red) and higher (in green) steady state, respectively

equilibrium points, x_{s_1} , x_{s_2} , such method allows to compute certificates for regions in the state space non containing any steady state, for some assigned ranges of variation of the parameters. The initial approximated partitions $\tilde{\mathcal{S}}_{1,2}$ have been computed by scaling the \mathcal{D}_1 and \mathcal{D}_2 as reported in normalized parallel coordinates in Fig. 6.8a. To reduce the computational burden required by the application of this method, the multi-parametric sensitivity analysis has been limited to the parameters subset $\Omega_p := \{k_1, k_2, k_5, k_7, k_8, k_9, \mu_{13}, \mu_{16}, \mu_{17}\}$, which, according to the local sensitivity analysis, are those exerting the greatest influence on the system dynamics, see Fig. 6.6. The global parametric robustness has been evaluated for increasingly larger ranges of parametric uncertainty, with respect to the nominal values given in Table 6.2. Figure 6.8b–d displays the computed robust steady state boxes for different cases. For parametric variations up to $\pm 20\%$ with respect to the nominal parameters value, the bistable behavior of the *GAL* regulatory network is guaranteed. For such uncertainty values, indeed, the computed subsets $\tilde{\mathcal{S}}_1$ and $\tilde{\mathcal{S}}_2$ are still disjoint, since the intervals of G_{3a} and $G_{80,3a}$ are not overlapping. For larger parametric variations the intersection of the two subsets is no longer empty (see Fig. 6.8d); hence, in the latter case it is no longer possible to guarantee that the system preserves bistability for all admissible values of the uncertain parameters.

The proposed procedure provides a powerful approach to the analysis and validation of any biochemical network model that is required to robustly reproduce bistable dynamics, underlying persistent memory, molecular switches and cell differentiation phenomena, without recourse to large-scale computational simulation.

References

1. Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435(7039):228–232. doi:[10.1038/nature03524](https://doi.org/10.1038/nature03524)
2. Ackers GK, Johnson AD, Shea MA (1982) Quantitative model for gene regulation by lambda phage repressor. *Proc Natl Acad Sci USA* 79(4):1129–1133
3. Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* 8(6):450–461
4. Amato F, Cosentino C, Merola A (2007) On the region of attraction of nonlinear quadratic systems. *Automatica* 43(12):2119–2123
5. Avery SV (2006) Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol* 4(8):577–587. doi:[10.1038/nrmicro1460](https://doi.org/10.1038/nrmicro1460)
6. Bagowski CP, Ferrell J Jr (2001) Bistability in the JNK cascade. *Curr Biol* 11(15):1176–1182
7. Bagowski CP, Besser J, Frey CR, Ferrell JE Jr (2003) The JNK cascade as a biochemical switch in mammalian cells: ultrasensitive and all-or-none responses. *Curr Biol* 13(4):315–320
8. Balas G, Chiang R, Packard A, Safonov M (2009) MATLAB robust control toolbox 3. Users guide. Natick, MA
9. Barabási A, Oltvai Z (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5(2):101–113
10. Becskei A, Séraphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J* 20(10):2528–2535. doi:[10.1093/emboj/20.10.2528](https://doi.org/10.1093/emboj/20.10.2528)
11. Bentele M, Lavrik I, Ulrich M, Stösser S, Heermann D, Kalthoff H, Kramer P, Eils R (2004) Mathematical modeling reveals threshold mechanism in cd95-induced apoptosis. *J Cell Biol* 166(6):839–851
12. Bhalla US, Iyengar R (1999) Emergent properties of networks of biological signaling pathways. *Science* 283(5400):381–387
13. Bhartiya S, Chaudhary N, Venkatesh KV, Doyle FJ 3rd (2006) Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series. *J R Soc Interface* 3(8):383–391. doi:[10.1098/rsif.2005.0103](https://doi.org/10.1098/rsif.2005.0103)
14. Bhat P (2008) Galactose regulon of yeast: from genetics to systems biology. Springer, New York
15. Blüthgen N, Herzog H (2003) How robust are switches in intracellular signaling cascades? *J Theor Biol* 225(3):293–300
16. Boyd S, El Ghaoul L, Feron E, Balakrishnan V (1987) Linear matrix inequalities in system and control theory, vol 15. Society for Industrial Mathematics
17. Chesi G (2004) On the estimation of the domain of attraction for uncertain polynomial systems via LMIs. In: Decision and control, 2004. CDC. 43rd IEEE conference on, IEEE, vol 1, pp 881–886
18. Chesi G (2011) Domain of attraction: analysis and control via SOS programming. Springer, New York
19. Cho K, Shin S, Kolch W, Wolkenhauer O (2003) Experimental design in systems biology, based on parameter sensitivity analysis using a monte carlo method: a case study for the tnfr α -mediated nf- κ b signal transduction pathway. *Simulation* 79(12):726–739
20. Chou CS, Nie Q, Yi TM (2008) Modeling robustness tradeoffs in yeast cell polarization induced by spatial gradients. *PLoS One* 3(9):e3103. doi:[10.1371/journal.pone.0003103](https://doi.org/10.1371/journal.pone.0003103)

21. Cosentino C, Bates D (2011) Feedback control in systems biology. CRC Press, Boca Raton
22. Cosentino C, Salerno L, Passanti A, Merola A, Bates DG, Amato F (2012) Structural bistability of the gal regulatory network and characterization of its domains of attraction. *J Comput Biol* 19(2):148–162. doi:[10.1089/cmb.2011.0251](https://doi.org/10.1089/cmb.2011.0251)
23. Delbrück M (1948) Biochemical mutants of bacterial viruses. *J Bacteriol* 56(1):1–16
24. Doyle J, Csete M (2005) Motifs, control, and stability. *PLoS Biol* 3(11):e392. doi:[10.1371/journal.pbio.0030392](https://doi.org/10.1371/journal.pbio.0030392)
25. Dubnau D, Losick R (2006) Bistability in bacteria. *Mol Microbiol* 61(3):564–572. doi:[10.1111/j.1365-2958.2006.05249.x](https://doi.org/10.1111/j.1365-2958.2006.05249.x)
26. Eissing T, Allgöwer F, Bullinger E (2005) Robustness properties of apoptosis models with respect to parameter variations and intrinsic noise. *Syst Biol (Stevenage)* 152(4):221–228
27. Feinberg M (1972) Complex balancing in general kinetic systems. *Arch Ration Mech Anal* 49(3):187–194
28. Feinberg M (1987) Chemical reaction network structure and the stability of complex isothermal reactors-i. The deficiency zero and deficiency one theorems. *Chem Eng Sci* 42(10):2229–2268
29. Feinberg M (1988) Chemical reaction network structure and the stability of complex isothermal reactors-ii. Multiple steady states for networks of deficiency one. *Chem Eng Sci* 43(1):1–25
30. Feng XJ, Hooshangi S, Chen D, Li G, Weiss R, Rabitz H (2004) Optimizing genetic circuits by global sensitivity analysis. *Biophys J* 87(4):2195–2202. doi:[10.1529/biophysj.104.044131](https://doi.org/10.1529/biophysj.104.044131)
31. Ferrell J Jr, Machleder EM (1998) The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Science* 280(5365):895–898
32. Ferrell JE Jr (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* 14(2):140–148
33. Freeman M (2000) Feedback control of intercellular signalling in development. *Nature* 408(6810):313–319. doi:[10.1038/35042500](https://doi.org/10.1038/35042500)
34. Gardner T, Cantor C, Collins J (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342
35. Genesio R, Tartaglia M, Vicino A (1985) On the estimation of asymptotic stability regions: state of the art and new proposals. *IEEE Trans Autom Control* 30(8):747–755
36. Gunawardena J (2010) Models in systems biology: the parameter problem and the meanings of robustness. In: Elements of computational systems biology. Wiley, New York, pp 19–47
37. Hasenauer J, Rumschinski P, Waldherr S, Borchers S, Allgöwer F, Findeisen R (2010) Guaranteed steady state bounds for uncertain (bio-) chemical processes using infeasibility certificates. *J Process Control* 20(9):1076–1083
38. Horn F (1972) Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch Ration Mech Anal* 49(3):172–186
39. Huang S, Eichler G, Bar-Yam Y, Ingber DE (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys Rev Lett* 94(12):128, 701
40. Ingalls BP, Sauro HM (2003) Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories. *J Theor Biol* 222(1):23–36
41. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–356
42. Jin Y, Sendhoff B (2008) Evolving in silico bistable and oscillatory dynamics for gene regulatory network motifs. In: Evolutionary computation, 2008. CEC, 2008. IEEE World Congress on computational intelligence. IEEE Congress on, IEEE, pp 386–391
43. Khalil H, Grizzle J (1996) Nonlinear systems, vol 3. Prentice Hall, Englewood Cliffs
44. Kitano H (2004) Biological robustness. *Nat Rev Genet* 5(11):826–837. doi:[10.1038/nrg1471](https://doi.org/10.1038/nrg1471)
45. Kitano H (2007) Towards a theory of biological robustness. *Mol Syst Biol* 3:137. doi:[10.1038/msb4100179](https://doi.org/10.1038/msb4100179)
46. Kurata H, El-Samad H, Iwasaki R, Ohtake H, Doyle JC, Grigorova I, Gross CA, Khammash M (2006) Module-based analysis of robustness tradeoffs in the heat shock response system. *PLoS Comput Biol* 2(7):e59. doi:[10.1371/journal.pcbi.0020059](https://doi.org/10.1371/journal.pcbi.0020059)
47. Kuznetsov I (1998) Elements of applied bifurcation theory, vol 112. Springer, New York

48. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee HJ, Sciammas R, Gantner BN, Diner AR, Singh H (2006) Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* 126(4):755–766. doi:[10.1016/j.cell.2006.06.052](https://doi.org/10.1016/j.cell.2006.06.052)
49. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002) Transcriptional regulatory networks in *saccharomyces cerevisiae*. *Science* 298(5594):799–804. doi:[10.1126/science.1075090](https://doi.org/10.1126/science.1075090)
50. Ma L, Iglesias PA (2002) Quantifying robustness of biochemical network models. *BMC Bioinform* 3:38
51. Maamar H, Dubnau D (2005) Bistability in the *bacillus subtilis* k-state (competence) system requires a positive feedback loop. *Mol Microbiol* 56(3):615–624. doi:[10.1111/j.1365-2958.2005.04592.x](https://doi.org/10.1111/j.1365-2958.2005.04592.x)
52. Müller-Hill B (1998) Some repressors of bacterial transcription. *Curr Opin Microbiol* 1(2):145–151
53. Monod J, Jacob F (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* 26:389–401
54. Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* 12:88–118
55. Newman M, Barabasi A, Watts D (2011) *The structure and dynamics of networks*. Princeton University Press, Princeton
56. Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci USA* 43(7):553–566
57. Paliwal S, Iglesias PA, Campbell K, Hilioti Z, Groisman A, Levchenko A (2007) MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* 446(7131):46–51. doi:[10.1038/nature05561](https://doi.org/10.1038/nature05561)
58. Ruhela A, Verma M, Edwards JS, Bhat PJ, Bhartiya S, Venkatesh KV (2004) Autoregulation of regulatory proteins is key for dynamic operation of gal switch in *saccharomyces cerevisiae*. *FEBS Lett* 576(1–2):119–126. doi:[10.1016/j.febslet.2004.09.001](https://doi.org/10.1016/j.febslet.2004.09.001)
59. Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31(1):64–68. doi:[10.1038/ng881](https://doi.org/10.1038/ng881)
60. Smits WK, Kuipers OP, Veening JW (2006) Phenotypic variation in bacteria: the role of feedback regulation. *Nat Rev Microbiol* 4(4):259–271. doi:[10.1038/nrmicro1381](https://doi.org/10.1038/nrmicro1381)
61. Stamatakis M, Zygorakis K (2011) Deterministic and stochastic population-level simulations of an artificial lac operon genetic network. *BMC Bioinform* 12:301. doi:[10.1186/1471-2105-12-301](https://doi.org/10.1186/1471-2105-12-301)
62. Stelling J, Sauer U, Szallasi Z, Doyle FJ 3rd, Doyle J (2004) Robustness of cellular functions. *Cell* 118(6):675–685. doi:[10.1016/j.cell.2004.09.008](https://doi.org/10.1016/j.cell.2004.09.008)
63. Strogatz S (1994) *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering*. Westview Press, Perseus Books, (Cambridge)
64. Thomas R (1978) Logical analysis of systems comprising feedback loops. *J Theor Biol* 73(4):631–656
65. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* 15(2):221–231
66. Vandenberghe L, Boyd S (1996) Semidefinite programming. *SIAM Rev* 38(1):49–95
67. Vannelli A, Vidyasagar M (1985) Maximal lyapunov functions and domains of attraction for autonomous nonlinear systems. *Automatica* 21(1):69–80
68. Veening JW, Hamoen LW, Kuipers OP (2005) Phosphatases modulate the bistable sporulation gene expression pattern in *bacillus subtilis*. *Mol Microbiol* 56(6):1481–1494. doi:[10.1111/j.1365-2958.2005.04659.x](https://doi.org/10.1111/j.1365-2958.2005.04659.x)
69. Waddington C et al (1957) *The strategy of the genes. A discussion of some aspects of theoretical biology. with an appendix by h. kacser. The strategy of the genes A discussion of some aspects of theoretical biology With an appendix by H. Kacser pp ix+-262*

70. Waldherr S, Findeisen R, Allgöwer F (2008) Global sensitivity analysis of biochemical reaction networks via semidefinite programming. In: IFAC proceedings volumes (IFAP-papers online), 17th World Congress, international federation of automatic control, IFAC, vol 17
71. Wiggins S (2003) Introduction to applied nonlinear dynamical systems and chaos, vol 2. Springer, New York
72. Wykoff DD, Rizvi AH, Raser JM, Margolin B, O'Shea EK (2007) Positive feedback regulates switching of phosphate transporters in *s. cerevisiae*. *Mol Cell* 27(6):1005–1013. doi:[10.1016/j.molcel.2007.07.022](https://doi.org/10.1016/j.molcel.2007.07.022)
73. Xiong W, Ferrell JE Jr (2003) A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* 426(6965):460–465. doi:[10.1038/nature02089](https://doi.org/10.1038/nature02089)

Part II
Synthetic Biology

Chapter 7

Meta-DNA: A DNA-Based Approach to Synthetic Biology

Harish Chandran, Nikhil Gopalkrishnan,
Bernard Yurke and John Reif

Abstract The goal of synthetic biology is to design and assemble synthetic systems that mimic biological systems. One of the most fundamental challenges in synthetic biology is to synthesize artificial biochemical systems, which we will call meta-biochemical systems, that provide the same functionality as biological nucleic acids-enzyme systems, but that use a very limited number of types of molecules. The motivation for developing such synthetic biology systems is to enable a better understanding of the basic processes of natural biology, and also to enable re-engineering and programmability of synthetic versions of biological systems. One of the key aspects of modern nucleic acid biochemistry is its extensive use of protein enzymes that were originally evolved in cells to manipulate nucleic acids, and then later adapted by man for laboratory use. This practice provided powerful tools for manipulating nucleic acids, but also limited the extent of the programmability of the available chemistry for manipulating nucleic acids, since it is very

This chapter is an extensively revised version of Harish Chandran, Nikhil Gopalkrishnan, Bernard Yurke, John Reif, Meta-DNA: Synthetic Biology via DNA Nanostructures and Hybridization Reactions, Journal of the Royal Society Interface, (2012). Published online 11 January 2012, pp. 1742–5662.

H. Chandran (✉) · N. Gopalkrishnan · J. Reif
Department of Computer Science, Duke University, Durham, NC, USA
e-mail: harish@cs.duke.edu

N. Gopalkrishnan
e-mail: nikhil@cs.duke.edu

B. Yurke
Department of Materials Science and Engineering and Department of Electrical and Computer Engineering, Boise State University, Boise, ID, USA
e-mail: bernardyrurke@boisestate.edu

J. Reif
Adjunct Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
e-mail: reif@cs.duke.edu

difficult to predictively modify the behavior of protein enzymes. Meta-biochemical systems offer the possible advantage of being far easier to re-engineer and program for desired functionality. The approach taken here is to develop a biochemical system which we call meta-DNA (abbreviated as mDNA), based entirely on strands of DNA as the only component molecules. Our work leverages prior work on the development of self-assembled DNA nanostructures (see [1, 2, 5, 9, 11, 18, 26] for excellent reviews of the field). Each base of a mDNA is a DNA nanostructure. Our mDNA bases are paired similar to DNA bases, but have a much larger alphabet of bases, thereby providing increased power of base addressability. Our mDNA bases can be assembled to form flexible linear assemblies (single stranded mDNA) analogous to single stranded DNA, and can be hybridized to form stiff helical structures (duplex mDNA) analogous to double stranded DNA, and also can be denatured back to single stranded mDNA. Our work also leverages the abstract activatable tile model developed by Majumder et al. [12] and prior work on the development of enzyme-free isothermal protocols based on DNA hybridization and sophisticated strand displacement hybridization reactions (see [6, 15, 16, 19, 21, 27, 28]). We describe various isothermal hybridization reactions that manipulate our mDNA in powerful ways analogous to DNA–DNA reactions and the action of various enzymes on DNA. These operations on mDNA include (i) hybridization of single strand mDNA (*ssmDNA*) into a double strand mDNA (*dsmDNA*) and heat denaturation of a *dsmDNA* into its component *ssmDNA* (analogous to DNA hybridization and denaturation), (ii) strand displacement of one *ssmDNA* by another (similar to strand displacement in DNA), (iii) restriction cuts on the backbones of *ssmDNA* and *dsmDNA* (similar to the action of restriction enzymes on DNA), (iv) polymerization chain reactions that extend *ssmDNA* on a template to form a complete *dsmDNA* (similar to the action of polymerase enzyme on DNA), (v) isothermal denaturation of a *dsmDNA* into its component *ssmDNA* (like the activity of helicase enzyme on DNA) and (vi) an isothermal replicator reaction which exponentially amplifies *ssmDNA* strands (similar to the isothermal PCR reaction). We provide a formal model to describe the required properties and operations of our mDNA, and show that our proposed DNA nanostructures and hybridization reactions provide these properties and functionality.

Keywords Enzyme · DNA · Meta-DNA (mDNA) · Isothermal · Hybridization · Strand displacement · Single strand meta-DNA (*ssmDNA*) · Double strand meta-DNA (*dsmDNA*)

7.1 Introduction

7.1.1 Synthetic Biology

A major goal of synthetic biology is to produce synthetic biochemical systems which have functions similar to the biochemical functions of living organisms. Considerable related work has been done in the fields of artificial and synthetic life, see [3, 10, 13]

for an overview of these fields. Two predominant approaches by researchers in this area have been to design protein-based or RNA-based biochemical systems. However, both proteins and RNA are very difficult to predictively design for given functions, and behavior of the resulting protein-based or RNA-based biochemical systems can be very complex, making their engineering highly challenging. An alternative approach we propose here is to produce synthetic biochemical systems based on a very well understood, relatively non-reactive molecule, such as DNA (DNA strands and DNA nanostructures). DNA–DNA interactions like hybridization and strand displacement are reasonably well understood and a vast literature exists that studies, models, predicts and even controls such interactions (see [8, 29]). This literature enables us to program simple DNA systems by controlling experimental conditions like pH, salt concentrations, temperature and DNA concentration.

7.1.2 Overview of Our Work on Meta-DNA

In this work, we engineer synthetic biochemical systems, called meta-DNA, consisting only of DNA that capture the properties and structure of DNA in biological systems. Our work is reductive: we use simple DNA chemistry to emulate more complex enzyme based DNA chemistry through ingenious use of DNA hybridization, strand displacement and hairpin systems. From a computer science perspective, our work can be thought of as using a lower level programming language to simulate programs encoded in a higher level programming language. This kind of simulation often leads to a better abstract understanding of the phenomena being simulated. From a synthetic biology perspective, we are building novel biochemical systems to emulate useful, well known natural biological systems and providing alternatives to protein enzymes. From an engineering perspective, our work is a minimalist approach to designing biochemical systems from simple, predictable yet powerful modules. Our systems are largely isothermal and autonomous which suggest that they may have applications for in vitro biochemical systems like transport devices, molecular motors, detection, signalling and computing systems.

7.1.3 Prior Work on Synthetic Replicators

Driven by the important role that replication plays in Biology, many self-replicating systems have been proposed, starting with Von Neumann. Early self-replicating systems were designed by von Kiedrowski [24], Tjivikua et al. [22]. For a review of various artificial replicators see [4, 23, 25]. A DNA-based artificial replicator was proposed by Zhang and Yurke [30]. Schulman and Winfree [17] study growth and evolution of simple crystals using DNA. Smith et al. [20] have independently developed abstractions for self-replication systems that can be thought of as tile-based and also rely on the idea of activation.

The major departure from this prior work on self-replicating systems and our meta-DNA approach, is that meta-DNA allows for a wide variety of key operations far beyond merely the operation of replication, enabling a much more extensive and complex set of synthetic biochemical systems.

7.1.4 Organization of this Chapter

In Sect. 7.2 we review the key properties of DNA, including its structure and enzymatic reactions. In Sect. 7.3 we list the desirable properties of meta-DNA and give a design for a meta-nucleotide, single and double stranded meta-DNA, along with its secondary structure. Section 7.4 states the assumptions we make while designing protocols for meta-DNA manipulation. In Sect. 7.5 we discuss various meta-DNA reactions such as meta-hybridization, meta-denaturation, meta-strand displacement, meta-polymerization, meta-restriction, meta-helicase denaturation and exponential amplification using a meta-DNA replicator. Section 7.6 discusses open problems and future work.

7.2 Review of DNA Structure and Reactions

DNA is considered to be an ideal material to construct nanoscale structures and devices and has been used as scaffolding material for complex shapes, fuel for molecular motors and aptamers for various organic and inorganic molecules. The key properties of DNA that enable these functionalities are programmability, predictable chemical interaction and secondary structure and simple laboratory protocols for its manipulation. Synthetic DNA is also cheaply and readily available from a variety of commercial sources. At the most abstract level, fabricating structures and devices with DNA is akin to working with smart bricks that fit together in a specific pre-defined way and then putting them in a bag, shaking it and waiting for the bricks to self-assemble. However, the process is probabilistic and correcting errors is a fundamental challenge for the field.

Most readers will be familiar with the basic structure and reactions of DNA. The following discussion will be useful to even these readers as it frames DNA and its reactions in the context of our work and also specifies the granularity at which we consider DNA, with particular focus on the biochemical properties that we wish to simulate using our synthetic mDNA systems.

7.2.1 Key Properties of DNA

Here we review the key properties of natural DNA that we wish to mimic in mDNA. Single stranded DNA is a linear polymer made from repeating units called *nucleotides*. The nucleotide repeats contain both a segment of the *backbone* of the molecule, which holds the chain together, and a *base*. Each nucleotide has directionality and hence a DNA strand is asymmetric. The asymmetric ends of DNA strands are called the 5' (five prime) and 3' (three prime) ends. Single stranded DNA is flexible and has a small persistence length and is generally modeled as a freely-jointed chain. In living organisms, DNA does not usually exist as a single molecule, but instead as a pair of molecules entwined like vines, in the shape of a double helix. In a double helix the direction of the nucleotides in one strand is opposite to their direction in the other strand: the strands are *antiparallel*. Double stranded DNA has much greater persistence length than single stranded DNA and is generally modeled as a worm-like chain. The DNA double helix is stabilized by hydrogen bonds between the bases attached to the two strands. The four bases are classified into two types, purines (A and G) and pyrimidines (T and C). Each type of base on one strand overwhelmingly prefers a bond with just one type of base on the other strand. This is called *complementary base pairing*. Here, purines form hydrogen bonds to pyrimidines, with A bonding preferentially to T, and C bonding preferentially to G. This arrangement of two nucleotides binding together across the double helix is called a *base pair*. As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can therefore be pulled apart like a zipper, either by a mechanical force or kinetic energy due to high temperature. Given a pH value and salt (Mg^{++} , Na^+) concentrations, the *melting temperature* is defined as the temperature at which half the DNA strands are in a double helical conformation while the rest are denatured. Melting temperature depends on the length of the DNA strand as well as its sequence, with longer strands having a higher melting temperature.

7.2.2 Key Reactions Involving DNA

We view DNA reactions as biochemical programs and attempt to emulate them. Hence, it is important to understand and abstract out these DNA reaction properties. With this goal in mind, we review a few key DNA reactions.

7.2.2.1 DNA–DNA Reactions

- *Hybridization* is the attachment of a pair of single strands of DNA via hydrogen bonds along their complementary regions to form a double helix. Note that the two

strands only attach in an anti-parallel manner and hence are *reverse complements* of each other.

- The reverse process, separation of a double helix into its constituent single strands by the breaking of hydrogen bonds is *DNA denaturation*. This can be achieved via mechanical shear forces, high temperature or presence of denaturing agents like urea or formamide.
- *DNA strand displacement* is the displacement of a single strand of DNA from a double helix by an incoming strand with a longer complementary region to the template strand. The incoming strand has a *toehold*, an empty single stranded region on the template strand complementary to a subsequence of the incoming strand, to which it binds initially. It eventually displaces the outgoing strand via a kinetic process modeled as a one dimensional random walk.

7.2.2.2 Enzymatic Reactions on DNA

- *DNA restriction* is the cleaving of the backbone at a sequence specific recognition site by a restriction enzyme.
- *DNA polymerases* are a class of enzymes that catalyze the *polymerization* of nucleotides into a DNA strand. The polymerase “reads” an intact DNA strand as a template and uses it to synthesize the new strand. The newly polymerized molecule is complementary to the template strand. DNA polymerases lack the ability to do de novo polymerization and can only extend a DNA strand already attached to the template strand, called a *primer*.
- Isothermal denaturation can be achieved by *helicases* which are motor proteins that move directionally along a DNA backbone, denaturing the double helix.
- *DNA ligation* is the rejoining of nicked double stranded DNA by repairing the backbone break. Enzymes belonging to this class are known as *ligases*.

7.3 Desired Properties of Meta-DNA

We wish to abstract the structure and reactions of DNA described in Sect. 7.2.2 and emulate them using only DNA–DNA interactions. In doing so, we would have circumvented DNA-enzyme chemistry with a synthetic biochemical system that uses only DNA hybridization. First we state the desirable properties of meta-DNA (mDNA).

7.3.1 List of Desirable Properties for mDNA

- We desire a set of $2k$ meta-nucleotides (where k is an integer ≥ 2). Each meta-nucleotide must be directional and have a 3' and 5' end. There must be two

types of meta-nucleotides: meta-purines and meta-pyrimidines, each k in number. There must exist a pairing between meta-purines and meta-pyrimidines where each edge indicates an overwhelming complementary preference of meta-base linking between the meta-purine and meta-pyrimidine it connects.

- The active 5' end of any meta-nucleotide should be able to bind to the active 3' end of any other meta-nucleotide and vice versa via a meta-backbone link.
- The meta-backbone link must be a *strong* bond (implemented as a long sequence of hybridized DNA) while the meta-base link must be a *weak* bond (implemented as a short sequence of hybridized DNA).
- *ssmDNA* should be a linear directional polymer chain of meta-nucleotides bound by meta-backbone linkages and has 3' and 5' ends.
- Two *ssmDNA* that are reverse complementary to each other must have the ability to meta-hybridize to form a *dsmDNA*. The strands of any *dsmDNA* must be anti-parallel.
- When an ensemble of identical *dsmDNA* is heated to a temperature known as its melting temperature, half of the ensemble must denature into its constituent *ssmDNA*. The melting temperature must depend on the sequence of meta-nucleotides of the *mDNA* and also on its length, with longer strands having a higher melting temperature.
- To support strand displacement protocols, the weak meta-base bonds must be continuously broken and remade in a kinetic process called breathing. The rate of breathing should be positively correlated with temperature.
- *ssmDNA* must be flexible and have a small persistence length when compared to *dsmDNA* of comparable length. *ssmDNA* must be like a freely-jointed chain while *dsmDNA* must be like a worm-like chain.
- Meta-nucleotides should not spontaneously form meta-backbone bonds in the absence of catalyst.

7.3.2 Abstract Description of *mDNA*

We model a meta-nucleotide as an activatable tile [12] having three activatable pads: a 5' pad, a 3' pad and a base pad and represent it by a square tile as illustrated in Fig. 7.1. The tile has directionality as indicated by an arrow from 5' to 3' which is imposed by the sequence in which the pads are activated, with 5' always activated before 3'. Tiles bind to each other via symmetric pad interactions called *binding* or *linking*. Each binding has a *strength* associated with it (1, 2 or 3) that depends only on the type of pads involved in the binding. The strength of a binding models the energy required to break the bond. Base pads can only interact with other base pads through strength 1 bindings, and are called *meta-base bindings*. 5' pads can only interact with 3' pads through strength 3 bindings and these are called *meta-backbone bindings*. The pads exist in one of four states: inactive, active unbound, bound and capped. Inactive pads do not bind with other pads. Any active unbound 5' pad of a meta-nucleotide can bind to any active unbound 3' pad of another meta-nucleotide via a bond of strength

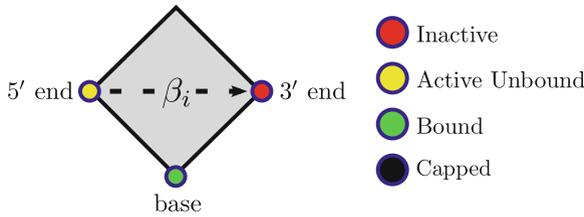


Fig. 7.1 Abstraction of the nucleotide



Fig. 7.2 Single and double stranded mDNA: **a** single stranded mDNA, **b** double stranded mDNA, **c** double stranded mDNA has a longer persistence length

3 after which these pads go to the bound state. There are $2k$ different base pads, each corresponding to a meta-nucleotide type. These are split into two complementary perfectly matched sets. Let $\mathfrak{P} = \{\beta_1, \beta_2, \dots, \beta_k, \bar{\beta}_1, \bar{\beta}_2, \dots, \bar{\beta}_k\}$ be the set of pads. For all i , pads β_i and $\bar{\beta}_i$ bind with each other. The face label on any tile indicates the base type of the meta-nucleotide. Any active unbound base pad of a meta-nucleotide can bind in an antiparallel manner to an active unbound complementary base pad of another meta-nucleotide by a strength 1 bond after which these pads go to the bound state.

A linear chain of these tiles held together by 5′–3′ pad bindings forms a directional polymer of meta-nucleotides and hence is *ssmDNA*. The 3′ base pad of the tile at the 3′ end of the *ssmDNA* and the 5′ base pad of the tile at the 5′ end are always in the capped state. A double stranded mDNA (*dsmDNA*) is a dimer of two *ssmDNA* held together by base pad bindings. *ssmDNA* is flexible with a low persistence length while *dsmDNA* is stiff and has a higher persistence length. These properties arise out of the geometric constraints imposed by the tile base structure of mDNA as illustrated in Fig. 7.2.

7.3.3 Strand Design for mDNA

Figure 7.3 shows an implementation of a meta-nucleotide tile as a DNA nanostructure. In Fig. 7.3a we have the tile with no *protection* and in Fig. 7.3b we have the tile with *protection* strands. The purpose of the protecting strands is to prevent spontaneous aggregation of tiles into mDNA. The protection mechanism is designed to impose the requirement that the 5′ pad is activated before the 3′ pad (this property

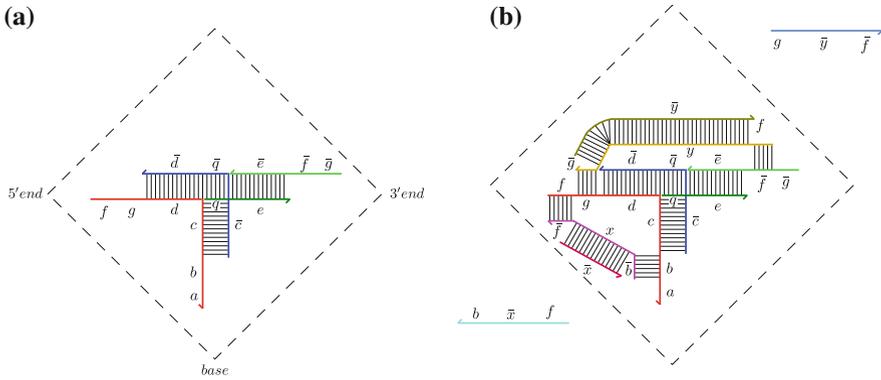


Fig. 7.3 Design of the meta-nucleotide

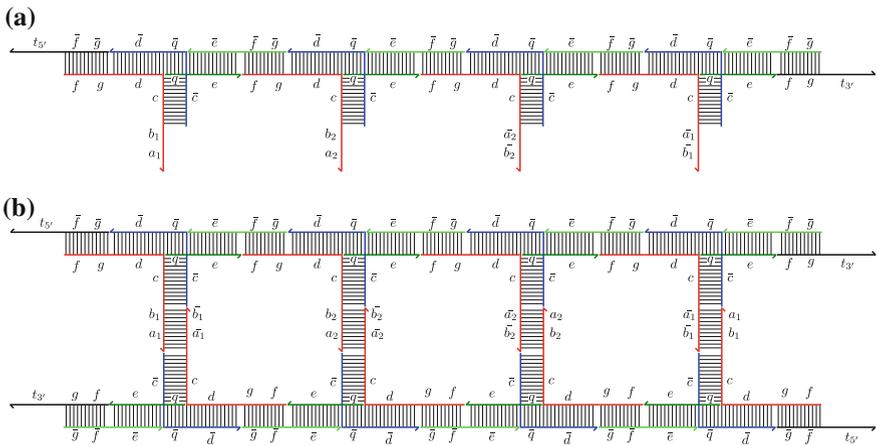


Fig. 7.4 Internals of mDNA: **a** internals of a single stranded mDNA, **b** internals of a double stranded mDNA

is used in the polymerization and replicator protocols). Each letter denotes a DNA sequence and a bar atop a letter indicates reverse complement of the sequence that the letter denotes. The red strand ($fgdcba$) contains the 5' pad fg and also the base pad ba and the light green strand ($\bar{g}\bar{f}\bar{e}$) contains the 3' pad $\bar{g}\bar{f}$. The blue strand ($\bar{c}\bar{q}\bar{d}$) and the dark green strand (qe) are bridging strands that hold the nanostructure together and give it the required geometry. We will have a detailed discussion of the secondary structure of the meta-nucleotide tile and mDNA in Sect. 7.3.4 (Fig. 7.4). The sequence of reactions that occur when a tile is deprotected are as follows. The pink strand $\bar{b}\bar{x}\bar{f}$ protects (renders inactive) the 5' pad. When ba binds to its complement on another tile, $\bar{b}\bar{x}\bar{f}$ is ripped away from the 5' pad by the invasion of the strand $f\bar{x}\bar{b}$ thus activating the 5' pad. The strand $f\bar{y}\bar{g}$ protects (renders inactive) the 3' pad. When the 5' pad binds to its complement on another tile, $f\bar{y}\bar{g}$ is ripped away from

the 3' pad by the invasion of the strand $g\bar{y}\bar{f}$ thus activating the 3' pad. Figure 7.5 illustrates all possible states a meta-nucleotide can exist in. We model *weak* base bonds and *strong* backbone bonds by making the sequences corresponding to the 5' and 3' pads much longer than the sequences corresponding to the base pads. The internal structure of both single and double stranded mDNA are shown in Fig. 7.4a and b. Note the black protection strands on the tiles at the ends of the mDNA to implement capping.

7.3.4 Secondary Structure of mDNA

Apart from simulating the reaction properties of DNA, we also wish to simulate its secondary structure, in particular:

- The flexible nature of *ssDNA* characterized by shorter persistence length
- The rigid nature of *dsDNA* characterized by a longer persistence length
- The double helical structure of *dsmDNA*.

Recently, Hamada and Murata [7] reported a novel self-assembled rigid T-shaped interconnected junction where each arm is a DNA double helix (Fig. 7.6a). They synthesized tile based structures like 1D linear ladders, 1D ringed structures and 2D lattices using the T-junction geometry. This suggests that their T-junction motif is a useful widget for designing stable, rigid, well behaved self-assembled objects. We use the T-junction as the key motif in achieving the required structural properties for mDNA. Each meta-nucleotide is a T-junction (Fig. 7.3a). *ssmDNA* (Fig. 7.4a) is a linear polymer of these T-junctions with consecutive base pad sections (the double helical structure $c \equiv \bar{c}$) not in the same plane, but slightly rotated so that they stick out of the plane of the paper. This rotation is controlled by carefully choosing the number of bases that make up the horizontal double helical section between consecutive vertical helical sections (Fig. 7.4a). This secondary structure for *ssmDNA* induces a helical twist for *ds* (imagined in Fig. 7.6b). We can think of the secondary structure of *dsmDNA* as a twisted 1D ladder. The 1D ladder design in [7] can be easily modified by adding or deleting a single base pair from the side rungs to induce a twist to get a double helix structure with approximately 10.5 meta-bases per turn of the double helix, mimicking the twist of *dsmDNA*. In particular, we choose the length of the repeating DNA double helical unit $efgdq \equiv \bar{q}\bar{d}\bar{g}\bar{f}\bar{e}$ to be either 41 (one less than the number of bases in four full turns of a DNA double helix) or 43 (one greater than the number of bases in four full turns of a DNA double helix). The following choice of lengths for the subsequences would potentially give us the required geometries and at the same time preserve the thermodynamic and kinetic properties that would allow our subsequent mDNA protocols to succeed ($|x|$ is the length of the DNA sequence represented by x): $|a_i| = |b_i| = |\bar{a}_i| = |\bar{b}_i| = 4$, $|c| = |\bar{c}| = |d| = |\bar{d}| = |q| = |\bar{q}| = 6$, $|f| = |\bar{f}| = |g| = |\bar{g}| = 12$ and $|e| = |\bar{e}| = 5$ or 7 .

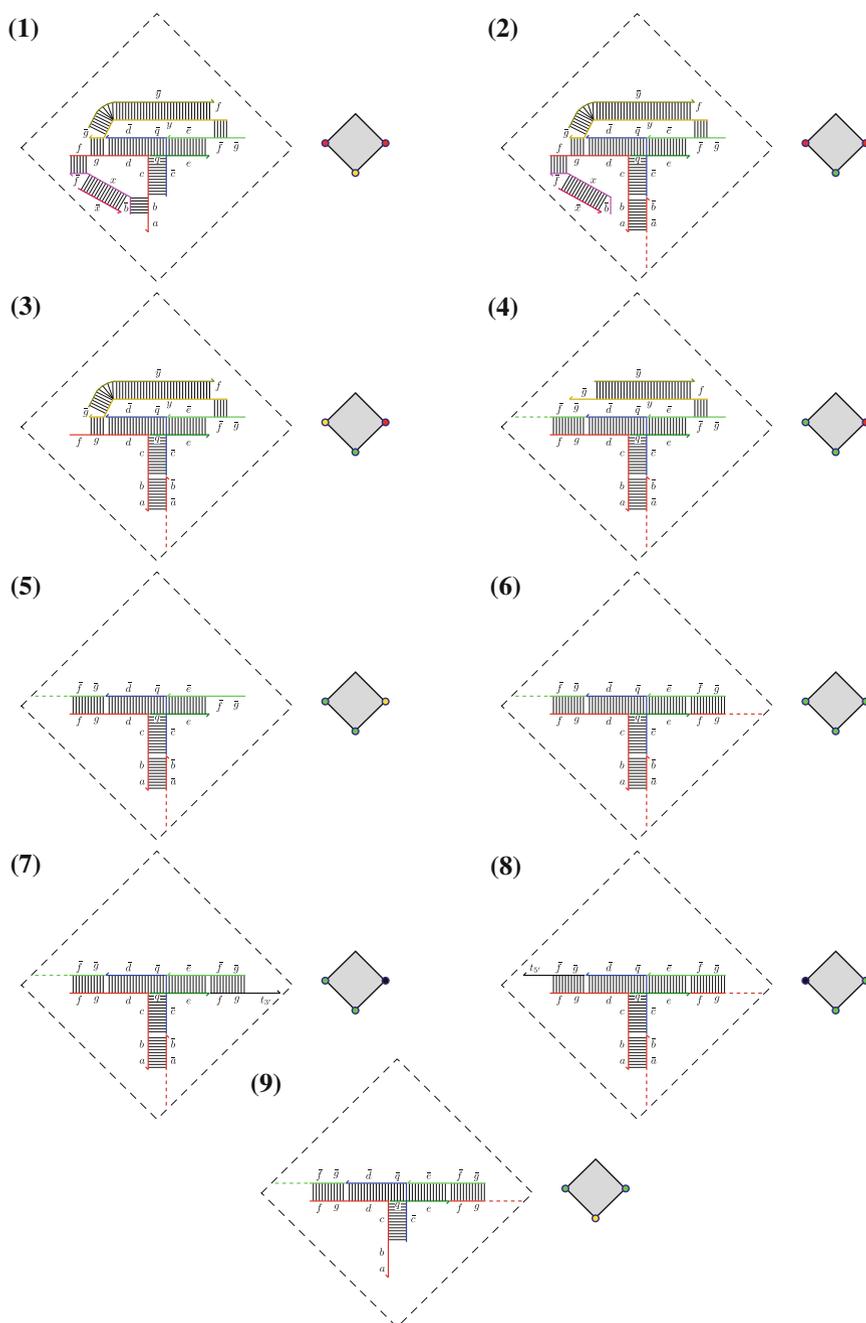


Fig. 7.5 States of the nucleotide

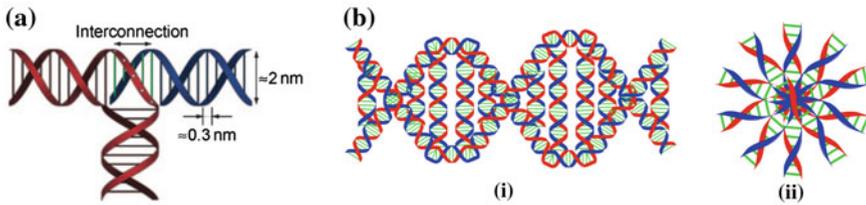


Fig. 7.6 **a** Design of rigid T-junction self-assembled from DNA. Figure from [7]. **b(i)** Secondary structure of *ds*mDNA forming a double helix (ratio of bases per axial rotation is programmable via strand design). Not to scale. **b(ii)** Axial view of *ds*mDNA. Only the rungs of the twisted ladder structure are shown

7.4 Assumptions for mDNA Reactions

Before describing our protocols involving mDNA we give here our assumptions under which these reactions proceed. We also describe some simple rules of DNA chemistry which are repeatedly used like subroutines in our mDNA protocols.

- Our systems are maintained at only three different temperatures characterized by a parameter τ . At room temperature or $\tau = 1$, both the meta-backbone bonds and the meta-base bonds are stable but breathing still occurs. Recall that breathing is the phenomenon of meta-base bonds spontaneously breaking and forming. At the melting temperature, $\tau = 2$, the meta-backbone bonds are stable but the base bonds are broken. At freezing temperature or $\tau = 0$, breathing does not occur.
- The pH, salt concentrations and other factors that affect hybridization are set to levels such that spontaneous hybridization between a DNA sequence and its complement can occur.
- Strand displacement, as defined earlier, always occurs and proceeds to completion. The strand that gets displaced out remains in the solution.
- We assume that each subsequence, denoted by a letter in the figures, only interacts with its perfect and full reverse complements and no other spurious interactions occur.

7.5 Reactions in mDNA

We set out protocols for mDNA that mimic DNA–DNA and DNA–enzyme interactions. We have two kinds of figures in the discussions that follow. The abstraction diagrams illustrate the protocols in the abstract activatable tile model while the internal structure diagrams illustrate the protocols in greater detail.

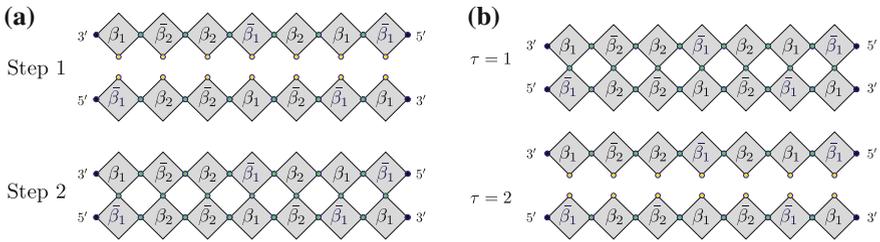


Fig. 7.7 Meta-hybridization and meta-denaturation by heating: **a** meta-hybridization reaction, **b** meta-denaturation at $\tau = 2$

7.5.1 Meta-hybridization and Meta-denaturation in mDNA

The simplest mDNA reactions are meta-hybridization and meta-denaturation, which are reverse reactions of each other. In meta-hybridization, at temperature $\tau = 1$ two complementary ssmDNA strands bind via complementary base pad bindings to give dsmDNA (Fig. 7.7a). Heating dsmDNA to temperature $\tau = 2$ meta-denatures the structure into its two component ssmDNA (Fig. 7.7b).

7.5.2 Meta-strand Displacement Process in mDNA

Meta-strand displacement for mDNA is defined as displacement of a ssmDNA from a meta-double helix by an incoming ssmDNA with a longer complementary region to the template strand. This reaction occurs at temperature $\tau = 1$. Figure 7.8 gives a high level view of meta-strand displacement using the activatable tile model. Two meta-strands compete to hybridize with a single ssmDNA. The shorter of the two is completely meta-hybridized to the template while the longer one comes in by gaining a toehold. Now, breathing of the meta-bases of the short strand gives an opportunity to the meta-bases of the competing incoming strand to meta-hybridize with the template. Note that the intermediate steps are reversible. However, once the incoming meta-strand completely displaces the outgoing meta-strand, the reaction stops as the outgoing meta-strand is extremely unlikely to come back in as it lacks a toehold.

Note that this reaction is made possible because the weak meta-base bonds can breathe at $\tau = 1$. Breathing in mDNA is expected to occur at a slower rate than in DNA because we require multiple bases to spontaneously denature for a single meta-base to denature. Also, contiguous bases in DNA are more immediately local than contiguous meta-bases in mDNA and hence the rate at which a meta-base occupies an empty spot on a complementary meta-base is also expected to be slower than for the corresponding process in DNA. Due to these reasons, we would expect meta-strand displacement in mDNA to proceed slower than strand displacement in DNA.

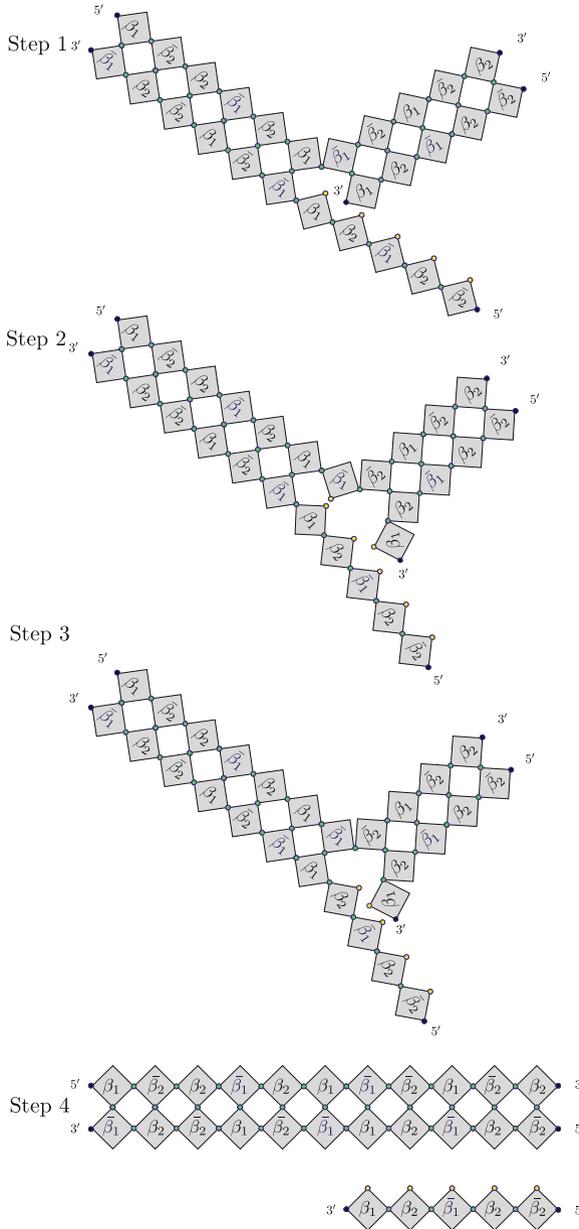


Fig. 7.8 Meta-strand displacement due to breathing at $\tau = 1$

7.5.3 Meta-polymerization in mDNA

Meta-polymerization in mDNA occurs by the extension of a *ssmDNA*, called a primer, by additions of free meta-nucleotides in the solution to the 3' end via meta-backbone bonds. The extension cannot occur *de novo*, it requires a primer meta-hybridized to a template *ssmDNA*. Neither can meta-nucleotides spontaneously aggregate, because of the protection strands (see Fig. 7.1). The meta-nucleotides added have base pads complementary to the corresponding base pads of the template strand.

Figure 7.9 gives an activatable tile model view of a single step in mDNA meta-polymerization. Each base pad of the template strand that is not bound to its complement is in the active unbound state. The 3' pad of the tile at the 3' end of the primer is in the capped state (step 1). It transitions to the active unbound state when its black capping strand is displaced by an initiator strand. The incoming meta-nucleotide (step 2) has its 5' and 3' pads in the inactive state, while its base pad is in the active unbound state. The complementary base pads bind (step 3), activating the 5' pad of the incoming nucleotide (step 4). The 5' pad then binds to the 3' pad of the previous meta-nucleotide (step 5) causing the 3' pad of the incoming nucleotide to transition to the active unbound state (step 6). This process occurs repeatedly till either no further free meta-nucleotides are available or the end of the template strand is reached. There is an alternate mechanism to stop the polymerization, which involves adding to the solution a black capper strand to transition the 3' pad of growing 3' end to a capped state, thus stopping further additions. Note that in a solution with an ensemble of these nanostructures, in general the stopping point of meta-polymerization cannot be carefully controlled and such an attempt would probably lead to various length subsequences of the fully complementary *ssmDNA*. However, we can exclude certain meta-base types from the meta-nucleotide mix in the solution ensuring that the polymerization halts when the complementary meta-base is encountered on the template strand. Since we have access to a larger alphabet of bases in our mDNA systems we can set aside a set of bases for such purposes, analogous to *stop codons* in translation.

We will now examine this protocol in greater detail by looking at the internal strand structure (Figs. 7.10 and 7.11). The 3' pad of the tile at the 3' end of the primer is capped by fgt_3 (step 1). The initiator strand $t_3\bar{g}\bar{f}$ binds to the capper strand and activates the 3' pad $\bar{g}\bar{f}$ (step 2). Now, the protected meta-nucleotide comes in (step 3) and binds to the complementary base sequence $\bar{a}_2\bar{b}_2$, displacing the strand $\bar{b}_2x\bar{f}$ (step 4). The toehold \bar{b}_2 is exposed on $\bar{b}_2x\bar{f}$ allowing $f\bar{x}\bar{b}_2$ to bind (step 5). This strand displaces $\bar{b}_2x\bar{f}$ exposing f on the 5' end of the incoming meta-nucleotide (step 6). Now, the 5' end of the incoming meta-nucleotide binds to the 3' end of the previous nucleotide, displacing $fy\bar{g}$ (step 7). The strand $fy\bar{g}$ is stripped away by its complement through the toehold \bar{g} , activating the 3' end ($\bar{g}\bar{f}$) of the incoming meta-nucleotide (step 8). This brings the 3' end of the growing *ssmDNA* back to the same state as in step 2 and thus the reaction can repeat till the end of the template strand is reached (step 9). At this point, we introduce the 3' capper sequence fgt_3

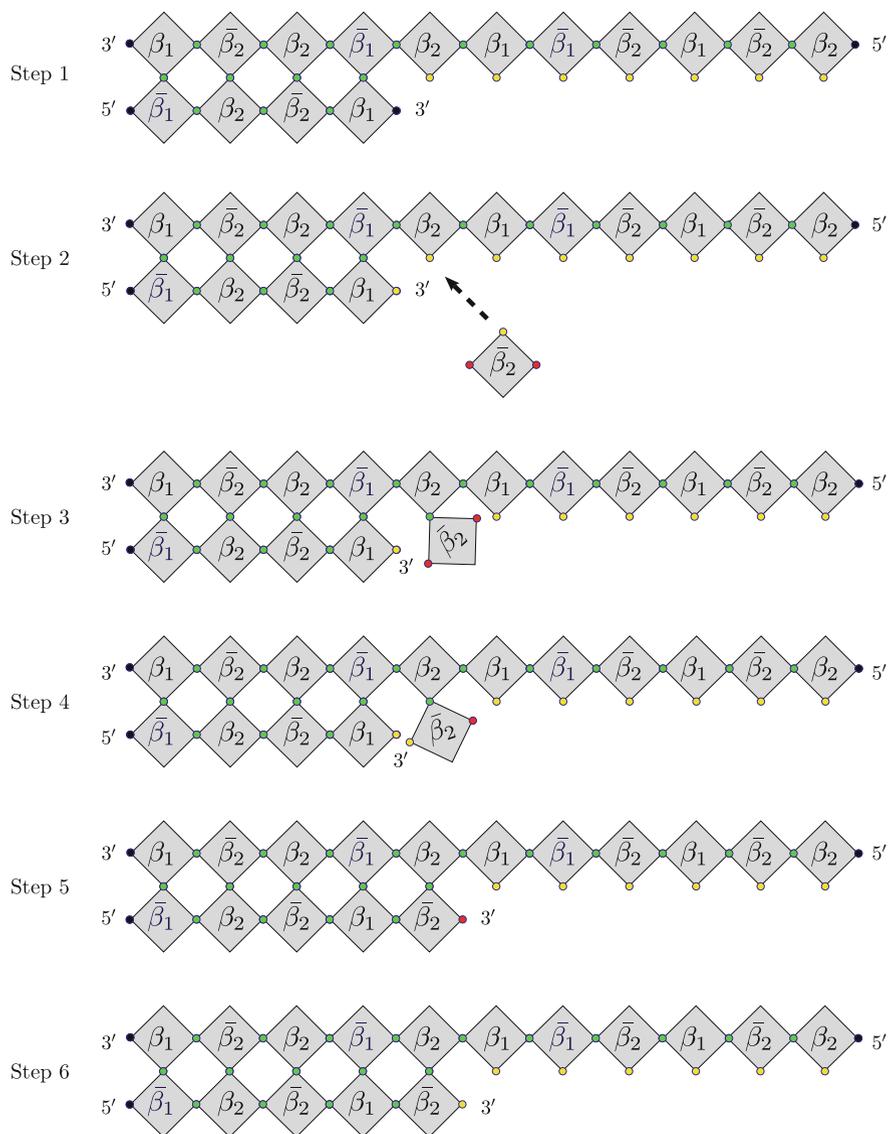


Fig. 7.9 Meta-polymerization reaction

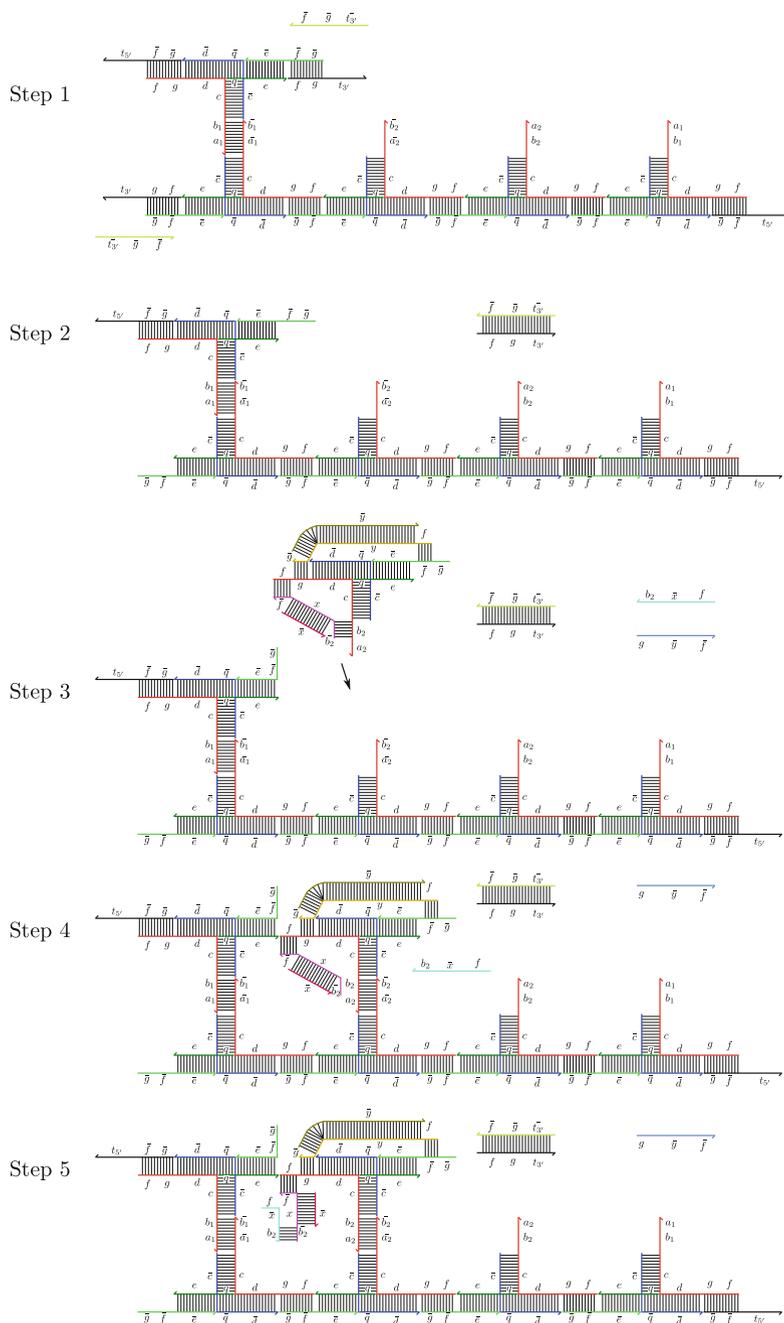


Fig. 7.10 Internals of meta-polymerization reaction-I

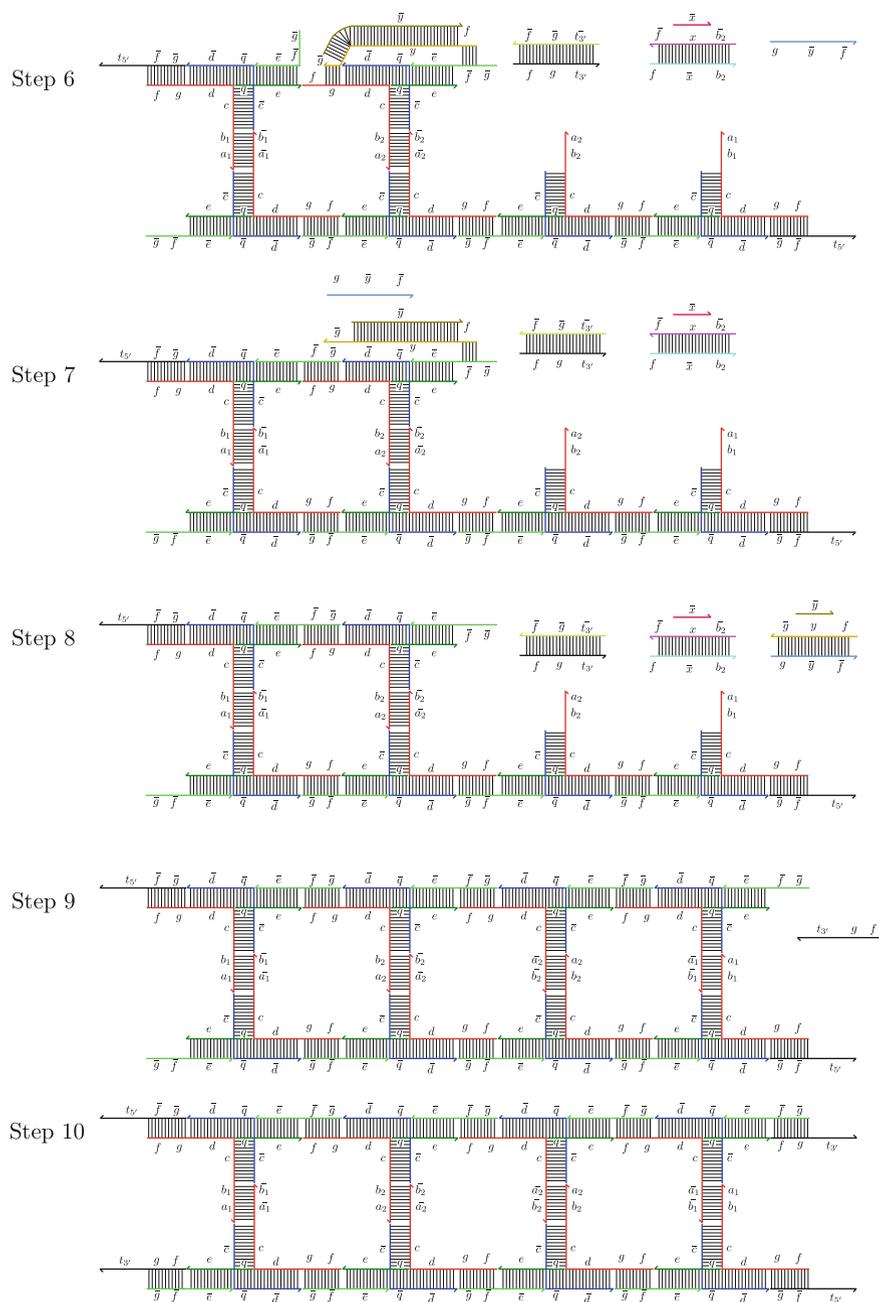


Fig. 7.11 Internals of meta-polymerization reaction-II

which binds to the 3' pad of the last meta-nucleotide added. This terminates the polymerization reaction. Note that in a solution with an ensemble of these nanostructures, we must wait for each copy of the reaction to proceed to step 9 before introducing the capper sequence or we must program a stopper sequence, analogous to *stop codons* in translation. DNA polymerases like $\Phi 29$ possess exceptional strand displacing capability which aids in isothermal amplification of DNA. However, our meta-polymerization protocol lacks this ability. Thus, if in the course of extending the primer an already bound *ssmDNA* is encountered then the meta-polymerization process halts. In Sect. 7.5.6, we show how to achieve isothermal amplification of *mDNA*.

7.5.4 Meta-restriction Cuts in *mDNA*

We can achieve site specific meta-restriction in both *ssmDNA* and *dsmDNA* by slightly modifying the internal structure of a nucleotide. This modification does not significantly affect the secondary structure of the *mDNA* and is compatible with all the other *mDNA* protocols described in this chapter. We prefix a sequence h_i to the 5' pad of the meta-nucleotide with base pad β_i . This is the sequence that will be recognized by cleaving strands that break the meta-backbone bonds. We illustrate restriction for *dsmDNA*. The protocol for *ssmDNA* is very similar and can be thought of as a special case of the meta-restriction of *dsmDNA*.

The abstract activatable tile model of meta-restriction is illustrated in Fig. 7.12. The 5' end of the meta-nucleotide β_1 is recognized, cut and sent to the capped state. In Fig. 7.13 the site h_2 associated with the meta-base b_2a_2 is recognized by the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$. There is a single recognition site h_2 on each of the strands of the *dsmDNA* (step 1). The cleaver strands bind to the toehold h_2 on each strand (step 2) and break the meta-backbone bond by strand displacing the 3' pad $\bar{g}\bar{f}$ (step 3). We now introduce the 3' capper sequence fgt_3' to cap the exposed 3' pads (step 4). We prevent interaction between the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$ and the 3' capper sequence fgt_3' by executing step 4 only after we are reasonably certain that step 3 is complete. This means that meta-restriction is not autonomous. The strand \bar{r}_2h_2fg is introduced to strip away the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$ from the 5' pad by exploiting the toehold r_2 (step 5). This exposes the 5' pads (step 6) which are then capped by introducing the 5' capper sequence $\bar{g}\bar{f}t_5'$ (step 7), completing the process of meta-restriction. Again, we prevent interaction between the strand \bar{r}_2h_2fg and the 5' capper sequence $\bar{g}\bar{f}t_5'$ by executing step 7 only after we are reasonably certain that step 6 is complete. We do not require that these extraneous strands do not interact at all. Rather, it is sufficient that even after interacting among themselves there are a sufficient concentration of them to perform the tasks described in Figs. 7.13 and 7.14.

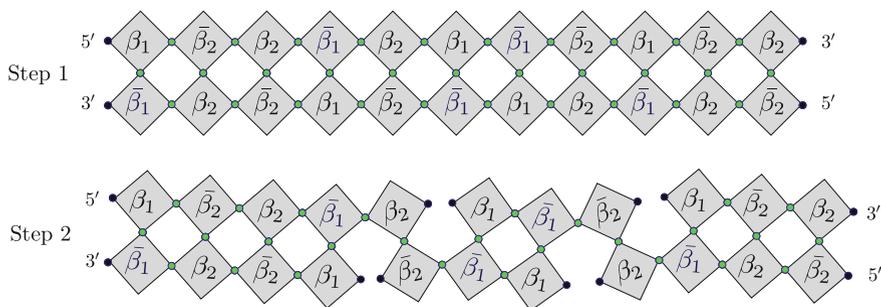


Fig. 7.12 Abstract description of meta-restriction at site β_1

7.5.5 Meta-helicase Denaturation Reaction in mDNA

We have previously described meta-denaturation for mDNA using temperature. The same result can be achieved isothermally, which we call meta-helicase denaturation. Consider the abstract view of meta-helicase denaturation given in Fig. 7.15. Denaturation is the breaking of meta-base bonds of a *dsmDNA*. When all the meta-base bonds are broken, the meta-strands float apart. Initially all the meta-base pads are in the bound state (step 1). Meta-helicase activity breaks the bonds and sends the base pads to the inactive state. This meta-helicase reaction doesn't necessarily act contiguously. Some meta-base bonds are broken before others (steps 2 and 3). When all the meta-base pads are broken the meta-strands float apart (step 4). We can reactivate the meta-base pads by transitioning them to the active unbound state (step 5) at which point the *ssmDNA* can recombine to form a *dsmDNA*.

Meta-helicase activity in mDNA is performed by a host of strands (colored pink in Fig. 7.16), two for each type of meta-base bond that must be broken. In Fig. 7.16 there are two types of meta-base bonds and hence we have four strands to perform helicase activity (step 1). We introduce a slight modification in the internal strand structure by appending a sequence p to the base pad of each meta-nucleotide. This modification does not significantly affect the secondary structure of mDNA and neither does it interfere with any other mDNA protocol described in this chapter. This sequence (p) will act as a toehold for the strands involved in the meta-helicase process. A pair of strands invade the meta-base pad bond and break them (step 2). Note that half of the meta-base bond is broken by one of these strands and half by the other, ensuring symmetry. Once each meta-base bond is broken the meta-strands drift apart (step 3) and meta-helicase activity is complete. We can strip off the protecting strand, for example $\bar{m}\bar{p}b_1$, by adding in its complement, say b_1pm (step 4). The *ssmDNA* can now recombine into *dsmDNA*. It is possible to repeat this process of meta-denaturation and meta-renaturation, however the protocol would not be autonomous.

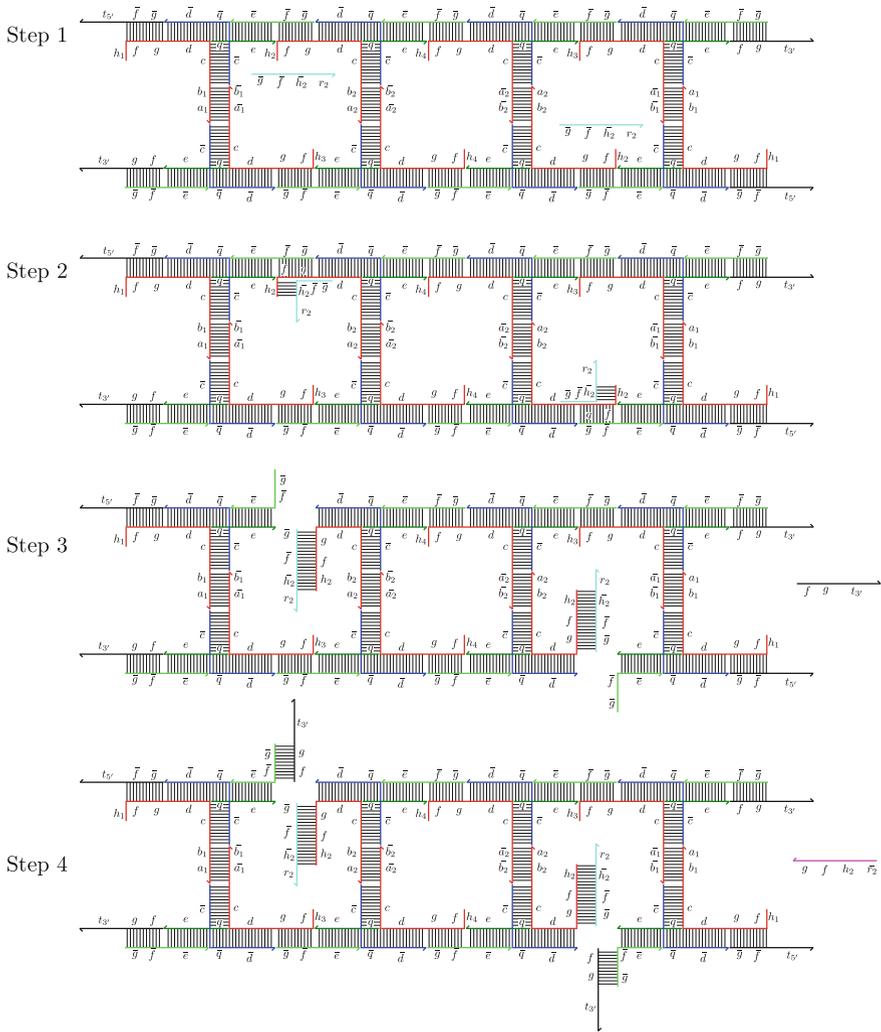


Fig. 7.13 Internals of meta-restriction at h_2 associated with meta-base b_{2A2-I}

7.5.6 Exponential Amplification in mDNA Using a Replicator

Our protocol for meta-polymerization in mDNA lacked meta-strand displacement capabilities and hence could not be used for isothermal PCR-like amplification. In this section we describe a method to get isothermal exponential amplification using a replicator mechanism. The protocol is similar to meta-polymerization, it involves linear contiguous extension of a primer by addition of meta-nucleotides, with the newly polymerized *ssmDNA* having the complementary sequence to that of the

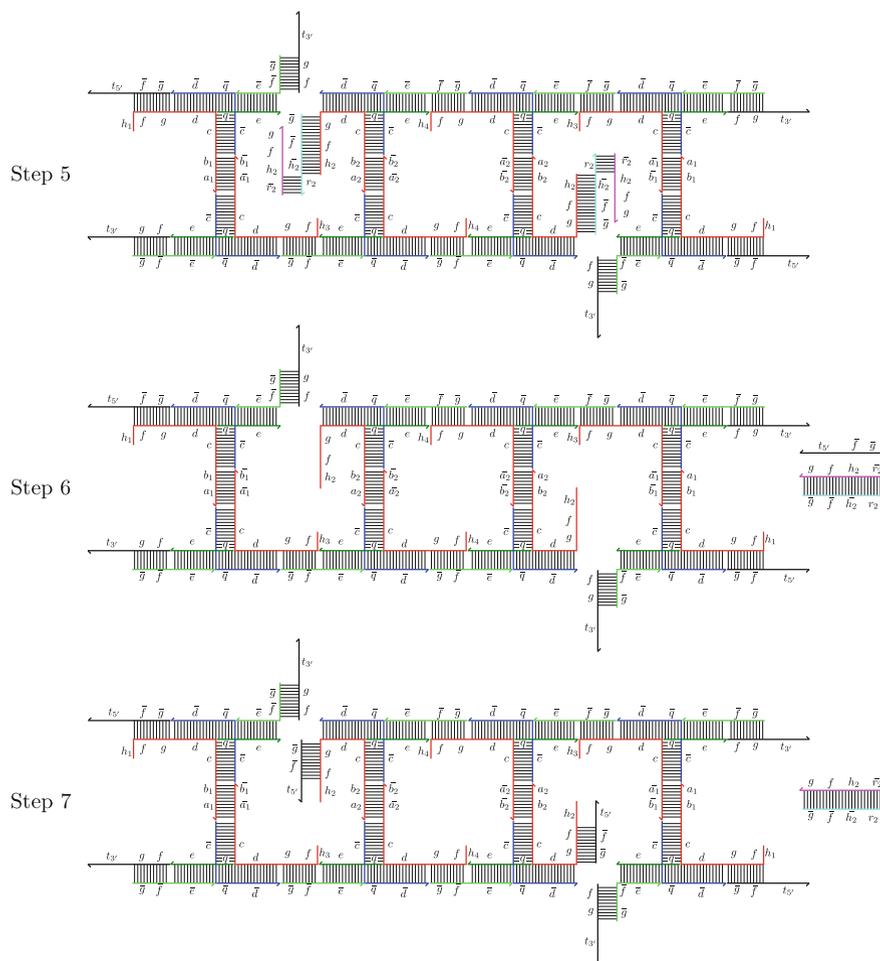


Fig. 7.14 Internals of meta-restriction at h_2 associated with meta-base b_{2a2-II}

template. The key difference in the two protocols is a mechanism to isothermally dissociate the newly synthesized *ssmDNA* from the template. We achieve this using a new idea, not present in any previous protocol described in this work, of active DNA sequences sequestered in hairpins that are released by strand displacement reactions. The release of the “hidden” DNA sequence inside the hairpin structure can be thought of as an activation step, setting off another strand displacement reaction. This idea of sequestering sequences within hairpins has been demonstrated previously by Zhang et al. [28] and Dirks and Pierce [6].

For the purposes of this protocol, we define a new state of the base pad in the activatable model of *mDNA*, which we call *semi active unbound* (shaded purple in Fig. 7.17c). If two base pads are in the semi active unbound state, they cannot bind to

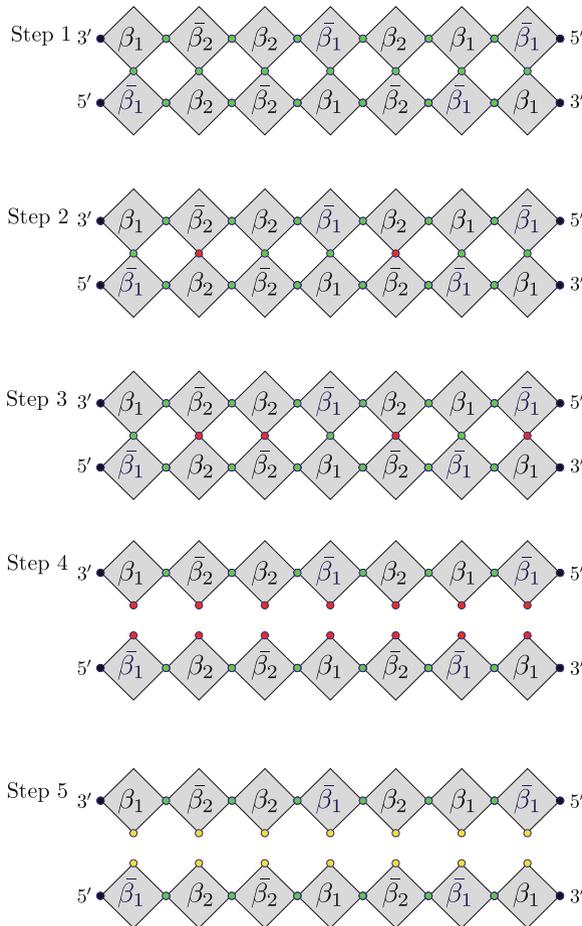


Fig. 7.15 Abstract description of meta-denaturation by helicase action

each other. However, a semi active unbound base pad can bind with an active unbound base pad, provided their sequences are complementary. There are also modifications (Figs. 7.17a, b) to how the other pad states are implemented, however the properties of the state do not change. Note the introduction of the hairpin structure $g_2 \bar{p} \bar{a} l \bar{g}_2$ at the 3' end of a meta-nucleotide (Fig. 7.17a) which will be used to cleave the growing strand from the template (Fig. 7.17c), sending the base pad to the semi active unbound state.

Figure 7.18 illustrates the replicator in the abstract activatable tile model of mDNA. The top strand is the template and the bottom strand grows from a primer. For initiating and terminating replication, we require the use of a pair of special tiles with complementary base pads, labeled $\beta_{3'}$ and $\beta_{3'}$ in Fig. 7.18. These are capped at one end and occur at the terminal ends of the template and hence in each replicated

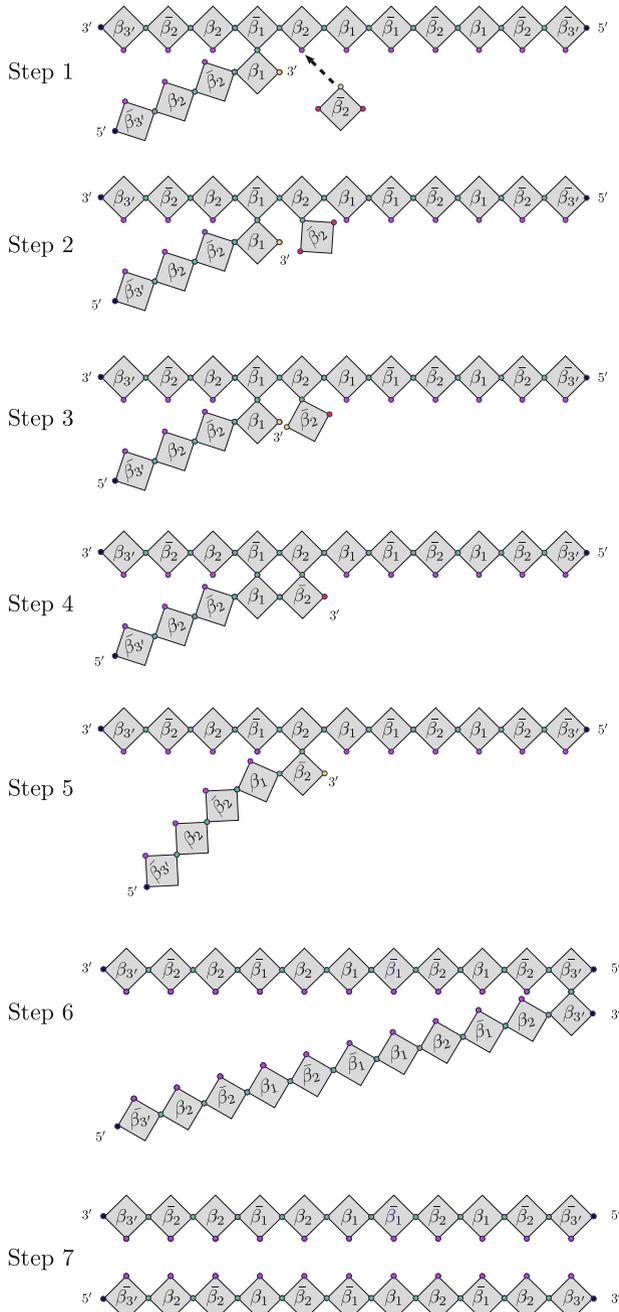


Fig. 7.18 Replicator abstraction

mDNA. $\bar{\beta}_{3'}$ acts as the primer, initiating the replication. Consider the replication process after a few meta-nucleotides have been added (step 1). The growing strand is attached via a single base pad (at the 3' terminal meta-nucleotide) to the template. The base pads of the rest of the meta-nucleotides are in the semi active unbound state and hence cannot bind with each other. A new meta-nucleotide comes in (step 1) with its base pad in the active unbound state and binds to the template (step 2). This activates its 5' end (step 3) and allows the 3' end of the growing strand to bind (step 4). This activates the 3' end of the meta-nucleotide added and also cleaves the meta-base pad binding between the template and the previously attached meta-nucleotide of the growing strand, sending the cleaved meta-base pads to the semi active unbound state (step 5). The process (step 1–5) repeats till the last meta-nucleotide $\beta_{3'}$ attaches. At this point, we want the two strands to separate. We achieve this by making the meta-base pad bond between $\beta_{3'}$ and $\bar{\beta}_{3'}$ relatively weak, allowing the strands to separate spontaneously due to breathing of the base pad bond. This is implemented by choosing a very short length sequence for the base pads of $\beta_{3'}$ and $\bar{\beta}_{3'}$. Both the template and the newly synthesized *ssmDNA* can now act as templates for further replication and hence we can achieve exponential amplification. Note that we can easily introduce mutations in the replicator mechanism by designing DNA sequences for certain base pads that do not have any exact complementary base pads but rather several partial complements. This allows us to probabilistically evolve a diverse sequence population of *ssmDNA*.

Let us examine the replicator protocol in greater detail, paying attention to the strand diagram (Figs. 7.19 and 7.20). We have a four meta-base mDNA sequence, with the first and last bases being the special terminator bases (step 1). Note that the sequence $g = g_1g_2$ and $\bar{g} = \bar{g}_2\bar{g}_1$. A meta-nucleotide with an active unbound base pad (step 1) comes in and binds to the template via the sequence b_2a_2 (step 2), strand displacing out the sequences $\bar{p}b_2$ (part of the light green strand) and \bar{b}_2 (part of the pink strand). The strand $f\bar{x}b_2$ now strips away $\bar{b}_2x\bar{f}$, activating the 5' pad of the incoming meta-nucleotide (step 3). This allows the 5' pad $f g_1g_2$ to bind with the 3' pad $\bar{g}_2\bar{g}_1\bar{f}$, displacing $\bar{g}_2\bar{g}_1$ and opening up the hairpin structure $\bar{p}\bar{a}_1l$ (step 4). This allows the strand $f y\bar{g}_2\bar{g}_1$ to be stripped away through the toehold $\bar{g}_2\bar{g}_1$, activating the 3' end (step 5). The released hairpin from step 4 can now cleave a_1 from \bar{a}_1 using the toehold p , which is half of the base pad binding, while the other half \bar{b}_1 of the base pad binding is cleaved from b_1 via the toehold p by the sequence $\bar{p}b_1$ (step 6). Note that cleaving the bond between b_1 and \bar{b}_1 is actually a reversible process, and could have occurred after step 1 itself. Only when it is combined with the cleaving of the bond between a_1 and \bar{a}_1 does it get biased towards the configuration depicted in step 6. This completes the addition of a single meta-nucleotide accompanied by cleaving the previous meta-nucleotide. The process repeats till the last meta-nucleotide is added at which point the short base sequence of $\bar{a}_3\bar{b}_3$ allows the *ssmDNA* to separate due to breathing. We note that both the separate *ssmDNA* are in a configuration that allows them to act as template strands for further replication, allowing exponential amplification of mDNA.

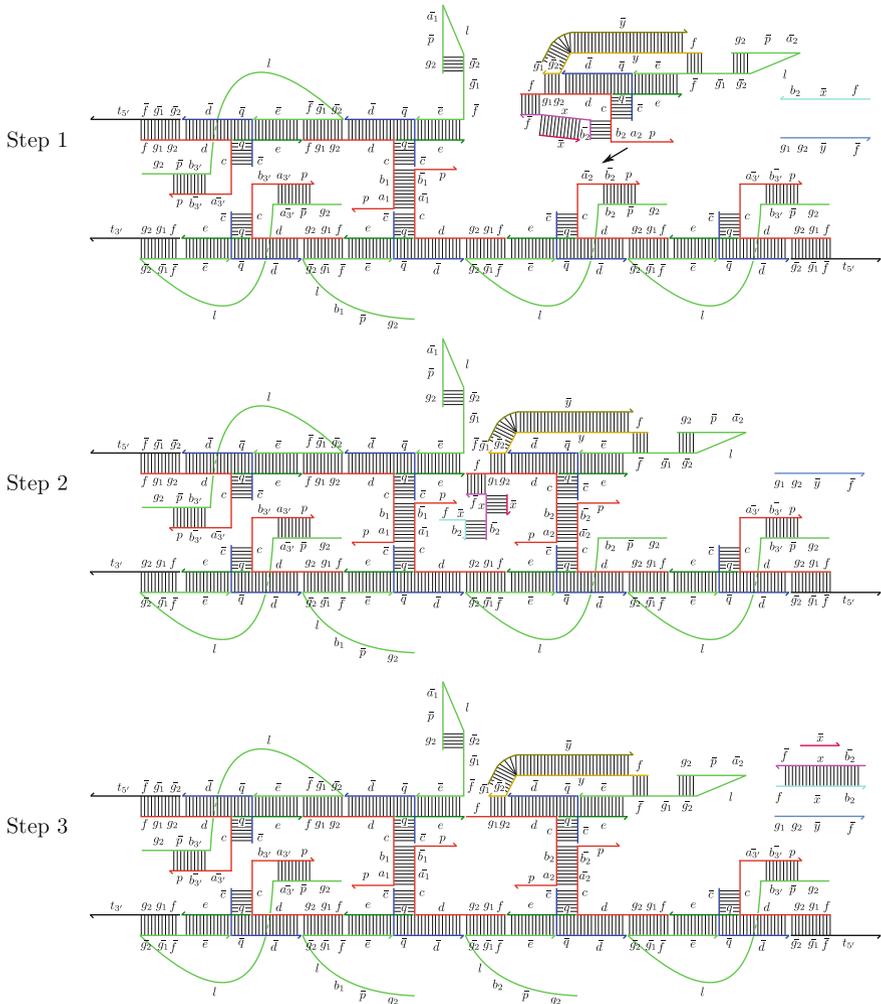


Fig. 7.19 Internals of exponential amplification by replication-I

7.6 Lessons Learnt

In this work we have outlined a synthetic biochemical system made purely from DNA strands that simulates the behavior of various protein enzymes acting on DNA. We believe mDNA systems are a powerful programming paradigm for designing complex biochemical systems. In Sect. 7.5.6 we provided a protocol for exponential self-replication of a meta-DNA sequence and pointed out how to introduce mutations during replication and thus achieving sequence diversity. Incorporating environmental selection pressure in mDNA to evolve functional biosystems is a major open

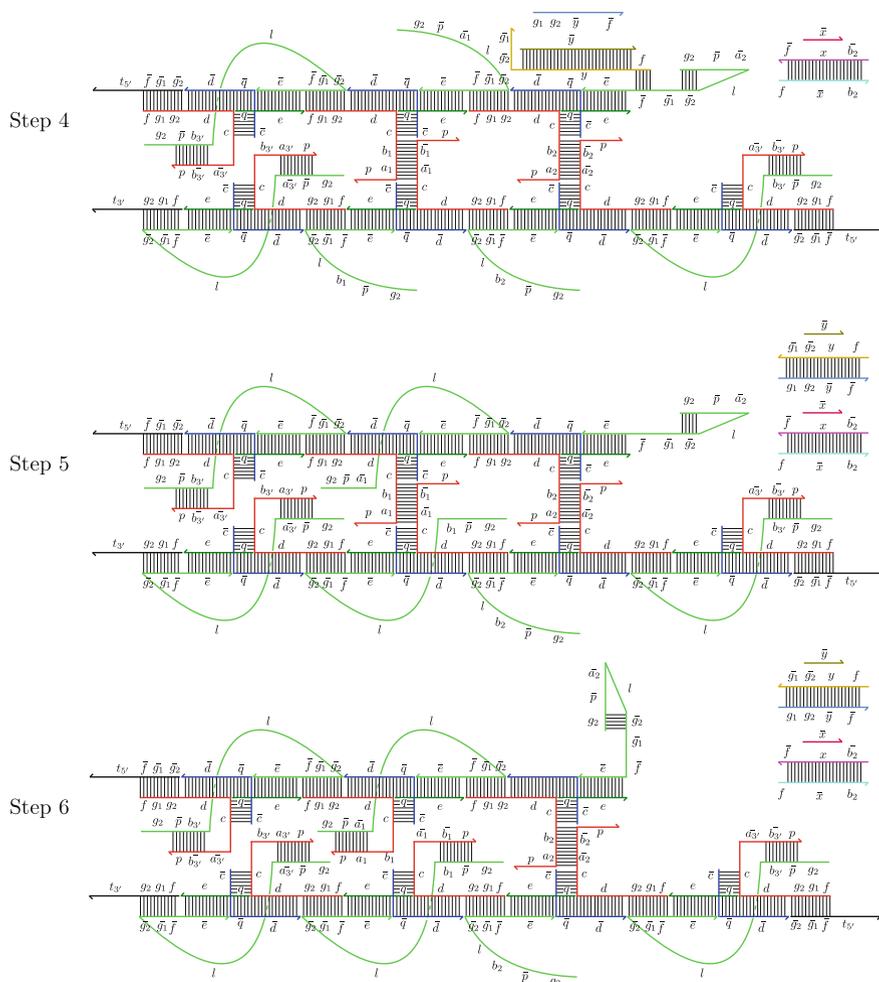


Fig. 7.20 Internals of exponential amplification by replication-II

challenge. The next step in our investigations of mDNA systems is development of a strand level kinetic simulation of mDNA protocols. We hope to use data from such simulations to enhance and fine tune our protocols and also perform in vitro experiments to validate them. A major facilitating factor in the field of DNA nanotechnology has been technology to efficiently synthesize synthetic DNA strands of defined sequence de novo from individual nucleotides. Analogously, we require methods to efficiently (high throughput with low error rates) synthesize mDNA strands of defined sequence from synthetic DNA. The current protection-deprotection DNA synthesis protocols can be adapted for the use of synthesizing mDNA from pre-formed meta-nucleotides. Note that meta-nucleotides can be synthesized using stan-

standard hierarchical assembly techniques from DNA self-assembly developed by Park et al. [14]. Other important protocols for mDNA systems are meta-polymerization with meta-strand displacing capability and meta-ligation of mDNA strands.

Acknowledgments We wish to acknowledge support by NSF grants CCF-0829797, CCF-0829798, CCF-1217457 and CCF-1141847.

References

1. Amin R, Kim S, Park SH, LaBean T (2009) Artificially designed DNA nanostructures. *NANO: Br Rep Rev* 4(3):119–139
2. Bath J, Turberfield A (2007) DNA nanomachines. *Nature Nanotechnol* 2:275–284
3. Bedau M, McCaskill J, Packard N, Rasmussen S, Adami C, Green D, Ikegami T, Kaneko K, Ray T (2000) Open problems in artificial life. *Artif Life* 6(4):363–376
4. Dadon Z, Wagner N, Ashkenasy G (2008) The road to non-enzymatic molecular networks. *Angewandte Chemie Int Ed* 47(33):6128–6136
5. Deng Z, Chen Y, Tian Y, Mao C (2006) A fresh look at DNA nanotechnology. *Nanotechnol: Sci Comput*, 23–34
6. Dirks R, Pierce N (2004) Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci USA* 101(43):15275–15278
7. Hamada S, Murata S (2009) Substrate-assisted assembly of interconnected single-duplex DNA nanostructures. *Angewandte Chemie Int Ed* 48(37):6820–6823
8. Koval V, Gnedenko O, Ivanov Y, Fedorova O, Archakov A, Knorre D (1999) Real-time oligonucleotide hybridization kinetics monitored by resonant mirror technique. *IUBMB Life* 48(3):317–320
9. LaBean T, Gothelf K, Reif J (2007) Self-assembling DNA nanostructures for patterned molecular assembly. *Nanobiotechnology II* 79–97
10. Luisi PL (2006) *The emergence of life—from chemical origins to synthetic biology*. Cambridge University Press, Cambridge
11. Lund K, Williams B, Ke Y, Liu Y, Yan H (2006) DNA nanotechnology: a rapidly evolving field. *Curr Nanosci* 2:113–122
12. Majumder U, LaBean T, Reif J (2007) Activatable tiles: compact, robust programmable assembly and other applications. In: *DNA Computing*
13. Packard N, Bedau M (2003) Artificial Life. *Encycl Cogn Sci* 1:209–215
14. Park SH, Pistol C, Ahn SJ, Reif J, Lebeck A, LaBean CDT (2006) Finite-size, fully addressable DNA tile lattices formed by hierarchical assembly procedures. *Angewandte Chemie Int Ed* 45(5):735–739
15. Reif J, Majumder U (2008) Isothermal reactivating Whiplash PCR for locally programmable molecular computation. *DNA Comput* 41–56
16. Sakamoto K, Kiga D, Momiya K, Gouzu H, Yokoyama S, Ikeda S, Sugiyama H, Hagiya M (1999) State Transitions by Molecules. *Biosystems* 81–91
17. Schulman R, Winfree E (2008) How crystals that sense and respond to their environments could evolve. *Nat Comput* 7(2):219–237
18. Seeman N (2004) Nanotechnology and the double helix. *Sci Am* 290(6):64–75
19. Sherman W, Seeman N (2004) A precisely controlled DNA biped walking device. *Nano Lett* 4:1203–1207
20. Smith A, Turney P, Ewaschuk R (2002) JohnnyVon: self-replicating automata in continuous two-dimensional space. *Comput Res Repository*
21. Tian Y, He Y, Mao C (2006) Cascade signal amplification for DNA detection. *ChemBioChem* 7(12):1862–1864

22. Tjivikua T, Ballester P, Rebek J (1990) A self-replicating system. *J Am Chem Soc* 112(2): 1249–1250
23. Vidonne A, Philp D (2009) Making molecules make themselves—the chemistry of artificial replicators. *Eur J Org Chem* 5:593–610
24. von Kiedrowski G (1986) A self-replicating hexadeoxynucleotide. *Angewandte Chemie International Edition* 25(10):932–935
25. von Kiedrowski G, Patzke V (2007) Self replicating systems. *ARKIVOK* 293–310
26. Winfree E (2003) DNA computing by self-assembly. *NAE's Bridge* 33:31–38
27. Yin P, Yan H, Daniell X, Turberfield A, Reif J (2004) A unidirectional DNA walker moving autonomously along a linear track. *Angewandte Chemie Int Ed* 116(37):5014–5019
28. Zhang D, Turberfield A, Yurke B, Winfree E (2007) Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* 318:1121–1125
29. Zhang DY, Winfree E (2009) Control of DNA strand displacement kinetics using toehold exchange. *J Am Chem Soc* 131(48):17303–17314
30. Zhang DY, Yurke B (2006) A DNA superstructure-based replicator without product inhibition. *Nat Comput* 5(2):183–202

Chapter 8

Towards Modeling Automation for Synthetic Biology

Chen Liao and Yizhi Cai

Abstract Rule-based modeling was first introduced to address the ‘combinatorial complexity’ problem in cellular signaling. A number of software tools and methods have been developed in recent years to make accurate predictions about the functional role of proteins in signaling transduction systems. Many of these approaches are based on formal languages, such as Kappa and BioNetGen (BNGL). Modeling also plays an integrant role in synthetic biology, a new interdisciplinary subject aiming to design novel biological systems. The specification of synthetic biology systems using high level languages is still a challenge. In this article, we proposed to extend the rule-based modeling from systems biology to synthetic biology and introduced a new model-specification language, which allows quickly generating mathematical models encoding the phenotypical behaviors of biological systems. Our approach (termed AutoModel) also takes into account the context dependencies of biological interactions, which makes it a desirable method for synthetic biology research. A software implementation of our approach is available at <https://github.com/cliao15/Rulebase2011>.

Keywords Automatic modeling · Designer sequences · Rule-based modeling · DNA · RNA · Species · Pattern · Wildcard · Qualifier · Reaction rule · Kinetic law · Mass action · Hill kinetics

C. Liao
Department of Bioengineering, University of Illinois at Urbana-Champaign,
Urbana, IL, USA
e-mail: cliao15@illinois.edu

Y. Cai (✉)
School of Biological Sciences, University of Edinburgh,
The King’s Buildings, Edinburgh EH9 3JR, UK
e-mail: yizhi.cai@ed.ac.uk

Y. Cai
Institute for Systems Biology and Synthetic Biology, University of Edinburgh,
Edinburgh, UK

8.1 Context-Dependency in Automatic Modeling

Recent progress has been made to develop rule-based approaches for modeling gene expression and regulation (GenoCAD [1, 2], GEC [22], *etc.*). However, the existing methods are limited to model the impact of nonlocal contextual information on biochemical processes. GEC associates a single property $rate(R)$ to each ribosome binding site (RBS) so a translational process can be represented as $m \rightarrow \{R\}m + p$, where m is for mRNA and p is for protein. This simplification assumes that the rate of mRNA translation is solely dependent on RBS (local information) while nonlocal information is ignored. However, this assumption does not work well all the time as the secondary structure of mRNA also plays an important role in determining the translation efficiency [5, 6, 15]. One possibility is that downstream genes may form secondary structure with RBS and then completely or partially block the translation initiation. For this respect, GenoCAD computes the translation rate at the cistron level by associating the rate with information of both RBS and its downstream gene [2]. Considering mRNA translation can be controlled by a number of mechanisms varied on different levels of complexity [9], merely pairing RBS with genes cannot meet the requirement of accurate modeling until more contextual genetic elements participating in the translational process are explicitly accounted for.

8.2 AutoModel

In this chapter, we will introduce a new tool AutoModel, which utilises the power of rule-based modeling to automatically generate and simulate the genetic network encoded by the designer sequences. Our approach partially overlaps with Kappa [4] and BNGL [7] by sharing the same idea that reactions can be identified between species possessing specific traits in molecular structures. Despite of similar modeling framework, our method is more suited for modeling genetic networks composed of DNA and RNA sequences, proteins as well as chemical molecules while theirs work better for modeling signaling transduction systems. The flow chart of our approach is shown in Fig. 8.1.

Section 8.2.1 defines reaction rules and describes every component of a reaction rule in a detailed manner. Section 8.2.2 gives a short summary of diverse biological interactions that can be described by our language through examples. Section 8.2.3 presents a step-by-step guide to writing an input file, in which we use genetic toggle switch as an example. Finally, Sect. 8.2.4 remarks our achievements and gives some new perspectives.

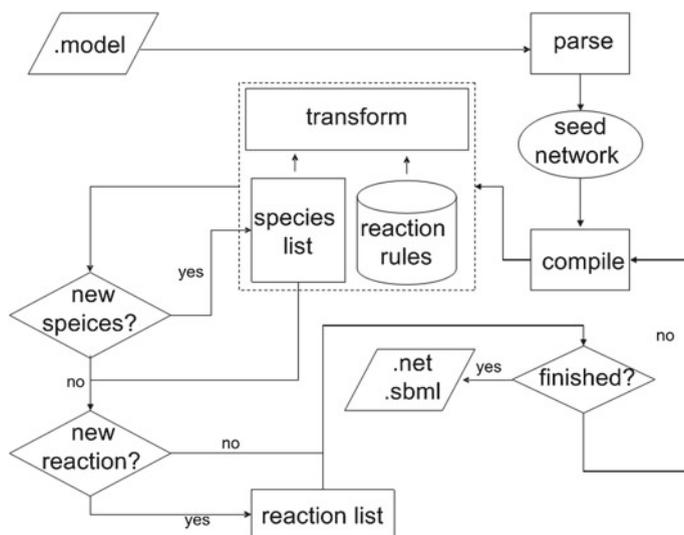


Fig. 8.1 Flow chart for the translation of a rule-based program to a network of reactions. Users construct genetic networks using our language in .model files. A .model file is then parsed to extract a network of seed species. The compiling process iteratively applies reaction rules to the seed network and generates a reaction network in both a native format and the SBML [12] format, which can be read by prevailing simulators. New species and reactions identified during transformation are added to a species list and a reaction list, respectively. Compilation is terminated either when all possible species and reactions have been identified or by enforcing some kinds of criteria, such as the number of species exceeds a maximum number limit

8.2.1 The Basics of Reaction Rules

Different from biochemical reactions, reaction rules only specify components necessary for a transformation of reactants to take place. This idea is best illustrated by functional group reactions in organic chemistry. For example, fatty acids containing a carboxyl group ($X\text{-COOH}$) can react with alcohols containing a hydroxyl group ($Y\text{-OH}$) to form esters ($X\text{-COO-Y}$). COOH and OH are reaction centers which undergo changes while X and Y are reaction contexts which are free from change. Because this chemical transformation can be ideally triggered independent of what fatty acids and alcohols are involved, a number of chemical reactions can be derived from such a single rule if we choose different X and Y groups. Therefore, a reaction rule defines a class of reactions whose reactants meet the conditions of triggering this rule.

Each reaction rule has four components: the left-hand side (LHS), the right-hand side (RHS), the rule directionality and the kinetic law. The LHS of a reaction rule specifies the conditioning of triggering this rule and the RHS specifies what changes may occur on the LHS. Patterns are used to identify species in both sides: reactant patterns define how reactant species are selected when reaction rules are

applied and product patterns specify the transformation of those selected species. Rule directionality can be either reversible (\leftrightarrow) or irreversible (\rightarrow). If a reaction rule is reversible and applied in the reverse direction, its product patterns are used to select reactant species and the results of transformation implied by this rule are reflected in the reactant patterns. The kinetic law of a reaction rule specifies at which rate all biochemical reactions derived from this rule proceed and how the reaction rate is related to the concentrations of reactant species. Both forward kinetic law and reverse kinetic law are required for reversible reaction rules.

8.2.1.1 Species

Proteins are structured objects in rule-based modeling. Each protein has a number of sites, which could bind other sites in protein-protein interactions and hold internal states for post-translational modifications. For example (see Listing 1 of [7]), EGFR has three sites: a binding site ‘L’ for its ligand EGF, a ligand-induced dimerization site ‘CR1’ and an autophosphorylation site ‘Y1068’, and would be represented as $EGFR(L, CR1, Y1068 \sim U)$, where $Y1068 \sim U$ indicates the Y1068 site is unphosphorylated. Similarly, EGF has only one receptor binding site ‘R’ and would be represented as $EGF(R)$. The ligand-receptor interaction between EGF and EGFR leads to a binding complex $EGFR(L!1, CR1, Y1068 \sim U).EGF(R!1)$, where the comma separates different components of a species and ‘!’ specifies a physical bond. The same number following ‘!’ shared by two sites indicates a bond formation between them. Note that a bond has exactly two terminal ends.

Another example is lac repressor (LacR), which is a homotetramer in structure. The tetramer has two subunits, each of which is composed of two monomers and binds to a DNA operator sequence. Each monomer has three domains that mediate its interactions with other species [18]:

- a DNA-binding domain, which binds to the lac promoter in a dimeric form;
- a regulatory domain, which binds allolactose and allolactose analogues, such as isopropyl- β -D-thiogalactopyranoside (IPTG);
- a tetramerisation domain, through which four monomers are joint.

A DNA loop is formed through the binding of two monomers of LacR to one operator sequence and the other two monomers to a second operator sequence far away from the former [21]. Each monomer of a lac dimer has a DNA-binding domain ‘dna’, an IPTG-binding domain ‘iptg’ and a dimerisation domain ‘dim’, and the lac dimer would be represented as $LacI(dna, iptg, dim!1).LacI(dna, iptg, dim!1)$. Additionally, a binding complex of IPTG and lac monomer would be similarly represented as $LacI(dna, iptg!1, dim).IPTG(laci!1)$, where a bond forms between the ‘laci’ site of the IPTG molecule and the ‘iptg’ site of the lac monomer.

DNA sequences can be viewed as compositions of individual functional parts. For example, a transcriptional unit may consist of four parts sequentially placed in order: a promoter, a ribosomal binding site (RBS), a coding sequence (CDS) and a terminator. AutoModel employs the hyphen character (‘-’) to connect one part with

its adjacent parts. Since CDS should also have their RNA and protein representations, we associate each representation with a specific character to indicate if it is a DNA, RNA, or protein. Four possible characters are ‘d’ for DNA, ‘r’ for RNA, ‘p’ for protein and ‘nb’ for molecules like IPTG (not a biological component). The representations of mRNA transcript (with RBS) and protein form of tetR gene after translation are *r:rbs()-tetr(dna,atc,dim)* and *p:tetr(dna,atc,dim)*, respectively.

Site-specific DNA inversion has been discovered in many organisms [14, 16, 17] and utilised in synthetic biology projects [10]. In the language of AutoModel, an asterisk is used to indicate a part has been inverted, such as *tetr*(dna,atc,dim)*.

8.2.1.2 Patterns and Wildcards

Patterns are similar to species in definitions but have ambiguous and incomplete specifications. The main difference between patterns and species is that patterns specify only information necessary for triggering reaction rules and can thus be taken as species selectors. For instance, *p:LacI(iptg)* is a pattern that describes any lacI monomer which has a free IPTG binding domain, while *p:LacI(dna,iptg,dim)* refers to a specific LacI monomer that contains a DNA binding domain, an IPTG binding domain and also a dimerization domain. In other words, *p:LacI(dna,iptg,dim)* is one instance of the pattern *p:LacI(iptg)* and besides lac monomer, the pattern *p:LacI(iptg)* selects any species containing lac monomer as a component as long as its IPTG binding site is free.

Using wildcards enables diverse patterns to be defined in various circumstances. The wildcard ‘+’ indicates that a match may occur at an occupied site but the identification of its binding partner is ignored. Another wildcard ‘?’, defined on the other way, indicates that a match may occur at a site regardless whether there is a bond present or not. For example, pattern *p:LacI(iptg!?)* selects both *p:LacI(dna,iptg,dim)* and *p:LacI(dna,iptg!1,dim).nb:IPTG(laci!1)* while *p:LacI(iptg!+)* only selects the latter. One advantage of using wildcards is that the writing of patterns can be simplified. Both patterns, *nb:IPTG(laci!+)* and *nb:IPTG(laci!1).p:LacI(iptg!1)*, are the same if LacI is the only binding partner of IPTG.

X parts are newly introduced to match any DNA sequences (d:X), RNA sequences (r:X) or protein (p:X). Besides used alone, each X part can be combined with other X parts or non-X parts (parts with defined sequences and functions) to form more complex patterns. The behavior of X parts can be modified by qualifiers, which is similar to the use of ‘+’ as a qualifier of ‘.’ to modify its replacement to one or more characters in regular expressions. Five qualifiers of X parts are given in Table 8.1, together with their meanings and examples of usage. The first three qualifiers are self-explanatory. Qualifier ‘>’ (‘<’) of an X part means, among all sequences that match this part, those that include its nearest part should be excluded. For example, the subsequence that matches the X part in *d:plac(laci1,laci2)-rbs()-X-term()* with regards to species *d:plac(laci1,laci2)-rbs()-tetr(dna,atc,dim)-term()-rbs()-laci(dna,iptg,dim)-term()* is underlined. However, if we replace ‘X’ with ‘X!>’ in the pattern, the underlined subsequence is not yet approved in pattern matching since it contains the part *term()*.

Table 8.1 Qualifiers of X parts

Qualifier	Meaning	Example
?	match zero or one part	<i>r:rbs()-X!?</i>
+	match one or more parts	<i>r:rbs()-X!+</i>
{n}	match exactly n part(s)	<i>r:rbs()-X!1-X2</i>
>	exclude downstream part	<i>d:term*()-X*!<-rbs*()-plac*(laci1,laci2)</i>
<	exclude upstream part	<i>d:plac(laci1,laci2)-rbs()-X!>-term()</i>

8.2.1.3 Kinetic Law

The kinetic law of a reaction rule defines the rate of a reaction. In the context of biological modeling, three most frequently-used kinetic models are the law of mass action, Hill kinetics and Michaelis-Menten kinetics. The law of mass action is concerned with the rate equations of elementary reactions. The reaction rate of equation 'aA + bB → cC + dD' is assumed to have a mathematic form $\alpha[A]^a[B]^b$, where α is a constant and the values for species enclosed by square brackets (*i.e.* A and B) are their concentrations. Hill kinetics describes the saturation of a receptor as a function of its ligand concentration. The mathematic form of this saturation is $V_m L^n / (K_{1/2}^n + L^n)$, where V_m , $K_{1/2}$, L , and n are the maximal saturation, the ligand concentration to achieve half maximal saturation, the ligand concentration and the Hill coefficient, respectively. Hill function is expected to increase slowly when $L < K_{1/2}$ and rapidly when $L > K_{1/2}$. If the Hill coefficient is very large, this function behaves like a step function. The third kinetic model is Michaelis-Menten equation, which is a special form of Hill kinetics when the hill coefficient is one. It models the enzymatic mechanism of a single substrate: $E + S \rightleftharpoons ES \rightarrow E + P$, in which E, S and P are enzyme, substrate and product, respectively. The mathematic form of a Michaelis-Menten type of reaction rate is $k_{cat} E_t S / (K_{1/2} + S)$, where the maximal rate is proportional to the total enzyme concentration E_t and the scale factor is k_{cat} .

Six candidate kinetic laws that are mostly encountered in biochemical reactions are utilised in our approach to define reaction rules (see Table 8.2). Calculating reaction rates from a formula requires that each parameter is assigned with a constant value and each variable is tied to the concentration of a particular reactant. A reaction rule having *mass_action_1(1,#1)* as its kinetic law implies the rates of selected reactions has a mathematical form ' $k * x$ ', where k is equal to one and x is the concentration of the first reactant. Similarly, '#2' means the second reactant and '#3' means the third reactant, *etc.* Parameter units in kinetic laws should be specified in a consistent manner, although they are not explicitly required.

Table 8.2 Kinetic laws used for reaction rule definitions

Kinetic law	Formula	Usage
basal rate	k	<i>basal_rate(k)</i>
MassAction-one	$k * x$	<i>mass_action_1(k,x)</i>
MassAction-two	$k * x * y$	<i>mass_action_2(k,x,y)</i>
Hill-Kinetics	$Vm * S^n / (Kp + S^n)$	<i>hill_kinetics(Vm,n,Kp,S)</i>
Henri-Michaelis-Menten	$kcat * E0 * S / (Km + S)$	<i>henri_michaelis_menten(kcat,E0,Km,S)</i>
Ordered-Bi-Bi	$kcat * E0 / (Ks/A/B + Kma/A + Kmb/B + 1)$	<i>ordered_bi_bi(kcat,E0,Ks,Kma,Kmb,A,B)</i>

8.2.2 A Short Summary: The Applications

8.2.2.1 Binding/Unbinding Interactions

Protein-DNA interactions can be very complex if multiple protein subunits bind to DNA sequences at different loci and function cooperatively. One simple situation is a homodimer specifically binds to the promoter region. For example, the binding of the lac dimer to the lac promoter in IPTG-free context would be represented as

```
d:X1-plac(laci1,laci2)-X2 + p:LacI(dna, iptg, dim!1) .p:LacI(dna, iptg, dim!1) <->
d:X1-plac(laci1!2, laci2!3)-X2.p:LacI(dna!2, iptg, dim!1) .p:LacI(dna!3, iptg, dim!1)
mass_action_2(1e8, #1, #2) mass_action_1(1e-2, #1)
```

where parameters in both forward kinetic law and reverse kinetic law are arbitrarily chosen but within a reasonable range (the same below). Besides, we can explore the possibility of other regulatory mechanisms. One such mechanism is that the lac monomer is able to bind to one of the two sites of the lac promoter to form reaction intermediates. Since the lac promoter has two distinct protein binding sites, two other reaction rules,

```
d:X1-plac(laci1)-X2 + p:LacI(dna, iptg, dim) <-> d:X1-plac(laci1!1)-X2.p:LacI
(dna!1, iptg, dim) mass_action_2(1e6, #1, #2) mass_action_1(1e-2, #1)
```

and

```
d:X1-plac(laci2)-X2 + p:LacI(dna, iptg, dim) <-> d:X1-plac(laci2!1)-X2.p:LacI
(dna!1, iptg, dim) mass_action_2(1e6, #1, #2) mass_action_1(1e-2, #1)
```

should be added. The use of small rate constants implies that the lac dimer is favored over the lac monomer for binding to the lac promoter. Reaction rules for heterodimers and other protein complexes binding to the promoter regions can be written in similar fashions.

8.2.2.2 Catalysis

Enzyme catalysis is a highly efficient mechanism that is instrumental for cellular activities. As we have discussed in Sect. 8.2.1.3, this type of interaction is usually

modeled by a two-step cascade reaction $E + S \rightleftharpoons ES \rightarrow E + P$. However, it can be reduced to a one-step reaction $E \rightarrow E + P$ if the amount of the substrate is very large. This is because the reaction rate is approximated to $k_{cat}E_t$ if $S \gg K_{1/2}$. The synthesis of 3OC6HSL by LuxI protein would be represented as

```
@p:LuxI() -> nb:3OC6HSL(luxr) mass_action_1(0.45,#1)
```

where we use '@' to label catalysts. Not only protein, but mRNA is able to catalyze biochemical processes. The Isaacs' riboregulator trans-activating key taR12 and cis-repressed lock crR12 are used together to regulate the mRNA-level genetic expression [13]. The RBS of crR12, which is initially blocked by a stem-loop structure, becomes exposed at the presence of taR12. The state change of the crR12's RBS can be described by the reaction rule

```
@r:taR12() + r:X1-crR12(rib~off)-X2 -> r:X1-crR12(rib~on)-X2
henri_michaelis_menten(7.6,#1,9e-4,#2)
```

which is a simplified version of the cascade of two reaction rules (in which kinetic laws are omitted)

```
r:taR12(crr12) + r:X1-crR12(tar12,rib~off)-X2 <-> r:X1-crR12(tar12!1,rib~off)
-X2.r:taR12(crr12!1)-> r:taR12(crr12) + r:X1-crR12(tar12,rib~on)-X2
```

8.2.2.3 Transcription and Translation

Transcriptional and translational rules have to be defined with the aid of X parts. A class of transcriptional reactions driven by the *lac* promoter can be defined by the reaction rule

```
@d:X1-plac(laci1,laci2)-X2!>-term()-X3 -> r:X2 mass_action_1(0.8,#1)
```

in which we assume transcription ends at the first terminator completely so the use of the qualifier '>' excludes the case of producing wrong transcripts when this rule is applied to DNA sequences containing more than one terminators. For example, the only mRNA transcript of the DNA sequence *d:plac(laci1,laci2)-rbs()-tetr(dna,atc,dim)-term()-rbs()-laci(dna,iptg,dim)-term()* is *r:rbs()-tetr(dna,atc,dim)*.

However, most terminators do not terminate transcriptions with 100% efficiency [20]. If the termination efficiency of the terminator *term()* is less than one, the DNA sequence may be transcribed to produce another mRNA transcript *r:rbs()-tetr(dna,atc,dim)-term()-rbs()-laci(dna,iptg,dim)* in which case RNA polymerase reads through the terminator. To take into account the incomplete transcriptional termination, we may add a similar reaction rule by removing the qualifier '>' of the part X2:

```
@d:X1-plac(laci1,laci2)-X2-term()-X3 -> r:X2 mass_action_1(0.1,#1)
```

If we apply both rules to the DNA sequence, the concentration ratio of the shorter transcript to the longer transcript is 9:1 and the termination efficiency of the terminator

is thus 90%. Different termination efficiencies can be obtained by balancing the reaction rates of the two rules.

On the other hand, translational rules are defined by replacing the open reading frame (ORF) with an X part which only selects one mRNA part:

```
@r:X1-rbs()-X2!1-X3 -> r:X2 mass_action_1(0.1,#1)
```

However, a translation of fused proteins may require developing new parts like start codon and stop codon, through which the length of an ORF can be determined.

8.2.2.4 Protein/mRNA Degradation

Protein and mRNA degradation are key for organisms to stabilize gene expression level. Each protein has a specific degradation rate and thus a specific reaction rule for its degradation. For example, the reaction rule for *lac* monomer degradation is

```
p:LacI(dna, iptg, dim) -> (nothing) basal_rate(5e-3)
```

Moreover, we know proteins can be partially degraded within protein-protein complexes or protein-DNA complexes. One approach of specifying partial protein degradations is replacing individual proteins with protein complex patterns in the reaction rule shown above. Each of the three reaction rules

```
p:LacI(dna!+, iptg, dim) -> (nothing) basal_rate(5e-3)
p:LacI(dna, iptg!+, dim) -> (nothing) basal_rate(5e-3)
p:LacI(dna, iptg, dim!+) -> (nothing) basal_rate(5e-3)
```

defines a class of protein degradation reactions. They are partial degradations of a *lac* monomer within its binding complexes containing a *lac* promoter, an IPTG molecule and another *lac* monomer, respectively.

Different from protein degradation, degradation rates are generally unknown for most polycistronic mRNAs. Although we could relate a mRNA pattern with a specific degradation rate, a more practical way is assuming mRNA degradation has a global rate and all mRNAs degrade at the same rate. The reaction rule for mRNA degradation under this assumption can be simply written as:

```
r:X -> (nothing) basal_rate(0.1)
```

8.2.2.5 Site-Specific DNA Recombination

Site-specific DNA recombination is a natural mechanism for rearrangements of DNA segments in living cells. Typical recombination systems have only one recombinase enzyme and a pair of recombination sites. One famous example is the Cre-lox system [23, 24]. The outcome of a Cre-lox recombination depends on the configuration of the flanking loxP sites. If the two loxP core sequences are oppositely oriented on the same DNA segment, Cre recombinase mediates the inversion of the flanked DNA segment:

```
@p:Cre() + d:X1-loxP()-X2-loxP*()-X3 -> d:X1-loxP()-X2*-loxP*()-X3
henri_michaelis_menten(7.6, #1, 9e-4, #2)
```

Besides, if the two loxP sites are located on different DNA segments, a Cre-mediated DNA segment translocation may occur:

```
@p:Cre() + d:X1-loxP()-X2 + d:X3-loxP()-X4 -> d:X1-loxP()-X4 + d:X3-loxP()
-X2ordered_bi_bi(7.6, #1, 9e-4, 3e-2, 3e-2, #2, #3)
```

where we assume this process is controlled by a two-substrate enzyme kinetic model. Finally, if the two loxP sites are oriented in the same direction on the same DNA segment, the flanked gene is deleted by Cre recombinase and a circularised DNA composed of this deleted gene and a loxP site is then formed.

8.2.3 Model Specification: Genetic Toggle Switch

A genetic toggle switch consists of two repressors and two promoters, each of which is inhibited by the repressor transcribed by the other promoter [8]. We choose *LacR*, *TetR* as the two repressors and lac promoter, tet promoter as the two promoters, correspondingly. The network diagram is drawn in Fig. 8.2 using TinkerCell [3]. Since *LacR* has a much higher binding affinity to lac promoter than that between *TetR* and tet promoter, a strong binding of *LacR* to the lac promoter would repress the transcription of the tetR gene and thus the amount of *TetR* will decrease, which in turn relieves its repression to the tet promoter and increases the amount of *LacR*. As a result, *LacR* is at a high expression level while *TetR* is at a low expression level. A certain amount of IPTG binding to *LacR* would relieve the repression of the lac promoter. Similarly, the toggle switch transits to another state at which *LacR* is at a low expression level while *TetR* is at a high expression level. To model this transition, we write a rule-based program by specifying four blocks: parameters, compartments, seedspecies and events, each of which describes one particular aspect of the toggle switch. Reaction rules for the lac and tet systems are specified in databases. By submitting the program to our software, a reaction network of the toggle switch, which is encoded in the SBML format, is generated within one second. The network generated automatically by AutoModel has 15 species and 33 reactions. We then simulate the network using COPASI [11] and the time course simulation of the network shows expected dynamics (see Fig. 8.3). We will show below how to construct a toggle switch model in our approach step by step.

8.2.3.1 Step 1: Specify Parameters

This block defines global parameters that can be used in other blocks. The syntax of parameter definition is `<name value>`. The *name* element gives each parameter a unique name and the value of each parameter is given by the *value* element, which can be either a number or a formula involving several other parameters. The order

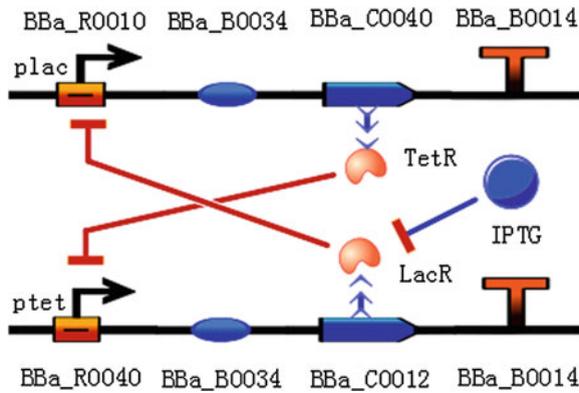


Fig. 8.2 Network diagram of the genetic toggle switch system. The CDS part BBa_C0040 (tetR gene) codes for TetR protein and the promoter part BBa_R0040 (tet promoter) is negatively regulated by this protein. Similarly, the CDS part BBa_C0012 (lacI gene) codes for LacR protein and the promoter part BBa_R0010 (lac promoter) is negatively regulated by this protein. The part BBa_B0034 is an RBS and the part BBa_B0014 is a terminator. All the parts shown here are from the MIT registry

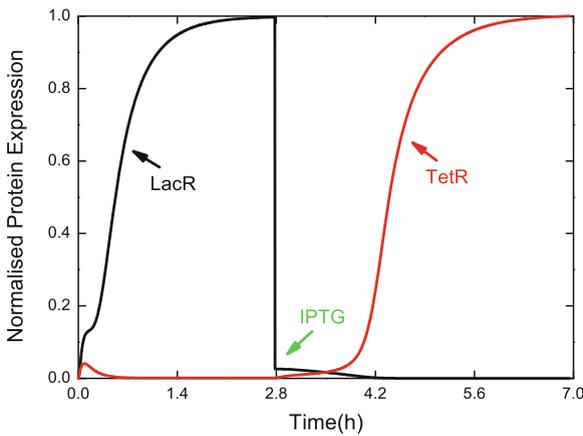


Fig. 8.3 Time course simulation of the genetic toggle switch model. At the beginning of the simulation, LacR is highly expressed while TetR is strongly repressed. A phase transition is observed about 1.4h after IPTG is added. Then, the expression of TetR goes up to a high level while the expression of LacR goes down to a low level. Here LacR and TetR are referred to the protein products of the lacI gene and the tetR gene, respectively

parameters are defined is not restricted and parameter units should be specified consistently. Here is the parameters block for the toggle switch model (comments start with a hash sign ('#')):

```
<parameters>
NA 6.02e23 # Avogadro
f 1
```

```

Ve f*0.4 # L
V f*7e-16 # L
N 100
iptg_init 0.0
dna1_init 1/NA/V # M
dna2_init 1/NA/V # M
pulse1_start 10000 # sec
pulse1_conc 1e-3 # M
</parameters>

```

8.2.3.2 Step 2: Specify Compartments

A compartment is a three-dimensional enclosed space where species locate in and usually takes a correspondence to some type of cells. The syntax of compartment definition is `<name outside table [volume] [population]>`. Its name and space volume are given by *name* and *volume*, respectively. Our approach does not take into account the cell growth and division but instead uses a fixed cell number, whose value is given by the *population* element. Both *volume* and *population* are optional (delimited by square bracket []) and the default value for both is one.

Topological relationships between compartments can be simply defined by specifying the *outside* element of each compartment, which should be the name of another compartment surrounding it. The hierarchically organized compartments have a tree structure, where each node, except the root, is allowed to have only one parent (outside compartment) and any number of children (inside compartments). The *outside* element of a root compartment is 'ROOT'.

A compartment is essentially a virtual boundary that confines the movement of species within it. The space enclosed by the boundary significantly reduces the possibility that molecules from different compartments collide and it is thus likely to isolate inside biochemical reactions from the outside. For this reason, we associate each compartment with a distinct set of reaction rules which are stored in a table whose name is given by the *table* element. The reaction rules for a particular compartment specify how the species in that compartment transform and the compartment develops its own seed network in parallel with other compartments.

Below is an example for *Escherichia coli* (*E. coli*) cells 'Ecoli' in the medium being an outside compartment 'Medium'. A set of reaction rules for 'Ecoli' can be found in the table named 'toggle_switch' in the database (see Sect. 8.2.3.5). The table 'medium' is empty because no reactions are expected to occur outside *E. coli* cells. Note volumes of both compartments and the number of *E. coli* cells have been defined in the parameters block.

```

<compartments>
Medium ROOT medium Ve
Ecoli Medium toggle_switch V N
</compartments>

```

8.2.3.3 Step 3: Specify Seedspecies

All species defined in this block constitute a seed network. The syntax of seed species definition is `<compartment name structure [init_conc] [is_const]>`. For each species, we need to specify *compartment* for its location, *name* for its name, *structure* for its molecular structure (see Sect. 8.2.1.1), *init_conc* for its initial concentration and *is_const* for if its concentration is kept constant. The *is_const* element can be either ‘True’ or ‘False’. If not specified, the *is_const* element has a default value ‘False’ and the *init_conc* element has a default value 0.0.

We assume single copy of each expression cassette shown in Fig. 8.2. A single-copy plasmid has an initial concentration $1/NA/V$, in which NA is the Avogadro’s number and V is the volume of the compartment ‘Ecoli’. The initial concentrations of both expression cassettes, ‘dna1_init’ and ‘dna2_init’, have been defined in the parameters block. The seedspecies block of the toggle switch model is given as follows:

```
<seedspecies>
Medium iptg nb:i0001(laci) iptg_init
Ecoli dna1 d:r0040(tetr1,tetr2)-b0034()-c0012(dna, iptg, dim)-b0014() dna1_init
Ecoli dna2 d:r0010(laci1,laci2)-b0034()-c0040(dna, atc, dim) -b0014() dna2_init
</seedspecies>
```

8.2.3.4 Step 4: Specify Events

An event in AutoModel is defined to change variables, which can be compartment volumes, seed species concentrations, and parameter values, instantly at a certain time point. The syntax of event definition is `<name trig_cond event_assignment_1 event_assignment_2 ...>`. The *trig_cond* element specifies the conditioning of an event to take place and the event is executed every time *trig_cond* turns from false to true during simulation. The effect of an event is specified in a list of event assignment elements, *event_assignment_1*, *event_assignment_2*, etc. Each event assignment has a syntax `<variable=expression>`, in which the expression can be a formula involving other defined variables. The only event for the toggle switch system is adding IPTG of amount ‘pulse1_conc’ to the medium at the time ‘pulse1_start’. The keyword ‘time’ used here is to track the simulation time.

```
<events>
pulse1 time>pulse1_start iptg=pulse1_conc
</events>
```

8.2.3.5 Step 5: Specify Reaction Rules

Two assumptions are made for the writing of reaction rules: (1) only dimers (both the lac dimer and the tet dimer) can bind to the promoter regions while the monomers cannot; (2) IPTG can bind to the lac monomer in whatever complexes as long as the DNA-binding site of the monomer is not occupied. Since we have given an

in-depth look at how to represent typical biological interactions using reaction rules in Sect. 8.2.2, we are not going to repeat it again but only list the reaction rules for the toggle switch system below.

```
#dimerisation of the tet monomer
p:c0040(dim,dna) + p:c0040(dim,dna) <-> p:c0040(dim!1,dna).p:c0040(dim!1,dna)
mass_action_2(1.79e7,#1,#2) mass_action_1(10,#1)

#tet dimer binds to the tet promoter
p:c0040(dna,atc,dim!1).p:c0040(dna,atc,dim!1) + d:X1-r0040(tetr1,tetr2)-X2 <->
p:c0040(dna!1,atc,dim!3).p:c0040(dna!2,atc,dim!3).d:X1-r0040(tetr1!1,tetr2!2)
-X2 mass_action_2(1e8,#1,#2) mass_action_1(1e-2,#1)

#IPTG binds to the lac monomer
p:c0012(iptg,dna) + nb:i0001(laci) <-> p:c0012(iptg!1,dna).nb:i0001(laci!1)
mass_action_2(1e6,#1,#2) mass_action_1(0.2,#1)

#dimerisation of the lac monomer
p:c0012(dim,dna) + p:c0012(dim,dna) <-> p:c0012(dim!1,dna).p:c0012(dim!1,dna)
mass_action_2(1.25e7,#1,#2) mass_action_1(10,#1)

#lac dimer binds to the lac promoter
p:c0012(dna,iptg,dim!1).p:c0012(dna,iptg,dim!1) + d:X1-r0010(laci1,laci2)-X2
<-> p:c0012(dna!1,iptg,dim!3).p:c0012(dna!2,iptg,dim!3).d:X1-r0010(laci1!1,
laci2!2)-X2 mass_action_2(2e10,#1,#2) mass_action_1(0.04,#1)

#transcription by the lac promoter
@d:X1-r0010(laci1,laci2)-X2!>-b0014()-X3 r:X2 mass_action_1(0.5,#1)

#leaky transcription by the lac promoter
@d:X1-r0010(laci1!+,laci2!+)-X2!>-b0014()-X3 -> r:X2 mass_action_1(0.0005,#1)

#transcription by the tet promoter
@d:X1-r0040(tetr1,tetr2)-X2!>-b0014()-X3 -> r:X2 mass_action_1(0.5,#1)

#leaky transcription by the tet promoter
@d:X1-r0040(tetr1!+,tetr2!+)-X2!>-b0014()-X3 -> r:X2 mass_action_1(0.0005,#1)

#translation
@r:X1-b0034()-X2!1-X3 -> p:X2 mass_action_1(0.01155,#1)

#mRNA degradation
r:X -> (nothing) mass_action_1(0.005783,#1)

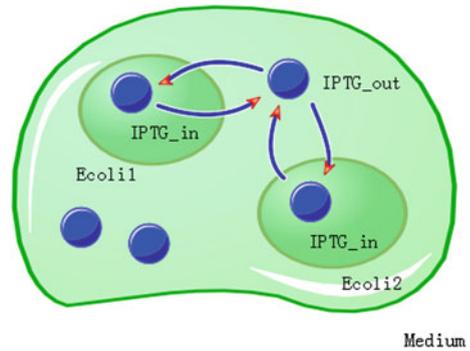
#(partial) degradation of the lac monomer
p:c0012(dim) -> (nothing) mass_action_1(2.31e-3,#1)

#(partial) degradation of the tet monomer
p:c0040(dim) -> (nothing) mass_action_1(2.31e-3,#1)
```

Additionally, IPTG molecules can diffuse through the boundary of compartments (see Fig. 8.4). Transport reactions of IPTG are modelled by two coupled ordinary differential equations (ODEs) [19]:

$$\begin{aligned}\frac{d[IPTG_{in}]}{dt} &= k_{in}[IPTG_{out}] - k_{out}[IPTG_{in}] \\ \frac{d[IPTG_{out}]}{dt} &= N_{cell}(k_{out}[IPTG_{in}] - k_{in}[IPTG_{out}])\end{aligned}$$

Fig. 8.4 Transport of IPTG molecules through cell membrane



where $[IPTG_{in}]$ and $[IPTG_{out}]$ are the IPTG concentrations in the compartment 'Ecoli' and 'Medium', respectively. N_{cell} is the number of *E. coli* cells and both k_{in} and k_{out} are 0.1.

8.2.4 Remarks

We have developed a new rule-based approach for modeling synthetic biological systems termed AutoModel. Reaction rules are introduced to represent various types of biological interactions. We have made a small step forward towards two goals: (1) high-level specification of contextual information for biological interactions and (2) high degree of automation in the modeling process. The first goal is achieved by explicitly modeling the behavior of functional components of a species rather than the behavior of the species as a whole. The functional components include but not limited to DNA-binding domains, protein-binding domains, phosphorylation domains, operator sequences within promoters, RBS, CDS and terminators. The second goal is achieved by separating the specification of reaction rules from the rest part of the model specification. A database of reaction rules for the frequently utilised mechanisms in constructing genetic devices, such as the lac system, the tet system, the lux system and the arabinose system, may be created for the public use. This makes biological models accessible to non-specialists, who are then able to obtain insights into some basic biological systems.

However, there is much room for our approach to improve. First of all, the selection capabilities of X parts can be further extended. For example, we may want to specify a DNA sequence which includes tetR gene but not lacI gene. We anticipate a logic expressoin like $X(lacI \text{ AND } (NOT \text{ tetR}))$ can be developed to provide these capabilities in the future. Additionally, our language is not formal yet so analysis like model checking cannot be performed on the models in the frameworks of both formal language theory and automata theory. Other focuses may include graphical user interface (GUI), web service and interactions with other CAD tools. By

developing a powerful and practical software tool, we can increase the ability for precise quantification of biological systems, which is the ultimate goal of the whole biology community.

References

1. Cai Y, Hartnett B, Gustafsson C et al (2007) A syntactic model to design and verify synthetic genetic constructs derived from standard biological parts. *Bioinformatics* 23:2760–2767
2. Cai Y, Lux MW, Adam L et al (2009) Modeling Structure-Function Relationships in Synthetic DNA Sequences using Attribute Grammars. *PLoS Comput Biol*. doi:[10.1371/journal.pcbi.1000529](https://doi.org/10.1371/journal.pcbi.1000529)
3. Chandran D, Bergmann FT, Sauro HM (2009) TinkerCell: modular CAD tool for synthetic biology. *J Biol Eng* 3:19
4. Danos V, Laneve C (2004) Formal molecular biology. *Theor Comput Sci* 325:69–110
5. Desmit MH, Vanduin J (1990) Secondary structure of the ribosome binding-site determines translational efficiency—a quantitative-analysis. *Proc Natl Acad Sci USA* 87:7668–7672
6. Desmit MH, Vanduin J (1994) Control of translation of messenger-RNA secondary structure in *Escherichia coli*—a quantitative analysis of literature data. *J Mol Biol* 244:144–150
7. Faeder JR, Blinov ML, Hlavacek WS (2009) Rule-based modeling of biochemical systems with BioNetGen. *Methods Mol Biol* 500:113–167
8. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342
9. Gebauer F, Hentze MW (2004) Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* 5:827–835
10. Haynes KA, Broderick ML, Brown AD et al (2008) Engineering bacteria to solve the burnt pancake problem. *J Biol Eng* 2:8
11. Hoops S, Sahle S, Gauges R et al (2006) COPASI—a COMplex PATHway SIMulator. *Bioinformatics* 22:3067–3074
12. Hucka M, Finney A, Sauro HM et al (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524–531
13. Isaacs FJ, Dwyer DJ, Ding C et al (2004) Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* 22:841–847
14. Klippel A, Mertens G, Patschinsky T et al (1988) The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA via a phosphoserine at amino acid position 9. *EMBO J* 7:1229–1237
15. Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324:255–258
16. Kutsukake K, Nakao T, Iino T et al (1985) A gene for DNA invertase and an invertible DNA in *Escherichia coli* K-12. *Gene* 34:343–350
17. Kutsukake K, Nakashima H, Tominaga A et al (2006) Two DNA invertases contribute to flagellar phase variation in *Salmonella enterica* serovar Typhimurium strain LT2. *J Bacteriol* 188:950–957
18. Lewis M (2005) The lac repressor. *C R Biol* 328:521–548
19. Melke P, Sahlin P, Levchenko A et al (2010) A cell-based model for quorum sensing in heterogeneous bacterial colonies. *PLoS Comput Biol*. doi:[10.1371/journal.pcbi.1000819](https://doi.org/10.1371/journal.pcbi.1000819)
20. Nojima T, Lin AC, Fujii T et al (2005) Determination of the termination efficiency of the transcription terminator using different fluorescent profiles in green fluorescent protein mutants. *Anal Sci* 21:1479–1481
21. Oehler S, Eismann ER, Krämer H et al (1990) The three operators of the lac operon cooperate in repression. *EMBO J* 9:973–979

22. Pedersen M, Phillips A (2009) Towards programming languages for genetic engineering of living cells. *J R Soc Interface* 6:S437–S450
23. Sauer B, Henderson N (1988) Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci USA* 85:5166–5170
24. Sternberg N, Hamilton D (1981) Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *J Mol Biol* 150:467–486

Chapter 9

An Investigation of Signal Transduction and Irreversible Decision Making Through Monostable and Bistable Switches

J. Krishnan and C. Liu

Abstract Highly nonlinear signal transduction is ubiquitous in cell signalling pathways with switch-like behaviour encountered repeatedly. Monostable and bistable switches represent distinct basic switches which are encountered. A number of contexts in cellular signalling (e.g. apoptosis and cell cycle) involve essentially irreversible transitions and decision making. In this article we examine signal transduction through prototypical monostable and bistable switches with a view towards understanding how irreversible signal transduction may occur through them and also examine the similarities and differences in signal transduction and decision making to classes of experimentally employed inputs. The study provides insights into how irreversible transitions may be orchestrated through different switches in cell signalling, the underlying design characteristics, capabilities and constraints involved, and the extent to which these switches can be distinguished based on irreversible decision making to experimentally available classes of inputs.

Keywords Signal transduction · Monostable · Bistable · Cellular signaling

9.1 Introduction

Cells respond to their environment and also regulate different aspects of their internal organization by highly complex signal transduction and gene regulatory networks. Such networks allow for cells to respond to multiple cues in their environment in an appropriate fashion. Examples of cellular responses include cell movement, programmed cell death (apoptosis), mitigating the effects of stresses and

J. Krishnan (✉) · C. Liu

Chemical Engineering and Chemical Technology,
Centre for Process Systems Engineering and Institute for Systems and Synthetic Biology,
Imperial College London, South Kensington Campus, London SW7 2AZ, UK
e-mail: j.krishnan@imperial.ac.uk

communicating with surrounding cells. Other aspects of the internal organization of cells such as the cell cycle and cell division, to name just two examples, are also robustly orchestrated through complex signalling and gene regulatory networks.

In order for signal transduction networks to perform their tasks and maintain cellular function and allow for robust decision making, highly non-linear signal transduction is necessary. Switch-like behaviour is a frequently observed signalling characteristic in signal transduction networks. This allows for strong responses to be elicited only when the stimulus strength is sufficiently high, whereupon a substantial difference in response is observed. Switching behaviour may be obtained in cellular signalling pathways through a multitude of mechanisms, including co-operative effects, zeroth-order ultrasensitivity [5], combined effects in a cascade [7], sequestration [13], multiphosphorylation [9, 14] and positive feedback [4]. Broadly, however, there are two classes of switching behaviour which are observed: monostable switches wherein the system possesses a single stable steady state for a fixed value of input and bistable (or more generally multistable) switches where the system possesses more than one stable steady state for a fixed value of input, which forms the basis for the switching behavior. Multiple combinations of factors can give rise to both kinds of switching behaviour, sometimes in the same pathways. Both kinds of switches have been investigated in considerable detail through modelling and experiments, both in specific contexts and also more generally.

A specific feature of signal transduction in cellular networks is that in specific instances it can be essentially irreversible. Examples include apoptosis leading to an irreversible decision of killing, or in the cell cycle, wherein essentially irreversible transitions occur for the cell to progress through different phases of the cell cycle. Others include decisions made by stem cells to proliferate or differentiate. All these processes are robustly orchestrated by networks to make these essentially irreversible changes. Irreversible changes can also occur in signalling where individual molecular species are targeted for degradation. It is thus important to understand how networks and pathways are organized and primed for realizing such irreversible changes appropriately and robustly.

Irreversible input-output behaviour is observed in dynamical systems with multiple steady states (or multiple attractors generally), and this has been the basis of a large body of modelling wherein irreversible signal processing and decision making in networks are modelled by bistable or multistable networks. Examples of this include modelling the apoptosis pathway as a bistable network, based on evidence of positive feedback in the pathways [3, 17], and modelling irreversible decision making in other processes similarly [10, 11, 16]. However in cellular systems, an irreversibility may result in a pathway from a new irreversible transition occurring when a particular threshold is crossed. Thus in this case cells may make use of natural machinery available to them to induce an essentially irreversible transition, such as cutting of the DNA. In some cases, the presence of multistationarity *in vivo* has been demonstrated, while in other cases it is unclear if it is actually present. Nevertheless it has been invoked to particularly capture the irreversible feature of signal transduction.

We thus see that there are qualitatively distinct ways in which irreversible transitions may be orchestrated, and it is of importance to understand the similarities and differences between them. Since there are substantial differences in the underlying signalling even qualitatively, it is important to identify the characteristics of each type of switch so that the modelling of such behaviour in concrete contexts is performed on a qualitatively correct basis. A particular challenge is the fact that experimentally only limited classes of inputs are used to elucidate signalling in many systems.

Motivated by such considerations, we investigate switch-like and irreversible decision making by examining simple, essentially minimal, models relevant to the issue at hand. We consider a series of four models which involve a switch-like process connected to a decision-module. Two possibilities for the switch are examined: a monostable and a bistable switch. Two possibilities for the decision-module are also considered, a reversible module and an irreversible module. A decision is taken when a key threshold is crossed in the decision-module. In this manner we obtain 4 modules, each of which has essential qualitative differences. It is worth pointing out that every one of these configurations has been invoked either explicitly or implicitly in specific contexts.

We examine these modules with respect to the conditions under which particular decisions are taken. Thus, we will focus on an essentially binary aspect of the signalling through the network: whether the particular decision is taken or not. Further, we examine these modules through classes of inputs (or extensions thereof) which are experimentally observed. The inputs we consider are step inputs, single pulse inputs of different intensity and duration, double pulse inputs and pulse trains of different strength, duration and spacing. We examine the response of these modules to understand where the similarities and differences lie.

The chapter is organized as follows. In Sect. 9.2 we discuss the models which are employed. The following section presents the results, and the final section summarizes and synthesizes the results and discusses insights and relevance to cell signalling.

9.2 Models and Methods

We now present the models which we base our investigation on. Since the focus is on monostable and bistable switches and reversible and irreversible signal transduction, we use appropriate models which focus on these essential aspects. At the outset we note that there are multiple factors which can give rise to monostable switches in signalling pathways: these include co-operativity, gradual combinations of steps in a cascade (e.g. MAPK cascades), zeroth order ultrasensitivity, sequestration/molecular titration and multisite phosphorylation. Likewise bistability typically results from strong positive feedback, which could include double negative feedback. Naturally there are different models for monostable switches, and also different model variants for bistable switches. This could be also generated via mechanisms such as multisite

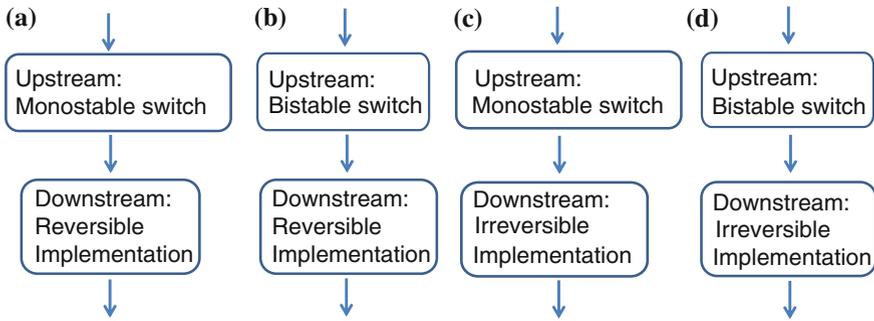


Fig. 9.1 Schematic of combinations of upstream and downstream modules. The models investigated are **a** A monostable reversible model **b** A bistable reversible model **c** A monostable irreversible model **d** A bistable irreversible model

phosphorylation which does not have an explicit positive feedback at the level of network components [9, 14].

Since our focus is to examine representative aspects of signal transduction involving monostable and bistable switches and irreversibility in particular, we employ relatively simple and tractable models of monostable and bistable switches. We are mainly concerned with the input-output behaviour of these pathways and thus the details of the models for the most part do not play any important role in our analysis. More complex switches with additional features encountered in specific contexts may possibly exhibit some differences, but this may then be understood both in terms of how generic switches function and the additional effects which are present there.

In order to focus on the issues of interest, we will focus on 4 models (see Fig. 9.1). The models involve a sequential interconnection between a switch module (either monostable or bistable) and a decision module (either reversible or irreversible). The decision module gives rise to a specific decision when a threshold in this module is crossed. Thus the 4 different designs we examine are a monostable-reversible model, a bistable reversible model, a monostable-irreversible model and a bistable-irreversible model. Examining these models together allows us to elucidate difference and similarities between monostable and bistable models with and without irreversibility.

The monostable module which we employ involves a simple feedforward Hill-type nonlinearity: the Hill coefficient determines the sharpness of the switch. The model describes the conversion of one species X (present in excess) into another X^* via an upstream signal S . The governing equation is

$$dX^*/dt = k(S^n/(A^n + S^n) - X^*) \quad (9.1)$$

When this switch is very sharp, it may be well approximated by a step function at the location $S = A$, giving rise to the equation

$$dX^*/dt = k(H(S - A) - X^*) \quad (9.2)$$

The monostable module has two parameters in addition to the Hill coefficient: A which sets the location of the threshold, and k which sets the timescale of the reaction. We will perform simulations with the Hill model, and analyze the case of the sharp switch above.

The bistable module, on the other hand is described by a simple model which captures all the features of bistability and which is tractable (also see Seaton and Krishnan [12]):

$$\begin{aligned} dX^*/dt &= -\alpha X^*, a + bS < 0 \\ dX^*/dt &= \alpha X^*(1 - X^*)(X^* - 1 + a + bS), 0 < a + bS < 1 \\ dX^*/dt &= \alpha(1 - X^*), a + bS > 1 \end{aligned} \quad (9.3)$$

The bistable module has an analogous parameter α which sets the time scale of the reaction. In addition it has two parameters a which sets the basal condition of the switch and b which determines at what signal intensity, the switch is triggered. Note that in this model, we will assume both a and b are positive. This means that $a + bS$ is always positive. Since we wish to compare the two switches on an essentially equal footing, we choose parameters so that the switching thresholds are at the same location. Thus in the monostable module, we choose $A = 1$ and choose a high Hill coefficient of about 10. This ensures that a stimulus strength of 1 (in non-dimensionless terms) results in switching. In the bistable switch, the parameter a is chosen so that the basal condition is well within the bistable region ($a < 1$). Note that the signal strength is strictly non-negative and that with positive a , the basal state is one in the bistable regime. Since the intrinsic threshold in the bistable switch is when $a + bS = 1$, a choice of $b = 1 - a$ ensures that the switching location is also $S = 1$. Now the parameters α and k in the above models are chosen so as to represent comparable time scales in the two modules.

It is worth pointing out that our models, which are essentially minimal models of monostable and bistable switches have identical amplitudes (they switch from 0 to 1) and identical switching thresholds, and comparable time scales.

These models are sequentially “connected” to a decision module

$$dR^*/dt = k_f X^*(1 - R^*) - k_r R^* \quad (9.4)$$

Here the output of the upstream module catalytically drives the downstream reaction. Implicitly it is assumed that a negligible amount of X^* is taken up in this conversion process, and that this reaction may be described by mass action kinetics. In other words there is negligible retroactivity. This is a very reasonable assumption since (1) The aspect of holdup is somewhat tangential to the focus of our investigations and (2) We are in any case employing simplified descriptions which implicitly coarse-grain many features. Ultimately our models captures the essential aspects of sequential signalling and information flow leading to decision making.

In the decision module, a decision is taken when R^* crosses a threshold R_o (since R^* cannot exceed 1, $R_o < 1$). If this happens in an irreversible fashion, this results in the irreversible decision module, whereas if it happens in a reversible fashion, this is a reversible decision module. Even though biochemical reactions in a cell are typically reversible, in specific instances crossing a threshold can initiate a series of events/transitions before the system reaches any steady state which is essentially irreversible. In other cases the nature of the decision is clearly irreversible (eg cell death). We thus see that the decision modules have the same threshold and only differ in whether they are by themselves capable of reversible or irreversible decisions. Sample parameters for this module are chosen so as to ensure that the threshold can indeed be reached through the driving upstream module.

Our focus will be to examine the behaviour of each of the four modules. Further, we will focus on the decision making in each of the 4 modules and in particular whether an irreversible decision (which we also refer to as a transition) is triggered. In particular we will examine the behaviour of these modules to typical well-characterized stimuli similar to those available in experiments. These are (1) Step inputs (2) Single (rectangular) pulses of differing amplitude and time duration (3) Double rectangular pulses where the time duration between pulses is an additional parameter and (4) Rectangular pulse trains. Simulations are performed in MATLAB.

Our investigation will thus yield insights into signal transduction and irreversible decision making in these modules subject to classes of well characterized inputs. This will also reveal to what extent each of these modules is similar and under which conditions, differences might arise.

9.3 Results

In this section we examine the dynamic response of our models to various inputs, analyze the model behaviour and examine the effects of model parameters and also briefly examine similar signal processing in more complex models. We start by examining the dynamic behaviour of the models in detail.

9.3.1 Signalling Dynamics

We first examine the behaviour of all our models to a series of inputs of increasing complexity. Parameter values in the models are fixed for this purpose. The only parameters which will be varied are those associated with the signal. The effects of model parameters are discussed later on in this section. The input signals considered are (i) Step input, with the size of the step being the associated parameter (ii) A rectangular pulse input which has two parameters, the height of the pulse and the width of the pulse (iii) A double rectangular pulse, which involves two rectangular pulses of identical structure separated by a time interval, which results in an additional

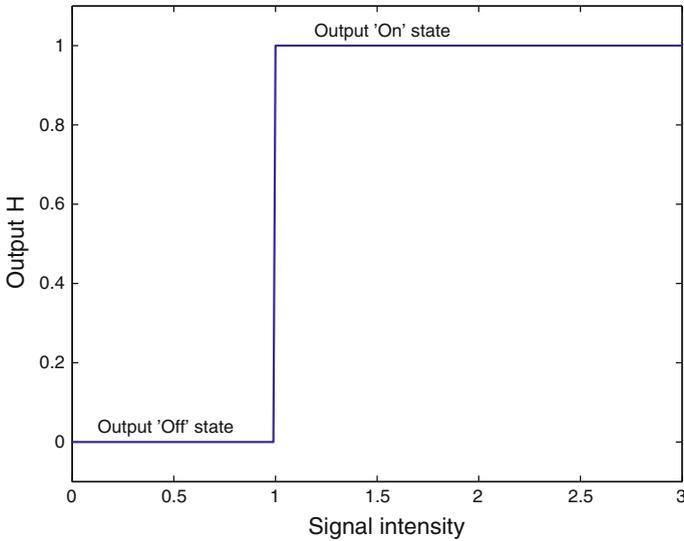


Fig. 9.2 Response of models to step signal. Output response under constant signal input for all models. A similar behaviour is seen for all models

parameter (iv) A pulse train of identical pulses which is characterized by the pulse height, width and spacing. The focus in our investigation is on the decision making in the models, which is essentially of a binary nature. In the case of the first three inputs a steady asymptotic state is indeed attained, while that is not necessarily the case for the last input. Our focus will be on when the output of the models corresponds to a transition being triggered. In other words we will examine when this binary switching decision is made.

Step inputs: We first investigate the response of the models to step inputs. The behaviour of the models is very clear: if the step height is above the threshold value, all models result in the asymptotic state corresponding to a binary switching transition. On the other hand if the step height is below the threshold value then the asymptotic state is the same as the basal state. Since the models have essentially the same threshold, all four models yield the same result. This is depicted in Fig. 9.2. It is worth pointing out that for monostable switches with moderately high Hill coefficient, this switch may not necessarily be very sharp. In this case (since the asymptotic amplitude of all switches is the same), the switching threshold for such a switch may slightly differ from that of sharp monostable/bistable switches.

Single pulse input: We now turn to the next of our inputs: a rectangular pulse. In this case the input regains its basal state eventually. For such an input the behaviour which we observe is as follows. A monostable-reversible model can never result in a binary transition, simply because the whole model is reversible, and there is a unique steady state (which is stable) for a given input. This is true, irrespective of the height and width of the pulse. In the case of a bistable-reversible switch, what we observe

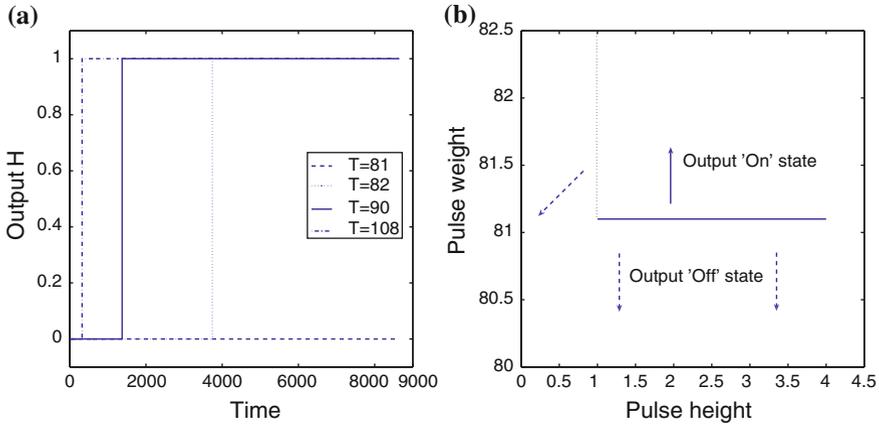


Fig. 9.3 Analysis of bistable reversible model in response to a single pulse. **a** Time course of output response to a pulse with various pulse durations for bistable reversible switch, revealing that a minimum pulse duration is needed to trigger the irreversible output response. **b** Critical curve for decision making in this model in terms of input characteristics: pulse amplitude and duration. The straight line suggests that the critical pulse duration is independent of pulse amplitude and divides the input parameter space into two sub-domains. The region above the line corresponds to the situation where a transition occurs. Parameters for the bistable model are $\alpha = 0.0044$, $a = 0.7$, $b = 0.3$, unless otherwise mentioned. Those of the decision module are $k_f = 0.1$, $k_r = 0.0044$. Unless mentioned, all parameters are dimensionless

is that if the pulse width (duration) is high enough, this can result in the bistable switch being permanently triggered, so that even when the input is switched off, the bistable module is permanently turned on, and this results in the threshold crossed in the downstream module resulting in a binary decision being made. This is depicted in Fig. 9.3a showing three pulses of different widths, clearly showing how an increase in pulse width can lead to the binary decision being triggered. This behaviour can be understood in terms of the interaction of a transient signal and a bistable switch, as studied in [12] (also see Bhalla and Iyengar [2]). It is worth emphasizing that of the two parameters, the parameter which determines whether the decision is triggered is the width of the pulse. This is further emphasized in Fig. 9.3b revealing a phase diagram in the input parameter space showing when the binary decision is triggered. The curve which separates the two outcomes is a straight line parallel to one of the axes, demonstrating that a threshold in pulse width needs to be crossed in order that that transition is triggered and that this width is the same, irrespective of pulse height as long as it is above a basic threshold.

We now examine the monostable irreversible switch (Fig. 9.4). In this case it is clear that the monostable switch module will revert to its basal state after the pulse is turned off. However the irreversibility in the decision module can still result in a transition being triggered. This happens if the pulse width is sufficiently large so that the monostable module output is high for a sufficiently long time, so that the threshold in the decision module is reached and a transition is irreversibly triggered.

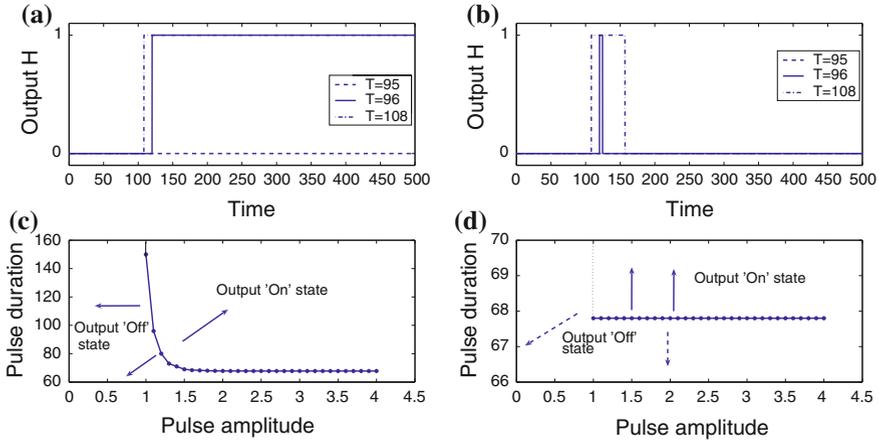


Fig. 9.4 Analysis of monostable reversible and irreversible models in response to a single pulse. Time course of output response to a rectangular pulse with various pulse durations for (a) monostable irreversible switch (b) monostable reversible models, under the same pulse amplitude $S = 1.1$. The two cases display dramatically different output responses. In particular, we see that in the monostable irreversible case, beyond a critical duration a transition is effected. **c** Critical curve for decision making in this model in terms of input characteristics: pulse amplitude and duration. The switching threshold is located at an input value of 1. A very similar trend to the bistable reversible case is observed. If a monostable switch involving moderate Hill coefficients is employed, the critical curve is practically flat as seen previously except when the pulse amplitude is close to the threshold. This is due to the non-sharpness of the switch. **d** The critical curve for a perfectly sharp monostable switch revealing a similar trend to the bistable reversible model in Fig. 9.3. Parameters for the monostable model are $A = 1, n = 10, k = 0.012$

Again here the pulse height plays practically no role, while the pulse width plays an important role.

It is worth pointing out that in both the above modules, the pulse height plays very little in the eventual outcome simply because both switches are sharp and have the same amplitude, which is independent of the input value, once it is above the switching threshold. The phase diagram for a perfect monostable switch is also a straight line parallel to the pulse height axis, as shown in Fig. 9.4. In our simulations we present results for a fairly sharp monostable switch of Hill coefficient 10 (see Fig. 9.4c). Here the phase diagram, for the most part is a straight line parallel to the pulse height axis. However for inputs in the narrow region where the monostable switch is not fully switched on (the transition zone), a sharp increase in pulse duration is needed. Thus the phase diagram has a sharply falling off curve which rapidly becomes parallel to one of the axes.

Finally, we examine the result of the bistable-irreversible module to a single pulse input. Here again, just as before, a pulse input can allow for the bistable switch to be triggered on, resulting in a binary decision being made in the module. The one difference between this module and the bistable reversible module is that once the threshold in the decision module is crossed, as a result of the bistable module being

essentially switched on, even if the bistable module is not close to a steady state, the binary transition is triggered. For the parameters chosen in the basal module, a single pulse which does not switch on the bistable module will not result in the binary decision being triggered. On the other hand the irreversibility embedded in the decision module implies that one can associate a precise time at which the decision is triggered. While the phase diagram in the input characteristics space (height, width and spacing) is the same as that of a bistable-reversible switch, different durations of pulses can substantially affect the time at which the transition is triggered. A longer duration pulse will result in the transition being triggered faster, as long as it is possible. The effect of varying the model parameters is studied in the next subsection.

Double pulse input. We now turn to a more complex input: double pulses, which are assumed to be identical and are characterized by the individual pulse characteristics as well as the time spacing between pulses. At the outset, we note that if a single pulse is unable to trigger a transition and the spacing between pulses is large, then the two pulses being effectively decoupled act as two isolated individual pulses and are unable to trigger a transition either. This is true for all models.

We first start with the monostable reversible model. Here again, since the model is reversible and the asymptotic state is the basal state, a double pulse input while inducing a transient is eventually unable to effect the transition. In the case of a bistable reversible switch, a double pulse input is indeed able to trigger a transition. There are some conditions for this: firstly the pulse height must be greater than a minimum value (the threshold and the width must be greater than a minimum value and for this case the spacing of the pulses must not be great). This is shown via simulations in Fig. 9.5, where the effect of two pulses of varying spacing are considered. We find that when the spacing is below a certain value the two pulse input can indeed induce a transition. Furthermore the two pulse input can induce a transition where the single pulse will not, and so this situation requires a certain memory of the bistable module from the first pulse which is used when the second pulse is input. This is therefore a situation where a two pulse input is able to effect the switching in the bistable module (without the single pulse doing so) and hence effect the transition. This result is more clearly seen in the phase diagram where the condition for a transition is shown with its dependence on the pulse height, width and spacing (see Fig. 9.5). The figure shows that pulse height plays a very minor role (the surface in the three parameter space is essentially a translate of a curve in the third direction. We also see that increasing the spacing can change the outcome making it possible for no transition to be effected.

When we examine the monostable irreversible switch (Fig. 9.6), we find that the double pulse can indeed effect a transition in this model too. Again the reason is similar, with the only difference being that the memory effect from the first pulse is used along with the second pulse to drive the decision module past the threshold. A more closely spaced pulse can induce a transition, and will do so more quickly. The three parameter diagram for this case is shown in Fig. 9.6, and we see a qualitatively similar plot to the bistable irreversible case. The only difference between these cases is that the bistable module is able to more effectively employ the memory effects from the first pulse to induce a transition. Therefore it generally will have a

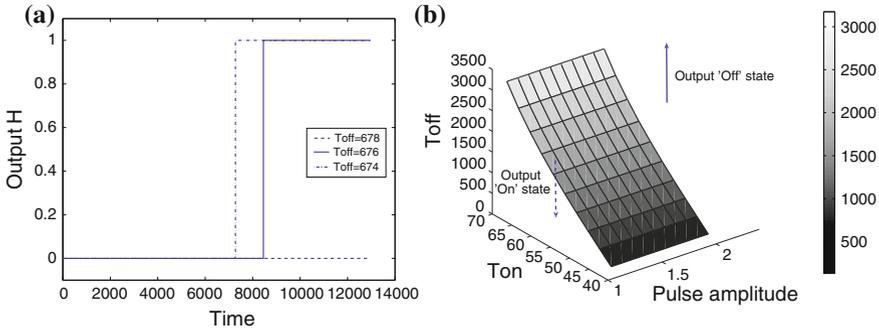


Fig. 9.5 Analysis of bistable reversible model in response to double pulses. **a** Time course of output response under double rectangular pulses (with the same pulse duration $T_{on} = 48$) with various spacing between pulses (T_{off}). This reveals a maximum T_{off} , above which a transition is not triggered. **b** Phase diagram depicting the effect of input characteristics (pulse amplitude, T_{on} , and T_{off}) in determining whether or not the transition occurs. An essentially identical behaviour is seen for the bistable irreversible model

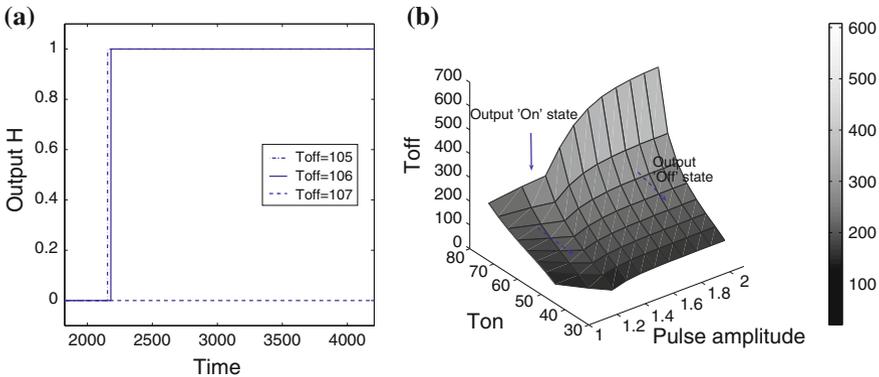


Fig. 9.6 Analysis of monostable irreversible model in response to double pulses. **a** Time course of output response under double pulses (with the same pulse characteristics: amplitude $S = 1.4$ and duration $T_{on} = 46$) with various spacing between pulses (T_{off}). It shows a maximum T_{off} , below which the transition is triggered. **b** Phase diagram depicting the effects of input characteristics (pulse amplitude, T_{on} , and T_{off}) on the transition. The behaviour is very similar to the bistable switch except for switch amplitudes just above the threshold

greater tolerable spacing between pulses for inducing transitions. This effect may be understood by noting that the bistability has a built in threshold at basal conditions, for the model variable, and once this is crossed, the system will asymptote to the elevated steady state.

We finally analyze the bistable irreversible case. For the parameters chosen, the bistable irreversible behaves in a very similar way to double pulse inputs as bistable reversible models. This is because, for the model considered, the only way for the threshold in the decision module to be crossed is by the bistable module variable

to become very close to the elevated steady state (and remain there for a fairly substantial period of time). This requirement guarantees that the bistable module will get switched on. Thus the behaviour of the module and the phase diagram is identical to the bistable reversible model.

Pulse train. We conclude this subsection by analyzing the signalling dynamics of the models when subject to a pulse train. Note that this is an example of a full fledged dynamic input, and the asymptotic state in general is not steady. In the case of the monostable reversible model, a pulse train elicits an oscillatory response. If the pulse train has a high frequency (relative to the time scale of the module) then the response has oscillations whose amplitude is severely diminished. When the pulse train has a low frequency, the module exhibits a periodic response of elevated phases separated by phases when the response is close to basal levels. In any case, in this model, there is no possibility of a transition.

When we examine the response of a bistable reversible module, we find that depending on the input it is possible for a period wave train input to induce the bistable switch module to oscillate very close to the elevated steady state. This is not surprising since we have already seen that single and double pulse inputs are capable of resulting in the bistable switch being triggered. A multiple pulse train allows more possibilities for the bistable switch to be switched on, since the presence of every additional pulse input only serves to increase the possibility of the bistable switch being switched on. What is also interesting is that the bistable module actually exhibits very small amplitude oscillations, which are not easily discernible. The reason for this is that the system hovers very close to the elevated steady state (which actually doesn't depend on the level of the signal) and asymptotically approaches it. The behaviour of this bistable module to different pulse inputs can be of two kinds: in the first case the pulse train is not actually able to effect a jump in the bistable switch. Thus a periodic pulse train (assumed to be of sufficiently high amplitude) results in a periodic oscillation, essentially with periodic jumps from the basal state. This is understood by noting that the bistable module remains close to the original basal steady state and every pulse in the pulse train results in a temporary elevation. In the second case, the pulse train can actually result in the system approaching asymptotically, the elevated state. Thus in this model, we may say that the transition has been triggered by the pulse train. A phase diagram depicting the effect of the input parameters in effecting a transition is shown in Fig. 9.7. As expected the pulse height plays a negligible role, while longer periods of pulse inputs can lead to a transition.

It is worth pointing out that for other bistable modules, the elevated steady state does have some (usually weak) dependence on the stimulus level and so one may in general expect small amplitude oscillations about the elevated state. In these cases while the asymptotic state is not steady, the response can be said to be one which has effected the transition, as long as the oscillation amplitude is not too high.

In the case of the monostable irreversible model (Fig. 9.7), we find that a pulse train can indeed induce a transition owing to the built-in irreversibility. Again this can be understood in terms of the response of this model to single and double pulse inputs and the fact that multiple pulses can only increase the possibility of a transition. The effects of the input parameters on the module response is shown in Fig. 9.7 and here

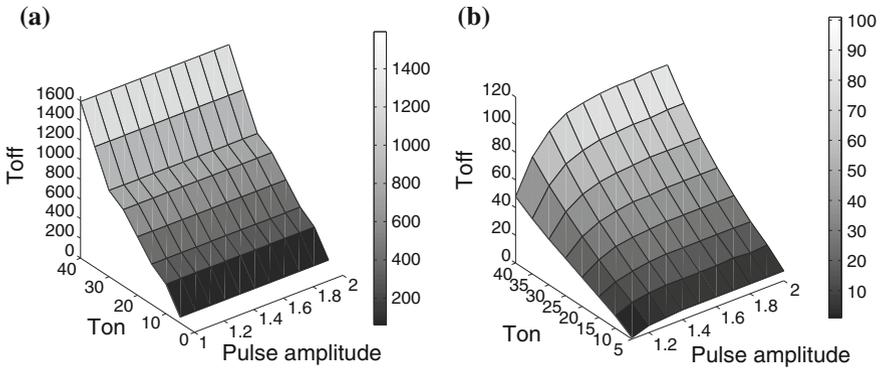


Fig. 9.7 Phase diagrams for bistable reversible and monostable irreversible switches in response to pulse trains. Phase diagram depicting the effects of input characteristic (pulse amplitude, T_{on} and T_{off}) in triggering transitions in the (a) bistable reversible (b) monostable irreversible models. The behaviour of the bistable reversible model is essentially the same as the bistable irreversible model in this case

again we find minimal dependence of pulse height, and a behaviour of the pulse duration which is qualitatively similar to that of the bistable reversible switch model. Finally in the case of the bistable irreversible model, we find that a pulse train can also induce a transition: this is only to be expected since this model has the bistability along with the additional irreversibility. An analysis of this model (results not shown) reveals that the phase diagram for the bistable irreversible model is essentially the same as the bistable reversible model. The behaviour of this model may be understood in a manner similar to the previous models.

Overall we see that step changes can induce transitions in all four models, while pulse inputs can induce transitions in models which have either bistability, irreversible in the decision module, or both. The qualitative dependence of the transition on input signal parameters is in fact very similar for these models.

In the above we have seen that the bistable reversible and the bistable irreversible models behaved in a very similar way which begs the question as to whether the downstream irreversibility plays any role. While the simulations above indicated no differences, simulations in the bistable model with a different parameter choice which changes the basal location (changing the parameter a in the model above) indicates that important differences can exist. Simulations were performed with a lower value of a , implying that at basal conditions, a higher barrier needed to be crossed to trigger the bistable switch. The behaviour of this bistable module coupled to both reversible and irreversible modules in response to a step input was exactly as before. When the two models were subject to a pulse input, we again see that the pulse needed a critical duration to trigger the switch. However a comparison between the bistable irreversible and bistable reversible switch reveals a clear difference: the minimum pulse duration needed in the bistable irreversible module is substantially less than the bistable reversible model. Thus in this case a series of pulse inputs

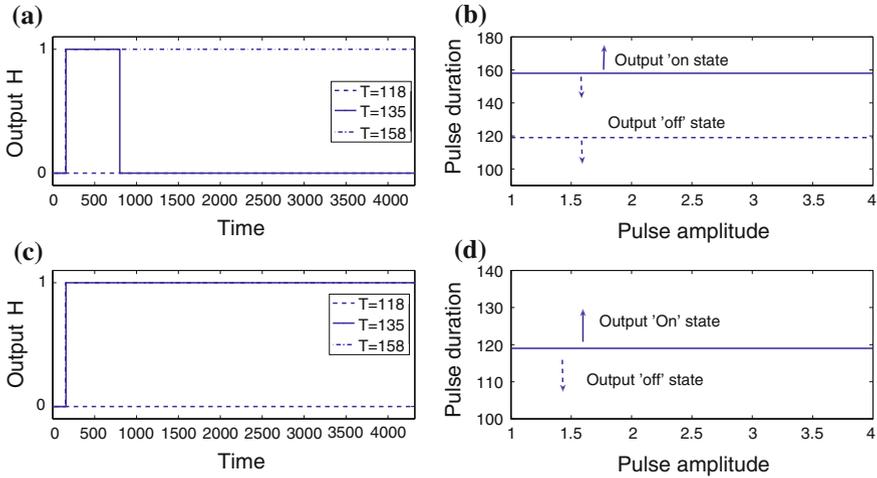


Fig. 9.8 Analysis of bistable reversible and irreversible models with different basal conditions in response to a pulse. **a** and **c** Time course of output response with different pulse durations. Again, a critical pulse duration is found to be necessary to trigger a transition. **b** and **d** Critical curve in terms of input characteristics (pulse duration and amplitude). The critical pulse duration in the irreversible implementation is shorter than that in the reversible implementation, which indicates that the downstream irreversibility partially relaxes the constraint of pulse duration that is required to trigger the irreversible output response. For this simulation the bistable model parameters are taken to be $a = 0.3, b = 0.7$

of intermediate duration were able to induce a transition in the bistable irreversible model but not in the bistable reversible model. For such inputs, the irreversibility of the decision module played the key role in effecting the transition even though the bistable switch was itself not triggered. For pulses of sufficient duration which trigger the bistable switch, the behaviour of both the bistable reversible and the bistable irreversible models is the same. This is depicted in Fig.9.8. Overall this is an example where the downstream irreversibility in the decision module plays a beneficial role in facilitating the transition and partially relaxing some constraints associated with the bistable module.

9.3.2 Effects of Model Parameters

We now examine the effects of model parameters to obtain some basic insights into how important parameters in the models can influence the possibility of a transition. Our focus will be on the roles of thresholds (including parameters in the basal state which effectively set thresholds) as well as the roles of the relative time-scales of the modules.

Effect of thresholds: At the outset, we note that the monostable and bistable modules have the same switch amplitude which is normalized to 1. We note that for this amplitude, if the downstream decision module threshold R_o is too high, then a transition can never be effected. Specifically if the steady state equilibrium value of the response for a switched on upstream module (of amplitude 1) is less than this threshold, i.e. $k_f/(k_f + k_r) < R_o$, then the transition will never be effected, irrespective of whether the downstream module is reversible or irreversible. If the downstream module is reversible, then the maximum steady state value it can attain corresponds to an upstream signal of 1 and this value is exactly $k_f/(k_f + k_r)$. Now even if the decision module is irreversible, by examining its dynamics we can rewrite the governing equation as

$$dR^*/dt = ((k_r + k_f)[k_f/(k_r + k_f) - R^*]) \tag{9.5}$$

From this, we immediately see that if R^* starts out less than $k_f/(k_f + k_r)$ (as it must since it is at steady state initially), then it can never attain a value greater than this transiently. Thus clearly the threshold must be low enough to be attainable, and we will assume that this is the case below.

Noting the above, and assuming an attainable threshold, we will focus on the level of the threshold, and the basal state in the bistable module (which is related). At the outset we note that for step changes, all modules are capable of exhibiting the transition and their behaviour is thus essentially the same here. The irreversible decision modules have associated with them a time of transition, and this decreases as the threshold is reduced.

We now focus on the response of these models to transient stimuli such as pulses. As mentioned before, a monostable-reversible module can never result in a transition. In the case of a bistable reversible module, decreasing the threshold does not actually have any important effect, since the only way to effect an irreversible transition in this module is to effect a switch in the bistable module, which is independent of downstream parameters. On the other hand changing the parameter a in the bistable module, keeping everything else fixed has an important quantitative effect. We note that at basal conditions, the bistable model variable has to attain a value of $1 - a$ to ensure that it will eventually be switched (in the absence of a signal the module has three steady states 0, $1 - a$, 1 and the intermediate unstable state sets the threshold which must be crossed for the switch to be triggered). Thus decreasing the value of a makes it more difficult for the module to be switched by transient signals. In particular the minimum time duration for a pulse to induce switching is $(1/\alpha)\ln(1/a)$ and so decreasing a implies the minimum duration of a pulse increases.

In the case of a monostable irreversible switch, decreasing the threshold R_o makes it easier to switch in a transient signal. This is also the case in a bistable-irreversible switch. If we define a critical signal strength S_o given by $k_f S_o/(k_f S_o + k_r) = R_o$, we see that this critical signal strength monotonically affects R_o . In particular we see that if the switch variable is maintained above this value, then the system is guaranteed to have a transition. Of course, signals which fall below this level can also lead to transitions. Now in the case of a bistable irreversible switch we see a combination

of two factors. If the switch variable can attain a height $1 - a$ it is guaranteed for the bistable switch to be switched on, while if the variable is maintained above S_o it is guaranteed to effect a transition. If $S_o < 1 - a$ it is indeed possible for a transient signal to not cause the bistable switch to switch on, while effecting the transition through the downstream module. This is simplest to see in the case where the decision module is much faster than the switch module. If a pulse input of sufficient duration causes the variable in the bistable module to become greater than S_o (but less than $1 - a$) then this can immediately drive the transition in the downstream module. This situation of having the downstream irreversibility be responsible for the transition expands the possibilities for the transition to be effected (in particular when the bistable switch is unable to do so) and is similar to the simulation result seen above. On the other hand if $S_o > 1 - a$ then any transient signal which causes the transition must of necessity drive the switch module above S_o and this makes it guaranteed for the bistable module to be switched on. Thus in this case, all signals which effect the transition also switch the bistable module on. For this case the module would switch on whether or not the downstream module was irreversible.

Effect of timescales: Having discussed the effect of thresholds, we now turn to the role of timescales. We consider two cases, one where the decision module is much faster than the switch module, and the other where the switch module is faster than the decision module. Note that the dynamics of the decision module can be altered by changing the parameters k_f and k_r keeping their ratio fixed. Firstly we note that if the decision module is reversible, the relative timescales do not play a role. In the reversible module, the only way to effect a transition is to have the upstream module switched on (essentially) permanently. In the case where the downstream module is irreversible the relative time scales play an important role.

Suppose the downstream module is much faster than the upstream module, then it is at a quasisteady state. Therefore transient variations in the switch module can drive the decision module and effect the transition. This situation is particularly suitable for a situation where the upstream module may not be permanently switched. All that is needed is to ensure that the variable in the upstream module (whether monostable or bistable) reaches a value S_o . In the case where the downstream module is slow, the downstream module is slowly regulated by the upstream module. For the transition to be triggered in the decision module, the upstream module must be at a high enough value (in particular higher than S_o) for a long enough time, the time requirements being dictated by the decision module. This is certainly possible in both monostable (for an input pulse of long duration) and bistable modules. For bistable modules however, the constraint of needing the variable value to be high enough for a long time results in the switch being effected in this module. Thus if the downstream module is slow, a monostable module can effect the switching if it is kept switched (by the driving signal) for a long time or by the bistable module being triggered (whether by a transient signal of short or long duration).

Taken together we can identify some essential design features in these models. A high threshold and slow decision module requires the switch module to be switched on for a long time if not permanently (and in the case of the bistable module switched on). A low threshold and a fast decision module is particularly amenable to relatively

weak transient signals driving the transition even bypassing the switching effects to some extent. In this situation, the irreversibility in the downstream module (when it exists) plays a dominant effect even if bistability is present upstream, and actually decision making in the system is quite insensitive to the nature of the upstream module. A high threshold and fast decision module needs the switch module to reach a sufficiently high level and even if the irreversibility in decision module triggers the decision, it invariably also involves the bistable module being switched on. The decision making is rate-limited by the upstream module. A low threshold and a slow decision module is a mixed situation. Here a monostable switch can remain switched on for a moderate time, while for a bistable switch depending on the relative magnitudes of S_o and the bistable threshold, it may be possible to switch without invoking the switching in the bistable module.

9.3.2.1 Model Analysis

To complement our discussion above, we present some analysis of our models. We start by noting that for a step input of sufficient magnitude, both monostable and bistable switches are triggered, and hence the response of all models at steady state is

$$R^* = k_f / (k_r + k_f) \tag{9.6}$$

As mentioned, the threshold R_o in the decision module must be below this value for a transition to be triggered. For all models, this value of R^* represents an upper bound (the initial condition for all models corresponds to steady state conditions with $S = 0$).

To proceed further with the analysis, we will make a simplification, purely for the sake of transparency in analysis. In the decision module, we will assume that $R^* \ll 1$. This will happen if k_f is much less than k_r (for a transition to occur, this naturally means that $R_o \ll 1$). This assumption simplifies the analysis of the dynamics in the decision module without making any important restrictions. In our analysis we will assume that the parameters k and α are not equal to k_r .

The dynamics of the decision module is now well approximated by

$$dR^* / dt = k_f X^* - k_r R^* \tag{9.7}$$

We consider the response of both monostable and bistable models to a pulse signal. The pulse corresponds to a step of amplitude $S_o > 1$ for a time duration T . The response of the monostable module is

$$\begin{aligned} X^* &= 1 - \exp(-kt), t < T \\ X^* &= (1 - \exp(-kT))\exp(-k(t - T)), t > T \end{aligned} \tag{9.8}$$

The switch module response attains an amplitude of $1 - \exp(-kT)$ before decreasing.

The response of the bistable module is

$$\begin{aligned}
 X^* &= 1 - \exp(-\alpha t), t < T \\
 dX^*/dt &= -\alpha X^*(1 - X^*)(X^* - (1 - a)), t > T
 \end{aligned}
 \tag{9.9}$$

The response for $t > T$ is obtained by solving the nonlinear differential equation, and while an analytical expression can be obtained, it becomes difficult to explicitly obtain X^* as a function of time for $t > T$. The key insight which is relevant here is that if $X^*(T) = 1 - \exp(-\alpha T)$ is greater than the basal bistable barrier of $1 - a$, then switching is guaranteed to a pulse input, and the system will subsequently evolve to the elevated steady state. This requires that $\exp(-\alpha T) < a$ or in other words $T > (1/\alpha)\ln(1/a)$.

Now the response of the decision module is given by

$$R^* = \exp(-k_r t) \int_0^t X^*(s) \exp(k_r s) k_f ds
 \tag{9.10}$$

We can use this to make some key inferences. Firstly, suppose the upstream module does get permanently switched on, then we have $X(t) \rightarrow 1$ as $t \rightarrow \infty$, in which case $R^* \rightarrow k_f/k_r$. This guarantees that the transition is triggered. On the other hand if the upstream module gets switched off, then $X^* \rightarrow 0$ as $t \rightarrow \infty$. In this case, a transition may be triggered if the threshold in the decision module is reached transiently.

The response of the decision module with an upstream monostable switch is

$$R^* = (k_f/k_r)(1 - \exp(-k_r t)) - (k_f/(k_r - k))(\exp(-kt) - \exp(-k_r t)), t < T
 \tag{9.11}$$

Note that the second term is negative. Clearly if T is very large then R^* will approach k_f/k_r very closely. If the decision module is irreversible, a transition will be triggered at a time instant t_o where $R^*(t_o) = R_o$. A positive t_o value less than T signifies the possibility of an irreversible transition. This is not of course guaranteed.

Note that it is possible for the transition to be triggered in some cases, after the initial duration of the pulse. For the subsequent dynamics after the input stimulus is removed, we have

$$\begin{aligned}
 R^*(t) &= R^*(T) \exp(k_r(T - t)) + (k_f/(k_r - k))(\exp(kT) - 1) \\
 &\quad [\exp(-kt) - \exp((-kT - k_r(t - T))] , t > T
 \end{aligned}
 \tag{9.12}$$

Clearly it is possible for a transition to be triggered for $t > T$.

One can obtain further insights from the above expressions in certain limiting conditions. In the first case, suppose $k \ll k_r$, then we notice from the expression for R^* for $t < T$ that the response is well approximated by

$$R^* = (k_f/k_r)(1 - \exp(-kt)), t < T
 \tag{9.13}$$

from which we immediately see that the time for the transition to be triggered is

$$t_0 = (1/k) \ln(k_f/k_r - R_o) \quad (9.14)$$

assuming it is less than T . It is also clear from the expression for R^* for $t > T$ in this case that R^* decreases after $t > T$. This means that an irreversible transition must be triggered before time T if it is to be triggered at all.

The opposite case $k \gg k_r$ can also be considered. Here from the expression for the response we obtain

$$R^*(T) = (k_f/k_r)(1 - \exp(-k_r T)) - (k_f/(-k))(-\exp(-k_r T)), t < T \quad (9.15)$$

From this we see that a transition can be triggered for a pulse of long enough duration T . It is also worth pointing out that if k_f and k_r are made small, keeping their ratio fixed, and $k_r \ll k$, this indicates that a pulse of duration which is of the order of the inverse of k_r ($T = O(1/k_r)$) is needed. The essential insight here is that a pulse of long enough duration (which is determined by the decision module) is needed to trigger the transition (Fig. 9.10).

Similar insights can be obtained for the bistable module. For $t < T$, the dynamics of X^* here is very similar to the monostable module (replace k by α) and exactly parallel conclusions can be drawn. However depending on $X^*(T)$, it is possible for the bistable module to be switched on (if $X^*(T) > 1 - a$) guaranteeing a transition, or for the transition to be triggered before $t = T$ (through the downstream irreversibility), or for the downstream irreversibility to trigger the transition after $t = T$, with the bistable switch not being switched. In general if $X^*(t) < 1 - a$ but high enough for R^* to exceed R_o a transition can be triggered. This can be analyzed similar to the analysis above, except that in the bistable case, explicitly obtaining X^* as a function of time is not possible.

Overall our analysis reveals the balance of factors in effecting a transition and the role of the thresholds and relative time scales.

9.3.3 Behaviour of Extended Models

We provide an example to illustrate the implications of our results. We consider a model where a signal enters the cell, and may also be pumped out. The signal can initiate an irreversible decision through the signalling cascades via a switch combined with a decision module. An example of this is when drugs are pumped into cells whereupon they may trigger an apoptotic response [8]. A model for this involves augmenting signalling models of the kind studied earlier with additional features. The additional features of the model are described by the following equations.

$$\begin{aligned} dc_e/dt &= P(Q.S - c_e) + V_2 c_i / (k_2 + c_i) - V_1 c_e / (k_1 + c_e) \\ dc_i/dt &= -v_2 c_i / (k_2 + c_i) + v_1 c_e / (k_1 + c_e) \end{aligned}$$

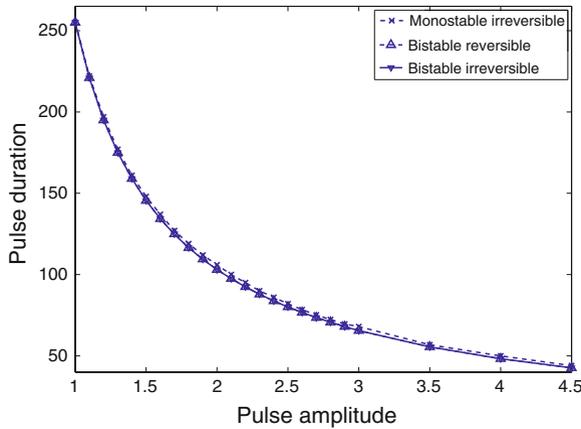


Fig. 9.9 Analysis of signal transduction and decision making in different configurations in response to pulse signal in a specific setting. Critical curves in terms of input characteristics (pulse amplitude and duration) for bistable reversible, bistable irreversible and monostable irreversible switches in a specific model setting, where the pulse signal is input to a signalling cascade involving a switch and a decision module. In addition, other features involving pumping the chemical signal in and out of the cell are also present. The same output response is obtained for both bistable reversible and irreversible pathways in response to a rectangular pulse, and together with monostable irreversible switch, all combinations give rise to strikingly similar phase diagrams. The model parameters are $P = 0.33$, $Q = 0.25$, $k_1 = 0.219$, $k_2 = 1.37$, $V_1 = V_2 = 4.67$, $v_1 = v_2 = 0.467$

This module describes the generation of drug c_e in the extracellular medium by an external signal S , its pumping into the cell via Michaelis Menten kinetics (maximum rate V_1) and the pumping out of intracellular drug c_i also via Michaelis Menten kinetics (maximum rate V_2). The intracellular drug is the input for a signalling cascade which involves a switch module combined with a decision module.

Simulations were performed with different kinds of inputs, for both kinds of switch modules. Here again, there is a critical value of S which can result in cell killing. In the case of a pulse, a pulse must have a sufficient duration to induce killing. A phase diagram of input parameters for a pulse signal to result in cell killing reveals that bistable-reversible, monostable-irreversible and bistable-irreversible modules have a very similar phase diagrams. This is illustrated in Fig. 9.9. This indicates that even in more complex settings it is not easy to obtain clear distinctions between these modules, and that the insights obtained above continue to hold good.

	(a) 	(b) 	(c) 	(d)
	Irreversible transition			
Step input	√	√	√	√
Transient input: Single pulse	×	√	√	√
		(1) Pulse intensity \geq threshold (2) Pulse duration \geq minimum pulse duration		
Double pulse/ Pulse train	×	√	√	√
		(1) Pulse intensity \geq threshold (2) Pulse duration \geq minimum pulse duration (3) Pulse interval \leq maximum pulse interval		
Key factors/ parameters	The downstream irreversibility is a dominant/necessary factor in providing the capability for an irreversible transition.		(1) Both models can give rise to the irreversible transition; either downstream irreversibility or upstream bistability can be a dominant and triggering factor, depending on the signal and parameters (threshold and relative time scales). (2) Under certain cases, downstream irreversibility provides a degree of redundancy in triggering an irreversible transition, and partially relaxes the constraint of pulse duration.	

Fig. 9.10 A tabular summary of essential results. The table summarizes the essential results obtained from the study

9.4 Conclusions

Switch-like behaviour and thresholds are ubiquitous in cell signalling. In addition, there are a number of contexts in which irreversible signal processing and decision making occurs: this includes transitions in the cell cycle, apoptosis, decisions to differentiate etc. Further, other signalling which triggers decisions such as degradation of particular proteins may be essentially irreversible (eg see Varedi et al. [15]). In this article we employed a systems approach to examine the dynamics of two

typical switches, monostable and bistable switches, and how they could be involved in irreversible information transformation in signalling pathways. The focus was to investigate how irreversible decision making may be realized, especially through different kinds of switch modules, and identify the capabilities and constraints involved. This approach provides a basis for understanding irreversible decision making in different cellular contexts.

Our models conceptualized the irreversible decision making through a sequential combination of a switching “module” and a decision “module”. Keeping in mind our simplified approach and focussing on only essential elements, we do not consider other elements such as retroactivity, or include details specific to particular cellular signalling contexts; likewise we employ generic descriptions of monostable and bistable switches which have typical input-output behaviour, rather than focus on specific examples. Our study involved studying combinations of monostable and bistable switches, in combination with simple decision modules which could either be reversible or irreversible. While different reactions in signalling pathways are reversible in principle, sometimes key transitions occur as a result of some threshold being crossed, triggering a particular new event which is the source of the irreversibility. We emphasize that every one of the model configurations we have studied has been invoked explicitly or implicitly in one context or another.

Our analysis of the modules reveals that while a monostable reversible module by itself is incapable of irreversible transitions, bistable-reversible, monostable-irreversible and bistable-irreversible models were all capable of irreversible decision making. All these models were capable of exhibiting transitions in response to transient stimuli such as pulse signals. The bistable models have irreversibility built in to them, while monostable models coupled to irreversible decision modules could result in irreversible decision making. Our analysis examined the signal transduction of all these models to common experimentally employed inputs such as single pulses, multiple pulses and pulse trains, and examined the key parameters in the input which would result in an irreversible decision. We found that all these modules were capable of irreversible transitions in response to these signals and that the qualitative effects of the input parameters on all these modules were similar. Furthermore, we found that the quantitative effects of the input parameters on whether the irreversible transition was realized could be made very similar, by appropriate choice of parameters. The implication for detailed models in concrete contexts is that it may be possible to relatively easily fit the steady state response to simple classes of inputs by suitable choices of a few parameters. Examining these models in the context of a model of drug induced cell killing, incorporating drug getting pumped in and out of the cell and eliciting an apoptotic response via signalling pathways involving switches reveals that the response of the models to the various inputs with respect to whether cells were killed was in fact strikingly similar, even though there were qualitative differences in all these models. This is also in agreement with the behaviour of such pathways in a distributed model of drug transport, uptake and drug induced killing [8].

Our examination of key parameters in the models allowed us to elucidate the key capabilities and constraints associated with each combination of modules with respect to irreversible decision making. In general, having a lower threshold, made it easier

for a transition to be triggered and in particular made it easier and more feasible for an irreversibility in the decision module to trigger the transition. Lower threshold not only made it possible for monostable irreversible models to more easily effect a transition, but also had the same effect in bistable irreversible models. It allowed for transitions to be effected in bistable irreversible models without the bistable switch being triggered, thus making the irreversibility in the decision module, the dominant factor.

With respect to time scales, in general a slower switch module and fast decision module, facilitated transitions, and in particular made transitions associated with irreversible decision modules much easier. Thus even here, a bistable irreversible module could effect the transition without a bistable switch being triggered. Therefore, the bistability notwithstanding, the module exhibited an irreversible transition owing to the downstream irreversibility, and in a manner very similar to a monostable irreversible switch. On the other hand slow decision modules implied that in the bistable case, a transition was effected in response to pulse stimuli, with the bistable module very likely switched on. This was because, for the irreversible transition a fair level of stimulus to the decision model was needed for a sustained and long period of time, which would inevitably (due to the intrinsic dynamics of the bistable module) cause the bistable switch to be triggered. Overall irreversibility in the decision module as the dominant factor was favoured by fast decision modules and not too high threshold. High threshold and slow decision modules implied that either the transition occurred due to a long transient input (monostable irreversible model) or inevitably resulted in the bistable switch being triggered and being the dominant factor. The presence of irreversibility in the downstream module when combined with a bistable module could thus be a redundant effect in this case, whereas in the previous case, the irreversibility downstream clearly enhanced the capability of the overall pathway for an irreversible transition, relaxing constraints associated with the upstream module.

9.5 Discussion

Since irreversible transitions occur in multiple contexts in cell signalling pathways it is important to carefully identify the source of the irreversibility. Very often in cell signalling modules, such transitions are modeled by bistable switches, even though the bistability is only shown to be plausible and not clearly demonstrated. In some cases it is acknowledged that there may be an additional irreversibility downstream, which was not considered important. In other situations such as drug-induced apoptosis, there is a debate as to what the mechanism is [1, 3, 6, 8, 17]. In individual contexts there are so-called validated models which match models to data but which do not actually deal with the key issue of the nature of the transition. In other cases bistable models are in effect modelled as a shorthand for an irreversible transition, though it is not clear if bistability actually exists. This situation is quite striking since monostable and bistable switches are signalling features with very different dynamical characteristics, even qualitatively.

Our results reveal that in order to discriminate between monostable and bistable models, careful experiments must be performed with a sufficiently broad and rich class of transient signals (perhaps designed) on one hand, and carefully monitoring and tracking the possibility of discrete (essentially) irreversible cellular events. It indicates that such events could by themselves (especially coupled with monostable switches and thresholds) be at work in different contexts and that bistability may either be not present or not the dominant factor. Further while from a modelling point of view bistable switches may be appealing as a robust structure which causes an irreversible transition, the robustness in actual cellular signalling may be realized by a host of other factors which for example act to prevent an irreversible transition from happening, which may in the end be realized by both monostable and multistable signalling circuits. Our results and insights also have implications for engineering signalling pathways to result in irreversible decision making and the capabilities and constraints involved.

The skeletal modelling framework we employed allows us to investigate and understand such signal transduction in different contexts and appreciate the role and relevance of additional layers of regulation involved. From a systems point of view our studies highlight the need to go well beyond steady state analysis not only experimentally but also theoretically to understand essential aspects of cell signalling and the complexity of the organization of information flow in living systems.

9.6 Lessons Learnt

1. Irreversible transitions can arise through both monostable and bistable signalling pathways. In the case of monostable pathways this has to be due to some irreversible downstream event being triggered
2. Insights into design capabilities and constraints and key factors involved in irreversible decision making were obtained by the analysis of a series of simplified models which focus on essential aspects.
3. The analysis and results can be used both in modelling and understanding complex signal transduction also and in designing synthetic circuits with irreversible decision making.

References

1. Albeck JG, Burke JM, Spencer SL, Lauffenburger DA, Sorger PK (2008) Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLoS Biol* 6:2831–2852
2. Bhalla U, Iyengar R (1999) Emergent properties of networks of biological signaling pathways. *Science* 283:381–387
3. Eissing T, Conzelmann H, Gilles ED, Allgower F, Bullinger E, Scheurich P (2004) Bistability analyses of a caspase activation model for receptor-induced apoptosis. *J Biol Chem* 279:36892–36897

4. Ferrell JE, Xiong W (2001) Bistability in cell signaling: how to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* 11:227–236
5. Goldbeter A, Koshland DE (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci USA* 78:6840–6844
6. Gu C, Zhang J, Chen Y, Lei J (2011) A trigger model of apoptosis induced by tumor necrosis factor signaling. *BMC Syst Biol* 5(Suppl 1):S13
7. Huang C, Ferrell J (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci USA* 93:10078–10083
8. Liu C, Krishnan J, Xu XY (2011) A systems-based mathematical modelling framework for investigating the effect of drugs on solid tumours. *Theor Biol Med Model* 8:45
9. Markevich NI, Hoek JB, Kholodenko BN (2004) Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol* 164:353–359
10. Novak B, Tyson JJ, Gyorfy B, Csikasz-Nagy A (2007) Irreversible cell-cycle transitions are due to systems-level feedback. *Nat Cell Biol* 9:724–728
11. Schittler D, Hasenauer J, Allgower F, Waldherr S (2010) Cell differentiation modeled via a coupled two-switch regulatory network. *Chaos* 20
12. Seaton D, Krishnan J (2011a) A modular systems approach to elucidating the interaction of adaptive and monostable and bistable threshold modules. *IET Syst Biol* 5:81–94
13. Seaton D, Krishnan J (2011b) The coupling of pathways and processes through shared components. *BMC Syst Biol* 5:103
14. Thomson M, Gunawardena J (2009) Unlimited multistability in multisite phosphorylation systems. *Nature* 460:274–277
15. Varedi KSM, Ventura AC, Merajver SD, Lin XN (2010) Multisite phosphorylation provides an effective and flexible mechanism for switch-like protein degradation. *PLoS One* 5:e14029
16. Yao G, Lee TJ, Mori S, Nevins JR, You L (2008) A bistable Rb-E2F switch underlies the restriction point. *Nat Cell Biol* 10:476–482
17. Zhang T, Brazhnik P, Tyson J (2009) Computational analysis of dynamical responses to the intrinsic pathway of programmed cell death. *Biophys J* 97:415–434

Chapter 10

From Specification to Parameters: A Linearization Approach

Heinz Koepl, Marc Hafner and James Lu

Abstract With the improvement of protocols for the assembly of transcriptional parts, synthetic biological devices can now be reliably assembled based on a design. The standardization of the parts open up the way for *in silico* design tools that improve the construct and optimize devices with respect to given formal specifications. The simplest such optimization is the selection of kinetic parameters and protein abundances such that the specified constraints are robustly satisfied. In this chapter we address the problem of determining parameter values that fulfill specifications expressed in terms of a functional on the trajectories of a dynamical model. We solve this inverse problem by linearizing the forward operator that maps parameter sets to specifications, and then inverting it locally. This approach has two advantages over brute-force random sampling. First, the linearization approach allows us to map back intervals instead of points and second, every obtained value in the parameter region is satisfying the specifications by construction.

Keywords Linearization · Parameters · Temporal logic · Linear temporal logic (LTL) · Specification functionals · Stoichiometric matrix · State space · Equilibrium point · Bifurcation point · Inverse bifurcation analysis · Inverse bifurcation synthesis

H. Koepl (✉)

ETH Zurich and IBM Research Laboratory, Zurich, Switzerland
e-mail: koepl@ethz.ch

M. Hafner

Harvard Medical School, Boston, MA, USA
e-mail: marc_hafner@hms.harvard.edu

J. Lu

F. Hoffmann-La Roche, Basel, Switzerland
e-mail: james.lu.jll@roche.com

10.1 Introduction

Synthetic biology places emphasis on small, standardized molecular parts and devices, mostly operating at the transcriptional level [1, 2]. With standardization comes the need for rigorous quantitative characterization of such devices and compositional theory to reliably build larger systems from small canonical circuits. For now most synthetic circuits implemented *in vivo* were constructed from a small number of components with topology and parameter values found by trial-and-error. The development of larger synthetic systems necessitates the use of appropriate design methodologies. *In silico* analyses can provide significant insights into the construction of complex synthetic systems, but due to the poor quantification of experimental and micro-environmental conditions, the predictive capability of *in silico* models for *in vitro* implementations remains limited. Apart from experimental limitations, modeling attempts to date most often make simplifying assumptions about all the perturbations that a synthetic construct is facing *in vivo*. For instance, only a few studies account for the large extrinsic noise [3–5] and in particular the one introduced by variations of plasmid copy number [6].

For a given circuit topology, kinetic parameters and other parameters that are involved in controlling the expression level of molecular species (e.g. promoter activity or number of ribosome binding sites) are important design parameters in synthetic biology. A major challenge is to find a set of parameters that satisfies the behavioral specification of a device [7]. Computer science offers various languages to formally define the proper functioning of a piece of code or hardware. Such specification languages of formal verification are used to check important behavioral properties, such as liveness, safety or fairness [8]. One convenient way to specify such properties is to use *temporal logic*, which is considered an extension of classical propositional reasoning, where propositional variables may change their truth values over time. A prominent such logic is the linear temporal logic (LTL), where the truth value of the propositions is interpreted over a linear timeline [8]. Such techniques were already applied to investigate robustness of computational models in system biology [9].

Mathematically, the design problem is an inverse problem and hence inherits the general feature of such problems, namely ill-posedness [10, 11]. More specifically, for a certain behavioral specification one aims to find the corresponding parameter set that gives rise to such behavior. An simple example for a quantity in feature space could be the concentration of a molecular species at particular time-points. The problem is closely related to parameter optimization and even more so to robust optimization, where an objective function—generally encoding some behavioral constraint (e.g. making model trajectories close to the measurements)—is optimized to yield the optimal parameter set. Ill-posedness refers to the observation that two close-by points in specification or behavioral feature space may map to very distant points in the parameter space, indicating that this mapping is generally not contractive but rather expansive. The inverse and corresponding forward problem is illustrated in Fig. 10.1.

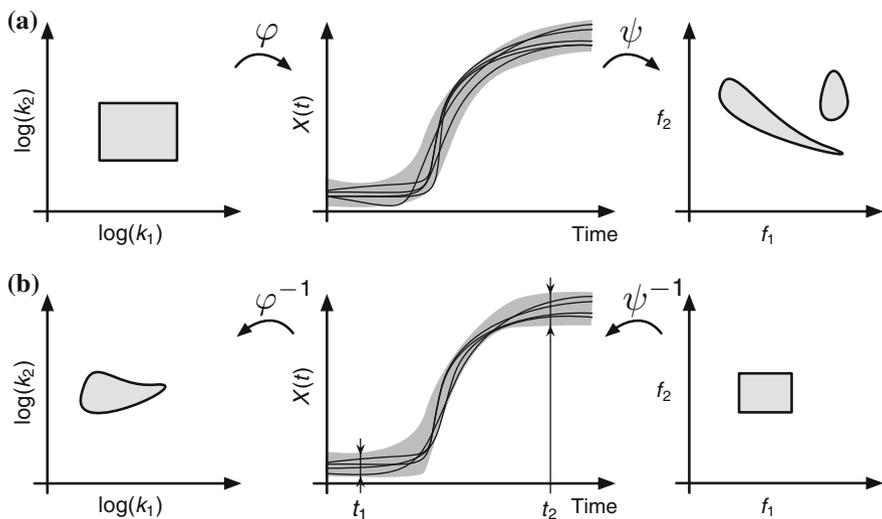


Fig. 10.1 **a** The forward problem of defining a parameter set from which trajectories and their behavioral features are computed. **b** The inverse problem of finding a parameter regions for a predetermined behavioral specification region S . Columns from left to right correspond to parameter space, trajectory space and behavioral feature space, respectively. Connected convex sets can map to nonconvex non-connected regions

In the chapter we restrict ourselves to models obeying the reaction rate equation and hence constitute a set of nonlinear ordinary differential equations. In general, connected domains may map to disconnected domains, for instance if the dynamical system contains bifurcation points (e.g. see Fig. 10.1). For the proposed linearization approach we will further restrict ourselves to connected domains in the respective image space. Moreover, we will not resort to specifying behavior through temporal logics but will define general *specification functionals*. These are mappings ψ from an appropriate function space \mathcal{X} of n -dimensional trajectories (e.g. $L_2([0, T], \mathbb{R}^n)$) to the m -dimensional reals and we choose the form

$$\psi(x) \equiv \int_0^T g(s, x(s)) ds$$

with $x \in \mathcal{X}$ and the feature kernel $g : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathcal{F}$, where $\mathcal{F} \subseteq \mathbb{R}^m$. A special and more tractable version of the kernel is the convolution, i.e. $g(t, x(t)) = h(T - t)x(t)$. In the following we will only require the map $x \mapsto g(\cdot, x)$ to be once-differentiable. With this, we can define the forward map from a p -dimensional parameter space to the feature space as the composition $F \equiv \psi \circ \phi$, with $\phi : \mathbb{R}^p \rightarrow \mathcal{X}$. The trajectories $x \in \mathcal{X}$ are generated by the reaction rate equation

$$\frac{d}{dt}x(t) = Nv(x(t), k) \quad \text{and} \quad x(0) = x_0, \tag{10.1}$$

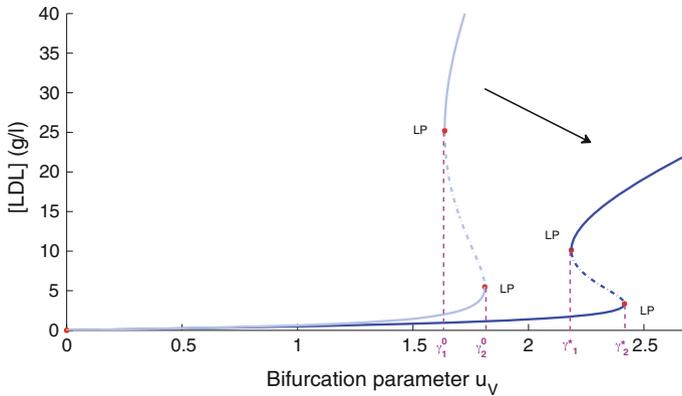


Fig. 10.2 Inverse bifurcation analysis or bifurcation design for a bistable system occurring in lipoprotein metabolism [14, 15]; original bifurcation structure characterized by bifurcation points γ_1^0 and γ_2^0 and desired new bifurcation structure determined by γ_1^* and γ_2^* . The design goal was to shift the abscissas of the bifurcation points to $(\gamma_1^*, \gamma_2^*) = 4/3(\gamma_1^0, \gamma_2^0)$

with the stoichiometric matrix $N \in \mathbb{Z}^{n \times q}$, the reaction flux vector $v : \mathbb{R}_{\geq 0}^n \times \mathbb{R}_{\geq 0}^p \rightarrow \mathbb{R}_{\geq 0}^q$ and $k \in \mathbb{R}^p$ the parameter set. In principle, specification functionals can also be the positions of equilibrium points, for instance with

$$\psi(x) = \lim_{t \rightarrow \infty} x(t),$$

where the support of x is enlarged appropriately. Hence, if initial conditions x_0 are treated as parameters as well the domain in specification space would show all equilibrium points and will in general also contain bifurcation points. The corresponding inverse problem or design problem of placing equilibrium points and bifurcation points at particular position of the state space is closely related to inverse bifurcation analysis [12, 13]. That is, given a desired coordinate for a bifurcation point one aims to determine its associated parameter set. An example illustrating the problem for a bistable system arising in a kinetic model for lipoprotein metabolism [14, 15] is shown in Fig. 10.2. Throughout the remaining part of this chapter we will not consider such asymptotic behavioral features and hence restrict the support for x to some finite interval $[0, T]$. Subsequently, we outline in Sect. 10.2 the proposed backward sampling approach and Sect. 10.3 provides a simulation study.

10.2 Backward Mapping via Linearization

The brute-force method of determining the parameter region that satisfies a certain behavioral specification $S \subseteq \mathcal{F}$ usually proceeds by Monte Carlo sampling of parameter sets, generating corresponding trajectories according to (10.1), checking

whether those satisfy S and finally retaining only those parameter sets that led to satisfied specification S . There are two immediate downsides of this approach. First, most draws will be unsuccessful for high dimensional parameter spaces, for tight specifications, or for both. Different approaches using an optimized sampling [16, 17] have been developed to mitigate this problem, but are not solving it as they require convergence of the sampling. Second, drawing parameter points in \mathbb{R}^p does not provide guarantees that those points belong to a connected domain of consistent parameter sets. Here we provide first attempts to tackle both problems.

The main idea is to locally linearize the forward map F around some point and then locally invert it. Hence, a small enough local patch in feature space can be mapped backward to a small patch in parameter space. By successively sampling expansion points in their neighborhoods (e.g. by the ball-walk algorithm [18]) we can systematically cover the entire specification S and obtain the corresponding parameter region. A series expansion of F around some initial parameter set k^0 reads

$$F(k^0 + dk) = F(k^0) + \left. \frac{\partial F(k)}{\partial k} \right|_{k=k^0} dk + o(dk)$$

Defining $df \equiv F(k^0 + dk) - F(k^0)$ we see that a neighborhood df in feature space to first order can be mapped backward using the Moore-Penrose pseudo-inverse

$$dk = L^\dagger df,$$

that we define with care as

$$L^\dagger \equiv \lim_{\lambda \rightarrow 0} (L^T L + \lambda I)^{-1} L^T = \lim_{\lambda \rightarrow 0} L (L L^T + \lambda I)^{-1}, \quad (10.2)$$

where L denotes the linearized forward map and hence is just the $m \times p$ matrix

$$L \equiv \left. \frac{\partial F(k)}{\partial k} \right|_{k=k^0}. \quad (10.3)$$

Note, that the limit in (10.2) exists even if the inverse of $L^T L$ and $L L^T$ do not exist. Such situations are encountered as soon as the number of specification features m are less than the number of parameters, i.e. the dimension p of the parameter space. Importantly, we can compute (10.3) efficiently using the variational equation for the system (10.1). Observe that

$$L = \frac{\partial}{\partial k} \psi \circ \varphi(k) = \int_0^T \left. \frac{\partial g(s, x)}{\partial x} \right|_{x=x(s, k)} \left. \frac{\partial x(s, k)}{\partial k} \right|_{k=k^0} ds,$$

where the last terms in the integral is just the sensitivity of the solution of (10.1) to perturbations in k around k^0 . According to the variational equation the sensitivity obeys the following ordinary $n \times p$ matrix differential equation

$$\frac{d}{dt} \frac{\partial x(t, k)}{\partial k} = N \left(\frac{\partial v(x, k)}{\partial x} \frac{\partial x(t, k)}{\partial k} + \frac{\partial v(x, k)}{\partial k} \right) \quad \text{with} \quad \frac{\partial x(0, k)}{\partial k} = 0, \quad (10.4)$$

where we skipped the explicit dependency on k^0 for brevity. Note, that (10.4) is equivalent to the transient sensitivity analysis of metabolic networks [19, 20], proposed as an extension of classical metabolic control analysis that only deals with steady state sensitivities. For a certain k^0 the sensitivity of the kernel g is a constant $m \times n$ matrix that can be computed explicitly. Thus, by jointly solving (10.1) and (10.4) for some k^0 together with

$$\frac{d}{dt} L(t) = \left. \frac{\partial g(s, x)}{\partial x} \right|_{x=x(t, k^0)} \left. \frac{\partial x(t, k)}{\partial k} \right|_{k=k^0} \quad \text{with} \quad L(0) = 0$$

up to time T we obtain the linearized map $L = L(T)$. Hence, for every sampled k^0 and associated feature point f^0 we propose to design a feature ball

$$\mathcal{B}_{f^0}(\delta) = \left\{ f \in \mathcal{F} \mid \|f - f^0\|^2 \leq \delta \right\}$$

and map it backward using L^\dagger . According to the singular value decomposition $L^\dagger = U \Sigma V$ with Σ a diagonal matrix with non-negative entries [11], the backward transformation needs to be a sequence of a rotation, a scaling and another rotation and hence the image of \mathcal{B}_{f^0} under L^\dagger can only be a ellipsoid in the parameter space

$$\left\{ L^\dagger f \mid f \in \mathcal{B}_{f^0}(\delta) \right\} \in \mathbb{R}^p.$$

Clearly, sampling a multivariate region with balls of same dimension allow for a complete coverage of the region—something that is can only be extrapolated when using point samples [16]. The question to efficiently sample a region with balls has been addressed in computational geometry and efficient randomized algorithms are available [18].

We remark that the map L is not the best local approximation to $F(k)$ in some norm sense. More specifically we can improve on L if we are giving additional samples of the neighborhood $\mathcal{B}_{f^0}(\delta)$. Consider we draw another $k^i \in \mathcal{B}_{k^0}$, then we can construct a rank-one update to L

$$\tilde{L}^i = L + \frac{\Delta F - L \Delta k}{\|\Delta k\|^2} \Delta k^T \quad (10.5)$$

where $\Delta F \equiv F(k^i) - F(k^0)$ and $\Delta k \equiv k^i - k^0$. In particular, the rank-one term (10.5) captures the nonlinear part of F . From (10.5) it follows that the matrix \tilde{L}^i satisfies the consistency property

$$\tilde{L}^i (k^i - k^0) = F(k^i) - F(k^0). \quad (10.6)$$

Thus, knowing how to construct rank-one updates over the domain of interest is equivalent to knowing $F(k)$. In fact, \tilde{L}^i is the matrix closest to L , with respect to the Frobenius norm, that satisfies (10.6). Subsequently we will use this improved linear approximation to F to bound the error that one can incur if one uses the pseudoinverse L^\dagger for the backward map. This will also provide means to determine the maximal ball size δ to stay below a certain error. We quantify the error in the feature space by the backward map followed by a forward map. That is, we want to find a δ such that

$$\|F(k^0 + L^\dagger(f - f^0)) - f\|_2 \leq \varepsilon \quad (10.7)$$

for all $f \in \mathcal{B}_{f^0}(\delta)$.

Now suppose we know a bound $\rho(\delta)$ for the Frobenius norm of the rank-one perturbation, i.e. $\|\tilde{L} - L\|_F \leq \rho(\delta)$ in the local domain of interest. Note, that $\rho(\delta)$ could and need to be estimated by sampling. Given a $f^i \in \mathcal{B}_{f^0}(\delta)$ the maximal error of the inverse-forward map is

$$\max_{\tilde{L}: \|\tilde{L} - L\|_F \leq \rho(\delta)} \|\tilde{L}L^\dagger(f^i - f^0) - (f^i - f^0)\|_2$$

which is known from robust linear squares [11] to be equivalent to the error

$$\|LL^\dagger(f^i - f^0) - (f^i - f^0)\|_2 + \rho(\delta)\|L^\dagger(f^i - f^0)\|_2.$$

Assuming that L has linearly independent rows, LL^\dagger is the identity matrix and thereby the error simplifies to

$$\rho(\delta)\|L^\dagger(f^i - f^0)\|_2.$$

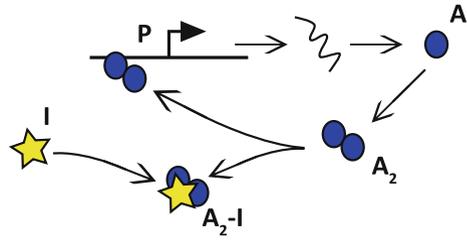
This result provides one way to determine the radius of the feature ball δ when relying on the pseudo-inverse

$$\begin{aligned} & \max_{\delta} \delta \\ & \text{subject to} \\ & \rho(\delta)\|L^\dagger(f - f^0)\|_2 \leq \varepsilon \\ & \|f - f^0\|_2 \leq \delta \end{aligned} \quad (10.8)$$

10.3 Simulation Results

As a proof of concept of our method, we applied it to a simple synthetic sensor construct [21]. The system is made of several gene copies (e.g. with plasmid transfection), expressing a protein that dimerizes and activates the gene by binding to the promoter. In presence of the inhibitor (input of the system), the dimer is trapped and

Fig. 10.3 Simple transcriptional sensor construct. The dimerized form (A_2) of a protein (A) is its own positive regulator; the inhibitor (I) tethers the dimer away in an inactive form ($A_2 - I$)



cannot bind to the promoter. A schematic of the involved reactions is depicted in Fig. 10.3.

The system is simulated according to mass-action and obeys

$$\begin{aligned}
 \frac{dx_1}{dt} &= k_1(x_5^0 - x_5) + k_2x_5 - k_3x_1 \\
 \frac{dx_2}{dt} &= k_4x_1 - 2k_5x_2^2 + 2k_6x_3 - k_{11}x_2 \\
 \frac{dx_3}{dt} &= k_5x_2^2 - k_6x_3 - k_7x_3y(t) + k_8x_4 - k_9(x_5^0 - x_5)x_3 + k_{10}x_5 - k_{11}x_3 \quad (10.9) \\
 \frac{dx_4}{dt} &= k_7x_3y(t) - k_8x_4 - k_{11}x_4 \\
 \frac{dx_5}{dt} &= k_9(x_5^0 - x_5)x_3 - k_{10}x_5 - k_{11}x_5.
 \end{aligned}$$

where the states x_i denote the concentration of mRNA, protein, protein-dimer and dimer-promoter complex, respectively. The quantities x_5^0 and $y(t)$ refer the total number of promoters and the external inhibitor concentration, respectively. The nominal value and the meaning of the model parameters are summarized in Table 10.1. We remark that such continuous state-space model have their limitations for transcriptional circuits because they require several gene copies in order to neglect the discrete Boolean nature of a single gene.

For the specified behavioral features, we expect the dimer to drop quickly after introduction of inhibitor and then quickly regain a high level after the inhibitor is washed out of the medium. We also constrain the monomeric protein. The specification functionals are the integral of the absolute difference to some target value $x^*(s)$ for the monomer and the dimer concentration over two small time intervals for each. More specifically,

$$\begin{bmatrix} \psi_1(x) \\ \psi_2(x) \end{bmatrix} = \begin{bmatrix} \int_0^T w_1(s)[x_2(s, k) - x_2^*(s)]^2 ds \\ \int_0^T w_2(s)[x_3(s, k) - x_3^*(s)]^2 ds \end{bmatrix} \quad (10.10)$$

where w is the temporal weight function chosen to be

Table 10.1 Nominal values and meaning of the kinetic parameters for the model of the synthetic sensor construct

Basal transcription rate	k_1	0.02 s^{-1}
Active-promoter transcription rate	k_2	0.4 s^{-1}
mRNA degradation rate	k_3	0.3 s^{-1}
Protein translation rate	k_4	3 (nm s)^{-1}
Dimerization rate	k_5	0.1 (nm s)^{-1}
Dimer dissociation rate	k_6	0.001 s^{-1}
Inhibitor binding rate	k_7	$0.011 \text{ (nm s)}^{-1}$
Inhibitor unbinding rate	k_8	0.2 s^{-1}
Dimer-promoter binding rate	k_9	0.21 (nm s)^{-1}
Dimer-promoter unbinding rate	k_{10}	0.2 s^{-1}
Protein degradation rate	k_{11}	0.2 s^{-1}

Values are based on [21] and slightly adapted to obtain a desired threshold behavior

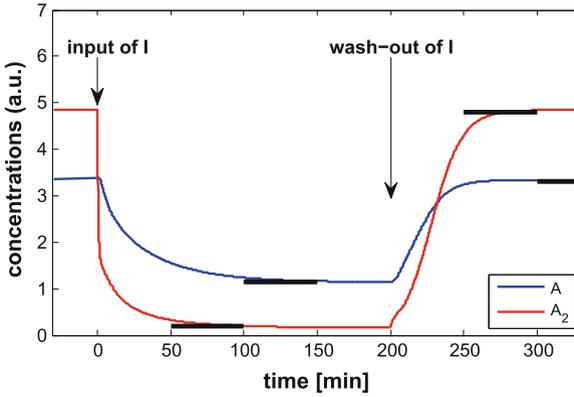


Fig. 10.4 Time courses of monomer (A , x_2) and dimer (A_2 , x_3) concentration of (10.9) for an addition and removal of the inhibitor (I , y); the target values and time intervals chosen for the specification functionals are indicated by *solid black lines*

$$w_i(t) = \begin{cases} 1 & \text{for } t \in [t_1, t_2] \cup [t_3, t_4] \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i = 1, 2 \quad (10.11)$$

The actual values for time-intervals for w_1 and w_2 , as well as the target values are shown together with the trajectories for the nominal system (10.9) in Fig. 10.4. For this case study we assume that we have means to design the binding rate of the inhibitor to the dimer k_7 and the binding rate of the dimer to the promoter k_9 . To assess the error incurred by the linearization we consider the reverse-forward mapping as described in (10.7). Hence for various size of δ we perform the inverse mapping with L^\dagger and the forward mapping with F . If the maps are exact we should obviously obtain a ball with the same δ . Any deviation ε thereof reflects the approximation of F^{-1} by L^\dagger . In Fig. 10.5 the images of $\mathcal{B}_{f_0}(\delta)$ under L^\dagger and $F \circ L^\dagger$ are shown

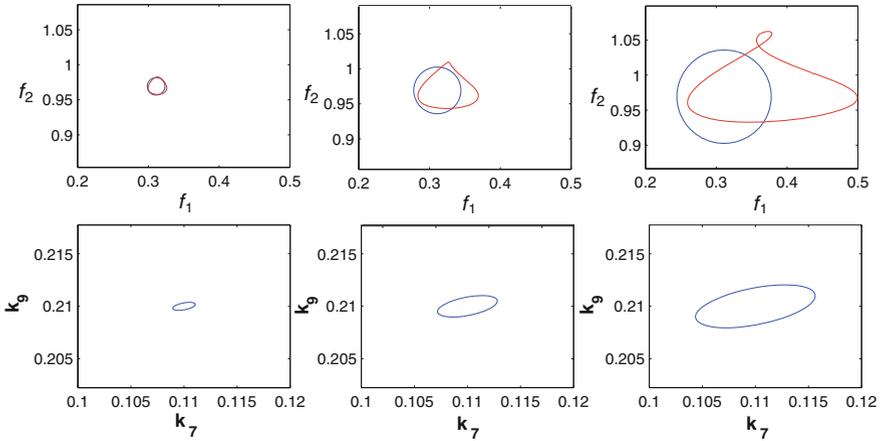


Fig. 10.5 Contours of $\mathcal{B}_{f_0}(\delta)$ (blue) in feature space (first row) are mapped back to the parameter space via L^\dagger (second row) and mapped forward using F (red) for increasing size of δ (from left to right)

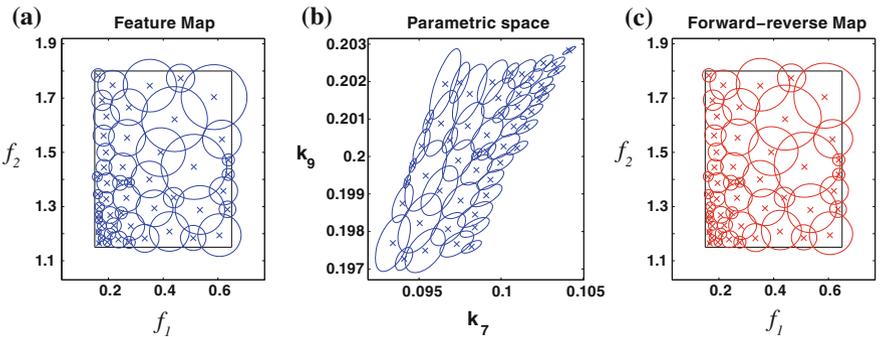


Fig. 10.6 Covering a certain specification range S (black rectangle) by overlapping balls, (a) which in turn yields overlapping ellipsoids in the parameter space (b). The precision of the mapping is illustrated by the reverse-forward map in (c). The centers of the balls are illustrated by crosses

for various radii δ . Hence, for an intermediate size of δ a good trade-off between approximation accuracy and sampling coverage is achievable. A systematic sampling of a predetermined specification area S would proceed by successively sampling overlapping balls with radii adapted to maintain ε under a certain value as illustrated in Fig. 10.6. In this example, the coverage of the region S is above 98 % using 50 balls of different radii. The lower left corner of the specification space (Fig. 10.6a) maps to a strongly nonlinear region of the parameter space (upper right corner in Fig. 10.6b) and therefore forces the use of smaller balls to keep the error in acceptable range. On the contrary, the upper right region of the specification space is more linear and larger balls can be used with limited relative error (Fig. 10.6c).

10.4 Conclusion

We presented a novel method to determine the parameter region of a biochemical reaction network that is consistent with a certain dynamical, behavioral specification. We defined specifications in a novel and general way that requires only the specification map to be once differentiable with respect to the states of the underlying differential equations. We showed that by locally linearizing this map we can solve the desired inverse problem of finding a parameter region for a given specification. As regions, instead of points, are mapped back to parameter space the scheme is in principle able to cover (given some regularity conditions) the feature and parameter space—something that is not possible with simple point sampling. We also discuss means for estimating the size of the local neighborhood in order to guarantee certain approximation errors. In this work, we considered classes of models based on ordinary differential equations, but the outlined framework can be extended to include stochastic dynamical models through the use of moment closure methods, for instance.

Acknowledgments H.K. acknowledges the support from the Swiss National Science Foundation (SNSF) grant number PP00P2_128503.

References

1. Nandagopal N, Elowitz MB (2011) Synthetic biology: integrated gene circuits. *Science* 333:1244–1248
2. Lu TK, Khalil AS, Collins JJ (2009) Next-generation synthetic gene networks. *Nat Biotechnol* 27:1139–1150
3. Bowsher CG, Swain PS (2012) Identifying sources of variation and the flow of information in biochemical networks. *Proc Natl Acad Sci USA* 109:E1320–E1328
4. Hilfinger A, Paulsson J (2011) Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proc Natl Acad Sci USA* 108:15004–15009
5. Zechner C, Ruess J, Krenn P, Pelet S, Peter M et al (2012) Moment-based inference predicts bimodality in transient gene expression. *Proc Natl Acad Sci USA* 109:8340–8345
6. Bleris L, Xie Z, Glass D, Adadey A, Sontag E et al (2011) Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Mol Syst Biol* 7:519
7. Miller M, Hafner M, Sontag E, Davidsohn N, Subramanian S et al (2012) Modular design of artificial tissue homeostasis: Robust control through synthetic cellular heterogeneity. *PLoS Comput Biol* 8:e1002579
8. Baier C, Katoen JP (2008) Principles of model checking. The MIT Press, London
9. Rizk A, Batt G, Fages F, Soliman S (2011) Continuous valuations of temporal logic specifications with applications to parameter optimization and robustness measures. *Theor Comput Sci* 412:2827–2839
10. Engl H, Hanke M, Neubauer A (1996) Regularization of inverse problems (Mathematics and its applications). Kluwer, Dordrecht
11. Christian HP (1998) Rank-Deficient and Discrete Ill-Posed Problems. *SIAM*, 51(3):455–500
12. Lu J, Engl HW, Schuster P (2006) Inverse bifurcation analysis: application to simple gene systems. *Algorithms Mol Biol* 1:11
13. Engl HW, Flamm C, Kügler P, Lu J, Müller S et al (2009) Inverse problems in systems biology. *Inverse Prob* 25:123014

14. Lu J, August E, Koepl H (2012) Inverse problems from biomedicine : inference of putative disease mechanisms and robust therapeutic strategies. *J Math Biol* 67(1):143–68
15. August E, Parker KH, Barahona M (2007) A dynamical model of lipoprotein metabolism. *Bull Math Biol* 69:1233–1254
16. Hafner M, Koepl H, Hasler M, Wagner A (2009) ‘Glocal’ robustness analysis and model discrimination for circadian oscillators. *PLoS Comput Biol* 5:e1000534
17. Zamora-Sillero E, Hafner M, Ibig A, Stelling J, Wagner A (2011) Efficient characterization of high-dimensional parameter spaces for systems biology. *BMC Syst Biol* 5:142
18. Vempala S (2005) Geometric random walks: a survey. *Comput Geom* 52:573–612
19. Hatzimanikatis V, Bailey JE (1996) MCA has more to say. *J Theor Biol* 182:233–242
20. Ingalls BP, Sauro HM (2003) Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories. *J Theor Biol* 222:23–36
21. Hooshangi S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci USA* 102:3581–3586

Chapter 11

Simplified Computational Design of Digital Synthetic Gene Circuits

Mario Andrea Marchisio and Jörg Stelling

Abstract One ultimate goal for synthetic biology is the complete computer-aided design of novel gene circuits. Here, we show how concepts and algorithms from electrical engineering can be exploited to set up a framework for the computational, automatic design of gene Boolean gates and devices. As in electrical engineering, the modular design of digital synthetic gene circuits can be automated via the Karnaugh map algorithm. However, differently from electronics, the circuit scheme corresponding to a Boolean formula is not unique since the wiring between gates can be established by transcription factors or small RNAs. In particular, we discuss a new, simplified version of our previous algorithm that is better tailored to wet-lab circuit implementation.

Keywords Digital circuits · Parts · Pools · Karnaugh map methods

11.1 Modular Design of Synthetic Gene Circuits

It is a central concept of synthetic biology, to design novel circuits—as in other engineering disciplines—by combining well-defined parts, devices, and higher-order functions [1]. For biological circuit design, however, this approach of bottom-up design of novel synthetic gene circuits faces several conceptual challenges. Fundamentally, the design requires the definition of functional modules, but it is not clear what the most suitable criteria for such definition in complex biological systems

M. A. Marchisio (✉) · J. Stelling
School of Life Science and Technology, HIT, Harbin, P. R. China
e-mail: marchisio@hit.edu.cn

J. Stelling
e-mail: joerg.stelling@bsse.ethz.ch

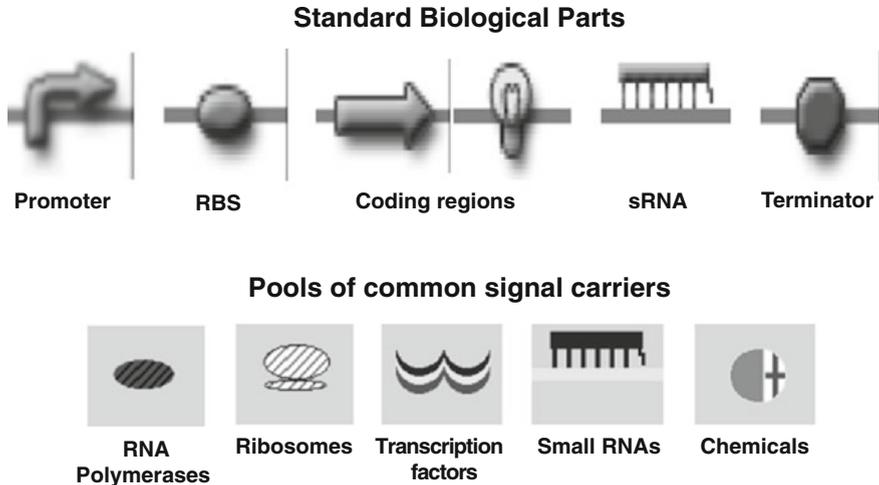


Fig. 11.1 Symbols of parts and pools. Symbols of standard biological parts (from the MIT registry) and pools used throughout the chapter (figure adapted from [3])

with non-linear dynamics could be [2]. The modules have to be *composable*, which means that they must share common input/output signals in order to be assembled into devices of higher complexity.

Such concepts often use analogies between biological and electrical engineering. In electrical devices, basic components such as resistors and capacitors are connected to each other inside circuits via the exchange of electrons. In a similar way, Standard Biological Parts are defined as DNA traits with well-characterized functions in transcription or translation. According to the MIT Registry of Standard Biological Parts (<http://parts.igem.org/>), parts are divided into categories such as promoters, ribosome binding sites (RBSs), coding regions, and terminators. These genetic parts are composable since they exchange several molecules referred to as *common signal carriers*. These carriers can be RNA polymerases and ribosomes [4], which lead to mRNA and protein synthesis. Similarly, transcription factors, small RNAs, and chemicals permit regulation among circuit components and interaction with the environment [5].

To enable computer-aided, rational design in line with these ideas, we previously developed a computational tool for the visual, modular design of synthetic gene circuits [5, 6]. In addition to standard parts, we introduced another kind of modules, namely *pools* that store free signal carriers, i.e., the circuit's biological *potentials* (see Fig. 11.1 for the graphical notation of modules and parts used in the following).

To represent the dynamic behavior of parts and pools, the modules are modeled according to full mass-action kinetics. These modular models are generated independently by running the corresponding Perl scripts that require, as input, parameter values and, when necessary, structural details such as the number of operators along a promoter sequence. Parts and pools are encoded into separate MDL (Model

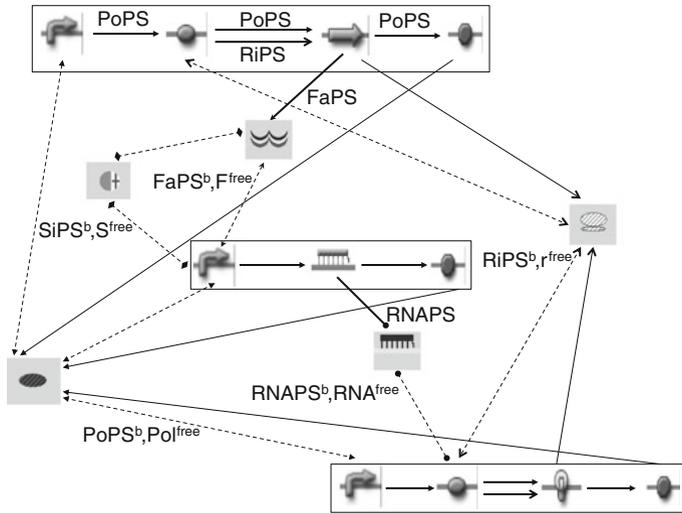


Fig. 11.2 Exchange of signal carriers among Parts and Pools. RNA polymerases flow through every standard biological part. The PoPS flux (Polymerase Per Second—*simple arrows, solid lines*) along each transcription unit (here represented as *boxes*) is originated by the interaction of free polymerases and promoters. Each promoter receives from the polymerase pool the amount of free molecules of polymerases (Pol^{free}) and it sends back a *balance flux* ($PoPS^b$) that arises from the binding and unbinding reactions of RNA polymerases with the DNA (*simple arrows, dashed lines*). Polymerases leave a transcription unit at the terminator and then return to their pool. Similarly, free ribosomes (r^{free}) interact with RBSs on the mRNA. Here, the RiPS flux (Ribosome Per Second—*line arrows*) starts and ends at coding regions where ribosomes are released again. Inside coding regions, moreover, transcription factors are produced. They flow directly to their pool (FaPS, Factor Per Second—*concave arrows*) and they are subsequently exchanged (F^{free} and $FaPS^b$) with the promoters they regulate. Furthermore, transcription factors can be activated or inhibited by chemicals (or environmental *signals*) inside the promoter or directly in their own pool. Free signal amount (S^{free}) and balance flux ($SiPS^b$, Signal Per Second—*diamond arrows*) are also relevant for RBSs when they contain riboswitches (not shown in the figure). Small RNAs are the last common signal carrier considered in our model. They are transcribed into an sRNA part and sent to their pool (RNAPS flux, RNA Per Second—*circle arrows*). From here, they get access to every RBSs they regulate (as RNA^{free} and RNA^b)

Definition Language) files [7] and can be handled within ProMoT [8]. In particular, ProMoT provides the user with a GUI where modules are deployed and connected via wires. Module interfaces are called *terminals*. Through them, parts have access to the current concentrations of free signal carriers in pools and parts exchange (with another part or a pool) fluxes of signal carriers, that is, biological *currents* (see Fig. 11.2 for a more detailed representation).

While the basic infrastructure for computer-aided design, thus, is available, circuit construction in most cases is still manual. In contrast to electrical circuit design, however, few methods for the automated design—in the sense of generating a circuit structure that allows to meet user-defined, functional performance specification—are

available for synthetic gene circuits. Here, we focus on the simplified computational design for a particular class of gene circuits, namely synthetic devices for arbitrary computations of Boolean logic.

11.2 Automatic Circuit Design

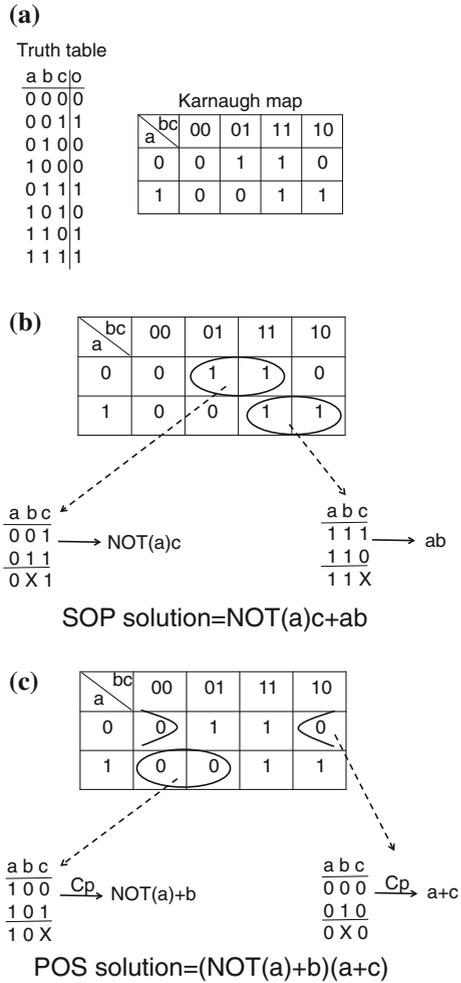
The design of a genetic circuit able to accomplish a specific task is, in general, not an easy operation. First, one has to determine the circuit structure. This step requires one to gather parts into bigger modules, such as motifs [9], whose dynamics is known; combinatorial effects may lead to an explosion of structural design alternatives here. Afterwards, genetic parts themselves have to be *identified*. This corresponds to finding those values of their kinetic parameters that make the overall circuit work as desired. Computational solutions that aim at the automatic design of genetic circuits starting from a *target* output function have been proposed [10–13]. They are all based on stochastic algorithms that perform a double optimization procedure in order to derive both circuit structure and parameter values. This approach, however, does not scale linearly (or polynomially) with the circuit size, and it might therefore require the adoption of very simplified biological models to describe circuit components (e.g., by representing translation as a single-step event).

As we showed in a previous publication [3], *digital* gene circuits can be designed in a simpler way. For this class of circuits, the truth table replaces any target function and it fully specifies the relation between circuit inputs (externally added chemicals or signals from other synthetic or natural biological circuits) and output (fluorescence, for instance). Moreover, via the Karnaugh map algorithm [14], the truth table can be converted directly into a Boolean formula from which the circuit scheme—where Boolean gates are organized in two or three layers—follows immediately. Hence, the circuit structure is given without a need for any optimization procedure. We will illustrate these concepts with a simple example in the following.

11.2.1 Bringing Electronics into Biology: The Karnaugh Map Method

The truth table of a logic circuit represents the relations between (multiple) inputs and outputs in terms of ‘0’ and ‘1’ states, and every row of the table pertains to a specific combination of discrete inputs and the resulting outputs. Every truth table is associated with two different minimal Boolean formulas: the CNF (Conjunctive Normal Forms) formula and the DNF (Disjunctive Normal Form) formula. In electronics, the former is called POS (Product Of Sums), the latter SOP (Sum Of Products). For any truth table, both Boolean formulas can (in principle) be derived with the Karnaugh map method.

Fig. 11.3 The Karnaugh map method. **a** A Karnaugh map is a rearrangement of the content of a truth table. In the first column *a* values are written, whereas the first row contains the *b* and *c* bits together. Their sequence 00, 01, 11, 10 is an example of a Grey code. **b** The SOP formula is obtained after grouping the 1s present into the Karnaugh map. **c** Different from the SOP formula, clauses in POS are determined by considering the complement (here indicated as *Cp*) of each variable whose digit is conserved within a group. Moreover, each clause contains a logic sum (OR) of input variables and the whole formula is a logic product (AND) of clauses



A Karnaugh map is simply a rearrangement of a truth table where two adjacent inputs differ at most by one digit only (a so called *Grey code* sequence—see Fig. 11.3a). In order to extract the SOP Boolean formula from a Karnaugh map, one has to group all the 1s in the table. Each group can contain only a power of two of 1s and it needs to have a square or rectangular shape, which means that 1s cannot be gathered in diagonal. First, one should try to construct groups as big as possible and then reduce their dimension, leaving groups of single 1s only at the very last step. Notice that overlaps between different groups are also allowed. As shown in Fig. 11.3b, in our example we have four 1s that can be gathered into two groups of dimension two. Each group corresponds to a *clause* in the formula. Only variables whose digit stays constant inside the group appear in the clause: if their binary value

is 0, they are negated. In this case, both clauses represent a logic product (AND) of their inputs. The whole SOP formula is finally given by the logic addition (OR) of the two clauses. As for the POS formula, an analogous procedure holds where, however, 0s have to be gathered instead (see Fig. 11.3c for further details).

Both for SOP and POS, a circuit can be organized in three layers of gates: an *input* layer containing YES (these gates exist only in biology and have no electronic counterpart) and NOT gates for *true* and *false* input values, respectively; an *internal* layer made of as many gates (AND in SOP, OR in POS) as the clauses in the Boolean formula; a *final* layer comprising a single gate: AND in POS and OR in SOP. However, in a SOP circuit, every AND gate may produce the circuit output, making the presence of the final OR gate unnecessary (see Fig. 11.10a). This *distributed output* architecture has been recently shown to be functional for the design of Boolean circuits [15].

Furthermore, recall that one of the *De Morgan* laws [16] claims that

$$\overline{(a + b)} = \bar{a}\bar{b}, \quad (11.1)$$

where \bar{a} denotes the negation (NOT) of signal a . Therefore, the POS formula in our example can be rewritten as

$$POS = \overline{(\bar{a}c) + (a\bar{b})}. \quad (11.2)$$

Hence, both SOP and POS require only three kinds of gates: YES, NOT, and AND. Moreover, the POS scheme that results from Eq. (11.2) looks like a SOP one where the final OR gate is replaced by a NOT that takes, as the only input, a negative regulatory factor produced by all the AND gates in the internal layer (see Fig. 11.11a).

Overall, thus, standard methods for logic circuit design allow both a systematic and a potentially compact (in terms of implementation) construction of alternative digital circuit design. The critical question, however, is to what extent these alternatives can be realized in biological systems, and how uncertainty, non-linearity, and other characteristics of biological systems affect their performance.

11.2.2 Gene Boolean Gates

Although different designs of diverse complexity are possible for engineering genetic Boolean gates (see for instance [17]), the most common ways to confer a digital behavior to genes and their products employ the transcriptional control via designed promoters [18, 19] and the translation regulation at the mRNA level using structures such as riboswitches/ribozymes [20] or antisense RNA base-pairing [21, 22].

Here, we consider chemicals as inputs for our digital circuits. They fall into two categories that, following the notation in [23], are referred to as *inducers* (i) such as IPTG (acting on the Lac repressor) and *corepressors* (c) such as tetracycline (binding to the Tet inverse repressor). Chemicals regulate transcription by binding

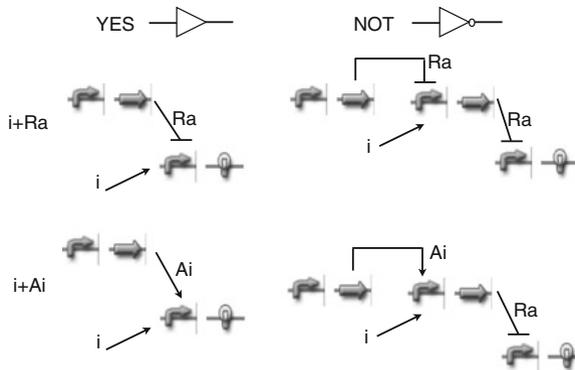


Fig. 11.4 Inducer-based gates. YES gates produce either active repressors (Ras) or inactive activators (Ais) to which inducers bind. NOT gates require one more gene in a two-step cascade. Notice that throughout the gate and circuit schemes in this chapter, terminators have been always omitted, and RBSs are only shown when they are regulated

to transcription factor proteins (repressors or activators) and modifying their state (active or inactive), and therefore their action on the DNA. Moreover, chemicals can control translation as well, by binding riboswitches or ribozymes, thereby causing structural changes of the mRNA.

We extend our previous Boolean gate design [3] by allowing inducers and corepressors to interact both with repressors and activators. In particular, inducers bind active repressors (Ra) and inactive activators (Ai). If we put a gene expressing a fluorescent protein under the control of either transcription factor, we will register fluorescence only in the presence of inducer molecules: this is what we call a YES gate. To have a NOT gate, on the contrary, we shall require that fluorescence is expressed only in the absence of inducers. This is achieved via a two-step cascade where the chemically regulated Ra or Ai acts on a downstream gene that encodes a negative regulatory factor (either a repressors or an sRNA) that, in its turn, controls the synthesis of the fluorescent protein (see Figs. 11.4 and 11.6). Only when inducers are present, the promoter in the first step of the cascade is either freed from repressors or bound by activators. In both cases, the negative factor is produced and fluorescence is switched off. We shall notice that in this work, small RNAs are assumed to behave only as translational repressors: this is, indeed, their most common function, although it was reported that they can also activate protein synthesis [24].

A complementary picture holds for corepressors. They bind either inactive repressors (Ri) or active activators (Aa) and realize “naturally” NOT gates, whereas YES gates require the two-step cascade configuration (see Figs. 11.5 and 11.6).

In analogy to [15], with this design of YES and NOT gates, AND gates are simply transcription units that are regulated by the outputs of the input gates, that is, regulatory factors that represent *true* and *false* variables (see Fig. 11.7). However, when a promoter is under the action of two activators, they have to act *cooperatively* in order to mimic the AND behavior properly.

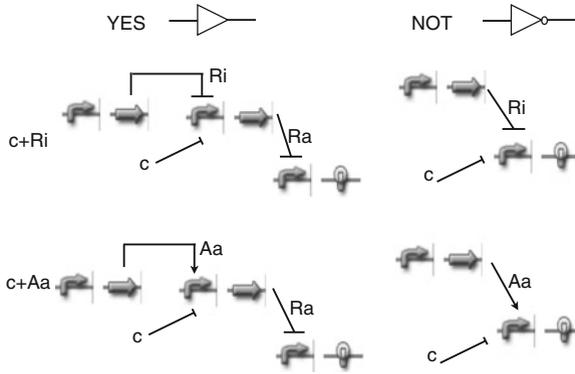


Fig. 11.5 Corepressor-based gates. NOT gates synthesize either *Ri*s or *Aa*s on which corepressors exert their action. YES gates require, on the contrary, the more complex two-step cascade structure

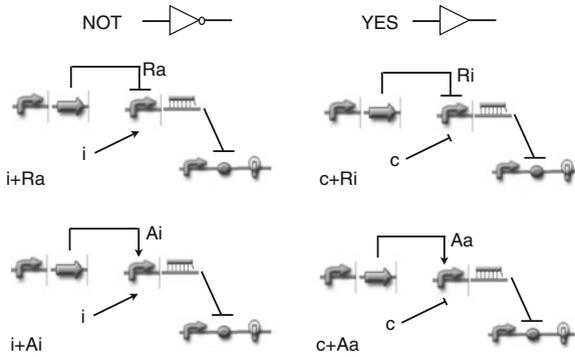


Fig. 11.6 sRNA-based gates. Configuration of NOT gates (for inducers) and YES gates (for corepressors) where the transcription unit producing the fluorescent signal is regulated by an sRNA at the RBS

Fig. 11.7 AND gates. The logic variable *a* is assigned to an inducer that inhibits a repressor *Ra*, whereas *b* is associated with a corepressor that inactivates an activator *Aa*. Their Boolean product, $(a\bar{b})$, is obtained simply by putting a promoter (which leads the expression of a fluorescent protein) under the control of both *Ra* and *Aa*

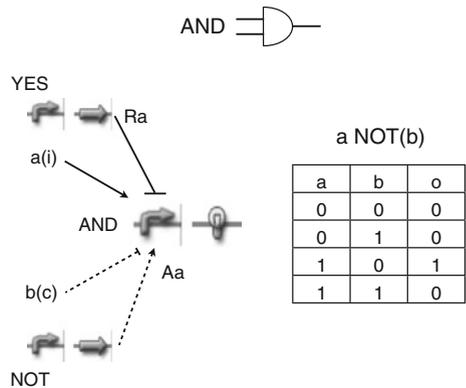
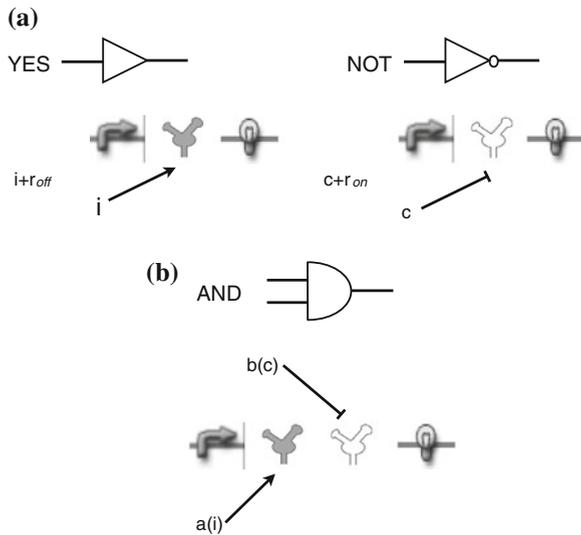


Fig. 11.8 Riboswitches-based gates. **a** Inducers activate riboswitches that are in the default ‘off’ state (r_{off}), allowing fluorescence production (YES). Corepressors, on the contrary, switch off the synthesis of fluorescent proteins by inactivating ‘on’ riboswitches (r_{on}). In the gates’ schemes, RBSs have been replaced with hairpins representing riboswitches. **b** Riboswitches-based configuration of the AND gate corresponding to the Boolean formula $a\bar{b}$. The truth table is shown in Fig. 11.7



Finally, in our representation, AND gates can contain riboswitches along their mRNA (on the RBS, since we consider bacterial cells). They are either activated by inducers (YES) or deactivated by corepressors (NOT—see Fig. 11.8).

Overall, thus, digital circuit design for biology provides different design alternatives, in terms of the delineation of circuit structures as well as of the basic implementations of biological gates. For practical applications, it is therefore essential to evaluate the performance of these alternatives systematically; this is the topic of the following section.

11.3 Design Alternatives for Digital Gene Circuits

For the evaluation of design alternatives, we will focus on the comparison of (dynamic mathematical models for) POS- and SOP-based design. Both of them are generated automatically through the computational methods described below.

11.3.1 The Computational Tool

In order to run our computational tool for the simplified automatic design of synthetic gene digital circuits, a user has to specify the inputs, that is, the chemicals’ type (i or c) and the corresponding regulatory factors (transcription factors or riboswitches) together with the circuit truth table. Although the Karnaugh map method works

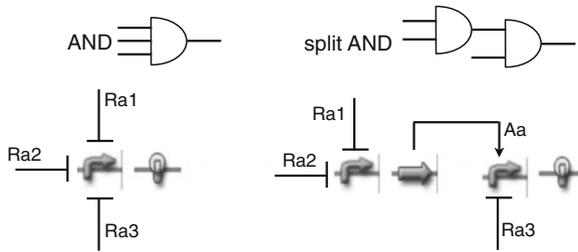


Fig. 11.9 Split AND gate configuration. In our model, AND gates cannot take three inputs of the same kind such as the tree repressors depicted here (each repressor is supposed to bind a different inducer; not shown). The first two repressors stay on the original AND gate that, however, does not produce any more a fluorescent protein but an activator. This activator regulates, together with the third repressor, a new AND gate where fluorescence is now produced

efficiently up to 6 inputs, current limitations in the implementation of engineered synthetic gene circuits (primarily due to the limited number of orthogonal components, see also [25]) forced us to restrict our algorithms to the cases of 2 or 3 input chemicals.

After calculating both SOP and POS circuit formulas, all the possible YES and NOT gates compatible with the rules presented in the previous section are generated. Hence, *false* inducers that act on a transcription factor are associated with two possible NOT gates that produce a repressor and an sRNA, respectively. Similarly, *true* corepressors will correspond to two diverse YES gates. Each combination of the available YES and NOT gates gives rise to a different circuit scheme. Provided that no riboswitches are involved, a solution is always *purely transcriptional*, that means it contains only repressors and activators whereas all the other solutions show a mixture of transcription factors and small RNAs to transduce the input signal from the input to the final layer of gates.

In our model, promoters and RBSs cannot be regulated by more than two factors. Whenever an AND gate requires either three different transcription factors or sRNAs, it is split into two new AND gates that are regulated by two factors only. The first AND gate receives two of the original inputs and it produces an activator that regulates the second AND gate together with the third original input (see Fig. 11.9). A combined control of transcription and translation is always allowed for AND gates.

As mentioned above, SOP circuits make use of the distributed output architecture, whereas POS designs require a final NOT gate. The latter, in general, is controlled by a repressor that is the common output of the internal layer. However, if none of the AND-gate RBSs is regulated either by sRNAs or by chemicals, a further solution is taken into account where all the AND gates produce a small RNA to control the final NOT gate.

Solutions both in SOP and POS are ranked according to the *complexity score* S defined as

$$S = 2^{n_R-1} + 2^{n_A-1} + n_S; \quad n_R, n_A \geq 1 \quad (11.3)$$

where n_R is the number of repressors in the circuit, n_A the one of activators, and n_s is the amount of small RNAs present in the network. This score aims at representing the practical effort for implementing a circuit and it is based on the assumption that transcription factors are more difficult to be engineered than small RNAs [3]. Hence, Eq. (11.3) favors the usage of translational control mechanisms. However, recent progress in DNA-binding-factor engineering due to zinc-finger proteins [26] and TAL effectors [27] might force us to revise, in the future, our definition of circuit complexity.

Finally, a user can choose a solution both in SOP and POS: they are encoded into MDL files and can be visualized (and modified, if desired) with ProMoT. From here, solutions can be exported into formats suitable for simulations such as SBML [28] and Matlab (MathWorks, Nantucket/MA).

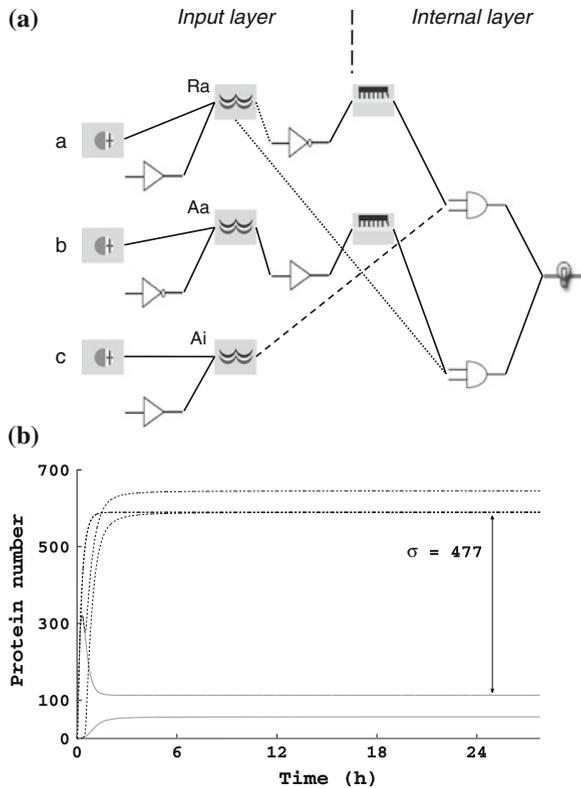
11.3.2 SOP and POS Solution: A Comparison

To illustrate the dynamics and the performance of the digital circuits designed by our tool, we simulated the time-dependent behavior of the SOP and POS solutions that minimize the complexity score S to reproduce the truth table shown in Fig. 11.3. Both the minimal circuits have a low complexity score of $S = 5$; note that the complexity scores for all design alternatives cover a range between $S = 5$ and $S = 9$. However, whereas the minimal SOP solution employs three transcription factors (two activators and one repressor) together with two sRNAs (Fig. 11.10a) and a total of seven genes, the POS scheme needs two activators, two repressors, and one small RNA, leading to eight genes overall (the extra gene is essentially due to the final NOT gate—see Fig. 11.11a). Therefore, the SOP solution should be preferable for a wet-lab implementation.

To characterize the circuit performance, we employ dynamic simulation of the ordinary differential equation-based model generated by the circuit design tool, using standard settings for all kinetic parameters. When we characterize the circuit performance with the signal separation between the 1 and 0 outputs (σ , calculated as the difference between the minimal steady-state protein number for a 1 and the maximal for a 0—for other possible parameters see [3]), the SOP circuit clearly outperforms the POS circuit (see Figs. 11.10b and 11.11b). Moreover, the SOP implementation seems to respond faster to changes in the input since all the eight possible outputs—one for each truth table entry—reach the steady state within 12 h, whereas several responses of the POS circuit take more than twice as much.

Despite all these advantages, the SOP solution shows the non-negligible drawback—due to the distributed-output architecture—of producing more than 50 proteins for some of its 0 outputs. This might be a problem if the circuit were integrated into a more complex structure where its output has to control some downstream components. In this case, the POS solution would be a better choice since almost no protein is present in correspondence to any 0 output.

Fig. 11.10 SOP solution. **a** Scheme of the minimal SOP solution. A pool of reporter proteins symbolizes the distributed-output architecture. **b** Time-dependent circuit simulations. For each truth table entry, the circuit reaches, first, a steady state in absence of chemicals (not shown) and afterwards it is fed with the specific combination of input molecules. Trajectories for input combinations that are designed to yield logic 1s in the output are *black, dashed lines* whereas the 0s for outputs are *gray, solid lines*



The example illustrates that, in synthetic circuit engineering, usually several performance criteria are important. Conflicting criteria and their corresponding trade-offs require the investigation of potentially large numbers of design alternatives, which is not possible by manual circuit design and analysis.

11.4 Lessons From Electronics

Methods and algorithms developed in electrical engineering for the design and construction of complex devices are now being applied to biology in order to implement novel gene circuits in a rational and systematic way. Despite the intrinsic complexity of biological systems, simple modules (i.e., repetitive patterns) are found in nature and they serve as basic building blocks for the synthesis of synthetic constructs based on DNA manipulation. Moreover, concepts such as signal carriers and pools (which represent “bio-electrons” and bio-batteries, respectively) permit us to give a visual representation of the interconnections and the possible cross-talk among gene circuit

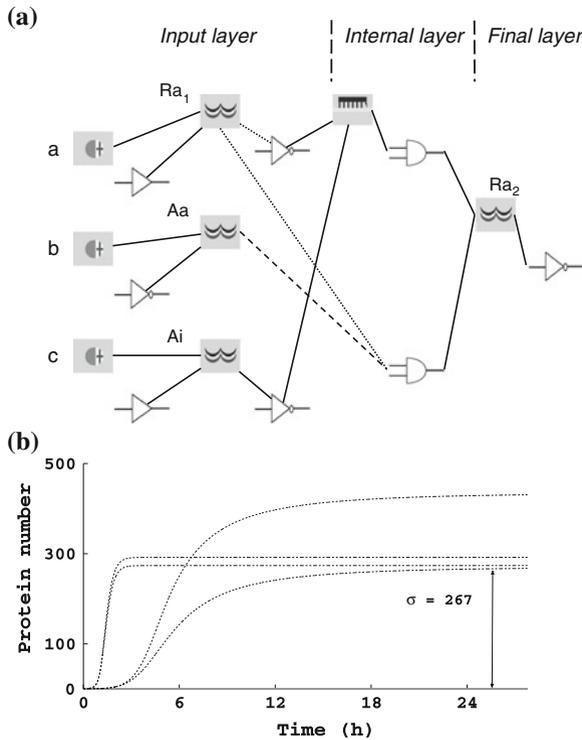


Fig. 11.11 POS solution. **a** Scheme of the minimal POS solution. The final NOT gate is the only gene responsible for the output production. **b** Time-dependent circuit simulations in analogy to the SOP solution (Fig. 11.10b)

components. Finally, algorithms such as the Karnaugh map method provide synthetic biologists with rules for the design of new systems that are able to perform well-defined functions such as sensing chemicals or carrying out computations [29].

11.5 Conclusions and Perspectives

We presented a simplified version of our original method for the automatic design of digital synthetic gene circuits [3]. Specifically, we reduced the maximal input number from four to three, but we allowed for more degrees of freedom in circuit design: the two kinds of chemicals considered (inducers and corepressors) may interact both with repressor and activator proteins. YES and NOT gates are constructed in a new and less complex way. They convert the input signals into transcription factors and/or small RNAs. AND gates are controlled by these regulatory factors or, alternatively, directly by chemicals when they host riboswitches on their RBSs. Finally, with a distributed

output architecture, circuits in SOP no longer require the final gate whereas circuits in POS, though still organized in three layers, do not need any NOR gate.

More generally, the example of digital circuit design shows that gene circuits could be re-shaped according to the experimental technology available, in our case just by varying the rules for engineering YES, NOT, and AND gates. At least for digital signal processing in engineered cells or cell populations, we anticipate that very complex Boolean logic networks might be built in the future by decomposing them into functional submodules that communicate within a cellular consortium [15]. For this, bio-chemical mechanisms different from mere transcription and translation regulation could be also exploited [30]. Ultimately, the most suitable circuit design, however, will be determined by (trade-offs) performance specifications as well as by practical implementability in biology.

References

1. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2:2006.0028. doi:10.1038/msb4100073. <http://dx.doi.org/10.1038/msb4100073>
2. Hans-Michael K, Stelling J (2012) Modular analysis of biological networks. *Adv Exp Med Biol* 736:3–17
3. Marchisio MA, Stelling J (2011) Automatic design of digital synthetic gene circuits. *PLoS Comput Biol* 7(2):e1001083. doi:10.1371/journal.pcbi.1001083. <http://dx.doi.org/10.1371/journal.pcbi.1001083>
4. Endy D (2005) Foundations for engineering biology. *Nature* 438(7067):449–453. doi:10.1038/nature04342. <http://dx.doi.org/10.1038/nature04342>
5. Marchisio MA, Stelling J (2008) Computational design of synthetic gene circuits with composable parts. *Bioinformatics*, 24(17):1903–1910. doi:10.1093/bioinformatics/btn330. <http://dx.doi.org/10.1093/bioinformatics/btn330>
6. Marchisio MA, Stelling J (2009) Synthetic gene network computational design. In: Proceedings of IEEE international symposium on circuits and systems, ISCAS 2009, pp 309–312
7. Ginkel M, Kremling A, Nutsch T, Rehner R, Gilles ED (2003) Modular modeling of cellular systems with ProMoT/Diva. *Bioinformatics* 19(9):1169–1176
8. Mirschel S, Steinmetz K, Rempel M, Ginkel M, Gilles ED (2009) PROMOT: modular modeling for systems biology. *Bioinformatics* 25(5):687–689. doi:10.1093/bioinformatics/btp029. <http://dx.doi.org/10.1093/bioinformatics/btp029>
9. Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* 8(6):450–461. doi:10.1038/nrg2102. <http://dx.doi.org/10.1038/nrg2102>
10. Dasika MS, Maranas CD (2008) OptCircuit: an optimization based method for computational design of genetic circuits. *BMC Syst Biol* 2:24. doi:10.1186/1752-0509-2-24. <http://dx.doi.org/10.1186/1752-0509-2-24>
11. François P, Hakim V (2004) Design of genetic networks with specified functions by evolution in silico. *Proc Natl Acad Sci U S A* 101(2):580–585. doi:10.1073/pnas.0304532101. <http://dx.doi.org/10.1073/pnas.0304532101>
12. Paladugu SR, Chickarmane V, Deckard A, Frumkin JP, McCormack M, Sauro HM (2006) In silico evolution of functional modules in biochemical networks. *Syst Biol (Stevenage)* 153(4):223–235
13. Rodrigo G, Carrera J, Jaramillo A (2007) Genetdes: automatic design of transcriptional networks. *Bioinformatics* 23(14):1857–1858. doi:10.1093/bioinformatics/btm237. <http://dx.doi.org/10.1093/bioinformatics/btm237>

14. Maurice K (1953) The map method for synthesis of combinational logic circuits. *Trans Am Inst Electr Eng* 72(9):593–599
15. Regot S, Macia J, Conde N, Furukawa K, Kjelln J, Peeters T, Hohmann S, de Nadal EL, Posas F, Sol R (2011) Distributed biological computation with multicellular engineered networks. *Nature* 469(7329):207–211. doi:[10.1038/nature09679](https://doi.org/10.1038/nature09679). <http://dx.doi.org/10.1038/nature09679>
16. Kuphaldt TR (2007) Lessons in electric circuits ,Vol 4(Digital). <http://www.ibiblio.org/obp/electricCircuits>
17. Terzer M, Jovanovic M, Choutko A, Nikolayeva O, Korn A, Brockhoff D, Zurcher F, Friedmann M, Schutz R, Zitzler E, Stelling J, Panke S (2007) Design of a biological half adder. *IET Synth Biol* 1(1–2):53–58
18. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Kuhlman T, Phillips R (2005) Transcriptional regulation by the numbers: applications. *Curr Opin Genet Dev* 15(2):125–135. doi:[10.1016/j.gde.2005.02.006](https://doi.org/10.1016/j.gde.2005.02.006). <http://dx.doi.org/10.1016/j.gde.2005.02.006>
19. Silva-Rocha R, de Lorenzo V (2008) Mining logic gates in prokaryotic transcriptional regulation networks. *FEBS Lett* 582(8):1237–1244. doi:[10.1016/j.febslet.2008.01.060](https://doi.org/10.1016/j.febslet.2008.01.060). <http://dx.doi.org/10.1016/j.febslet.2008.01.060>
20. Win MN, Smolke CD (2008) Higher-order cellular information processing with synthetic RNA devices. *Science* 322(5900):456–460. doi:[10.1126/science.1160311](https://doi.org/10.1126/science.1160311). <http://dx.doi.org/10.1126/science.1160311>
21. Lucks JB, Qi L, Mutalik VK, Wang D, Arkin AP (2011) Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc Natl Acad Sci U S A* 108(21):8617–8622. doi:[10.1073/pnas.1015741108](https://doi.org/10.1073/pnas.1015741108). <http://dx.doi.org/10.1073/pnas.1015741108>
22. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y (2007) A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol* 25(7):795–801. doi:[10.1038/nbt1307](https://doi.org/10.1038/nbt1307). <http://dx.doi.org/10.1038/nbt1307>
23. Lewin B (2000) *Genes VII*. Oxford University Press, New York
24. Isaacs FJ, Dwyer DJ, Collins JJ (2006) RNA synthetic biology. *Nat Biotechnol* 24(5):545–554. doi:[10.1038/nbt1208](https://doi.org/10.1038/nbt1208). <http://dx.doi.org/10.1038/nbt1208>
25. Purnick PEM, Weiss R (2009) The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* 10(6):410–422. doi:[10.1038/nrm2698](https://doi.org/10.1038/nrm2698). <http://dx.doi.org/10.1038/nrm2698>
26. Lohmueller JJ, Armel TZ, Silver PA (2012) A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res.* doi:[10.1093/nar/gks142](https://doi.org/10.1093/nar/gks142). <http://dx.doi.org/10.1093/nar/gks142>
27. Bogdanove AJ, Voytas DF (2011) TAL effectors: customizable proteins for DNA targeting. *Science* 333(6051):1843–1846. doi:[10.1126/science.1204094](https://doi.org/10.1126/science.1204094). <http://dx.doi.org/10.1126/science.1204094>
28. Hucka M et al (Mar 2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19(4):524–531
29. Benenson Y (2012) Biomolecular computing systems: principles, progress and potential. *Nat Rev Genet* 13(7):455–468. doi:[10.1038/nrg3197](https://doi.org/10.1038/nrg3197). <http://dx.doi.org/10.1038/nrg3197>
30. Marchisio MA, Rudolf F (2011) Synthetic biosensing systems. *Int J Biochem Cell Biol* 43(3):310–319. doi:[10.1016/j.biocel.2010.11.012](https://doi.org/10.1016/j.biocel.2010.11.012). <http://dx.doi.org/10.1016/j.biocel.2010.11.012>

Chapter 12

Synthetic Biochemical Devices for Programmable Dynamic Behavior

Jongmin Kim and Elisa Franco

Abstract The ability to harness biomolecules as tools for systematic engineering is fundamental to future developments of biotechnology and nanotechnology. Especially suitable for such applications are nucleic acid-based circuits with predictable interactions, allowing for rational design of circuit functions and dynamics. Here, we focus on synthetic transcriptional circuits utilizing the modular architecture of nucleic acid templates and the catalytic power of natural enzymes. The programmability of dynamic behaviors for synthetic circuits is illustrated through elementary circuits such as an adapter, a bistable switch, and several oscillators. Further, the effect of downstream processes on the central dynamical system illustrates the need for systematic methods of composing biomolecular circuits. We present insulating and amplifying devices as a solution for scaling up biomolecular networks much as in electronic circuits. Future applications of biomolecular programs will open up new possibilities in nanorobotics, nanomedicine, and artificial cells.

Keywords Nucleic acid-based circuits · Watson-Crick base pair interactions · Nucleic acid · Enzyme · Transcriptional circuit · Protein · DNA · RNA · Nucleotide · RNA polymerase · Promoter · Hybridization

J. Kim (✉)

Department of Bioengineering, California Institute of Technology, 1200 California Blvd,
Pasadena, CA 91125, USA
e-mail: jongmin@dna.caltech.edu

E. Franco

Department of Mechanical Engineering, University of California at Riverside,
900 University Avenue, Riverside, CA 92521, USA
e-mail: efranco@enr.ucr.edu

12.1 Introduction

Biology is the science of the century [9]. Understanding the molecular processes underlying cellular functions is fundamental to advance our ability to cure diseases and improve the quality of life. In parallel, our increasing biological knowledge has opened up unprecedented opportunities to use biological molecules as components to build new, synthetic devices for a variety of purposes, including molecular computing, nanofabrication, and new biomaterials [1, 4, 11, 19, 24].

In this picture, nucleic acids have not only emerged as fundamental molecules for storing and propagating genetic information, but also as an excellent biochemical material to build synthetic biodevices, including nanostructures and circuits. The specificity of Watson-Crick base pair interactions allows us to rationally design systems of strands that bind according to desired spatial and temporal patterns. By simply choosing the sequence of bases in each strand, we determine the binding affinity, and thus the overall supramolecular configuration, of the strands ensemble. Thanks to refined knowledge of base-pairing thermodynamic parameters, models allow us to both predict the structure corresponding to a given sequence, and also to compute candidate optimal sequences that will yield a desired structure [34, 64]. Interactions of nucleic acids with other ligands cannot be rationally designed, but they can be achieved through directed evolution techniques [10, 54]. Recently, a variety of modified nucleic acids, presenting improved thermodynamic properties, have been chemically synthesized [22]. Finally, custom synthesis of nucleic acid sequences is inexpensive (from \$0.2 per base-pair) and reliable. Thus, nucleic acids are one of the most versatile self-assembling materials for nanotechnology: we can use them to rationally build circuits [41] and nanostructures [40], and we can interface these devices with a variety of materials and other biological molecules [30, 35, 56].

In this chapter, we focus our attention on synthetic transcriptional systems, dynamic molecular circuits that can be built using nucleic acids and off-the-shelf enzymes. These circuits recapitulate gene regulatory processes in cells with a reduced number of components: therefore, they are a promising toolkit to rationally reproduce many functions that are unique to cells and not yet achievable by any human-designed material or system, such as spatial adaptability, self-repair, and replication. Thus far, transcriptional circuits have been employed to build basic components including bistable, oscillating, and adapting systems [26–28]. The integration of all these components will allow us to build multi-layered devices of growing complexity; however, systematic methods allowing for a modular integration are needed to prevent undesired interactions and perturbations typical of molecular networks in the absence of compartments [15].

We begin this chapter with a general introduction to nucleic acids and proteins as molecular materials, and then we describe how they can be employed to build synthetic reaction primitives and thus dynamics. An overview of available transcriptional modules, focused on oscillators and adapting circuits, is followed by a discussion on

scalability and signal transmission in interconnected circuits. Finally, we provide an outlook on the future applications of this toolkit.

12.2 Molecular Parts

In this section we briefly and broadly introduce nucleic acids and proteins as components to construct dynamic, artificial biochemical circuits.

Nucleic acids A nucleic acid strand is a long polymer made from repeating units called nucleotides. Each nucleotide contains both a backbone segment, which holds the chain together, and a base, which interacts with other bases forming hydrogen bonds. Natural nucleic acids are RNA and DNA, which differ in the backbone chemical composition (Ribose in RNA and Deoxyribose for DNA) and in one element in the set of bases: Adenine, Guanine, and Cytosine (A, G, and C) are common to both DNA and RNA; Thymine (T) is found in DNA, and Uracil (U) in RNA. The thermodynamic interactions causing bases to bind are well characterized [32]; the canonical Watson-Crick binding patterns are A–T (A–U in RNA) and C–G (non-Watson-Crick interactions e.g. G–U can contribute significant energy as well). The backbone molecule is asymmetric, and its direction is defined by the geometry of the sugar molecule, in particular the 3' and 5' position of carbon atoms in the ring, see Fig. 12.1a. Thus, a sequence of several nucleotides is always written in the 5'–3' direction. Binding of bases results in the well-known double helix structure of DNA (in nature, DNA typically occurs in a double-stranded form); however, more generally, nucleic acid strands “fold” into secondary structures that result in the configuration with minimum free energy from a thermodynamic point of view. Thus, single-stranded nucleic acids can fold onto themselves forming loops and hairpins, and bind to other strands forming multi-stranded complexes. These folding interactions occur spontaneously, thus secondary and tertiary structures of single or ensembles of strands are said to self-assemble.

Due to the well-understood specificity of Watson-Crick base-pairing interactions, DNA and RNA have emerged out of their biological role of information storage and transport, and are repurposed as an easily programmable material for nanoscale construction [8, 45]. Notable examples of self-assembled nanostructures include systems of tiles, origami, and 3-D lego-like blocks [7, 23, 45, 46]. As we will detail in the next sections, the principle of self-assembly can be used to build logic and dynamic circuits. A significant advantage of nucleic acids is that desired binding patterns among strands can be easily specified by designing complementary domains at an abstract level; domain examples, highlighted in different colors, are shown in Fig. 12.1b, c. Once domain-level strand interactions are chosen, specific candidate sequences can be designed systematically with one of the many available compilers [31] and design toolboxes such as NUPACK [64], DSD [39] and DNADesign (developed in the E. Winfree lab at Caltech). These software toolboxes operate much like optimization suites, where base sequences are optimized to guarantee that the minimum folding energy of the system occurs at the desired binding configuration,

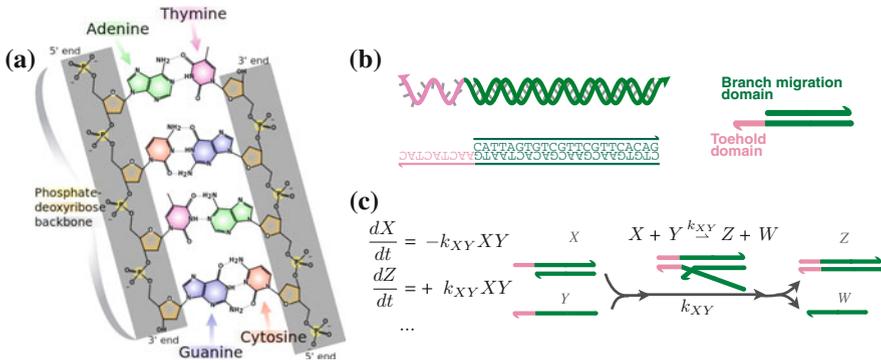


Fig. 12.1 Nucleic acids as programmable substrates. **a** Structure of a DNA double helical molecule (image adapted from www.wikipedia.org). **b** Partially double-stranded complex with a toehold domain (pink). **c** Schematic of a branch migration reaction and corresponding ordinary differential equations

satisfying the domain complementarity constraints. Nucleic acid strands can specifically bind to other molecules and materials (metals and non-metals) as well; however, as previously mentioned, in this case sequences cannot be designed rationally through software interfaces, but can be selected through directed evolution techniques such as SELEX [10, 54].

Proteins According to the central dogma of molecular biology, proteins are generated by complex cellular machinery through a process called RNA translation: a chain of amino-acids is produced starting from a chain of messenger RNA; RNA is in turn transcribed from a DNA template. Each amino acid corresponds to a specific triplet of bases, or codon. It is not our purpose to provide detailed information on protein composition and “encoding”, which is rather complex. Indeed, proteins are molecules whose chemical and spatial organization is extremely difficult to predict [33]. Using the efficient transcription and translation processes recapitulated *in vitro*, several synthetic biomolecular circuits have been demonstrated [20, 38, 50, 51]. Because this chapter focuses primarily on the use of nucleic acids to build circuits and systems, we briefly introduce polymerases and nucleases, classes of proteins that directly interact with nucleic acids, and are particularly simple to manufacture and utilize *in vitro*.

Polymerases are a class of proteins that catalyze the formation of nucleic acid polymers starting from nucleotides (dNTPs for DNA, rNTPs for RNA). Several types of DNA and RNA polymerases are commercially available; they differ in their molecular structures, specific functions (e.g., repair or replication), and substrates. Most reactions involving bacteriophage polymerases are relatively simple to recreate *in vitro*. In particular, the circuits described in this chapter will rely on transcription reactions carried out by the bacteriophage T7 RNA polymerase (T7 RNAP). We choose this polymerase because of its simple architecture (a single unit 99kDa protein), short promoter region (only 17 bases), and strong activity (it is typically used in

many in vitro applications). T7 RNAP is extremely promoter-specific and transcribes only DNA downstream of a T7 promoter where templated RNA synthesis proceeds from the 5' end to the 3' end. However, T7 RNAP also generates abortive transcripts (shorter than the full-length RNA signal) [21] and catalyzes aberrant untemplated reactions at a slow rate [57]. Some of the known side reactions can be taken into account at the design stage to minimize adverse effects on circuit operation [27].

Nucleases cleave the phosphodiester bonds between the nucleotide subunits of nucleic acids. Thus, in contrast with polymerases that catalyze formation of nucleic acid chains, nucleases cut or degrade the chains in various manners. In particular, we will consider RNA degrading enzymes such as RNase H and RNase R. *Escherichia coli* RNase H is used to selectively degrade RNA signals within RNA-DNA hybrid substrates [15, 27, 28]. However, as an endonuclease requiring stable RNA-DNA hybrid as substrates, RNase H cannot degrade RNA signals completely, leaving chunks of RNA pieces as incomplete degradation products [43]. Another RNase specific to single-stranded RNA, *E. coli* RNase R, can be utilized for circuit functions and/or to clean up short RNA fragments [55].

12.3 Reaction Primitives and Circuit Elements

Computations or operations with molecules can be performed by first identifying systems of chemical reactions exhibiting the desired dynamics, and then by implementing such reactions using appropriate substrates. Nucleic acids are the biomolecule of choice for constructing reaction networks, due to their programmable self-assembly properties; together with commercially available enzymes, they can be used as building blocks to generate or approximate biochemical reaction primitives. Without presumption of rigor we will describe some simple examples. We address the reader to [27, 53] for specific details.

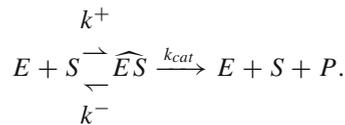
Nucleic acid hybridization kinetics naturally implement second-order chemical dynamics. Two complementary strands A and B bind at a rate k_{AB} that depends on their conformational states (secondary structures) and environmental factors such as temperature and salt conditions [16], and form a complex X with the following dynamics:

$$\frac{d}{dt}A = -k_{AB}A \cdot B = \frac{d}{dt}B, \quad \frac{d}{dt}X = k_{AB}A \cdot B.$$

Now, if perfectly double stranded, complex X is inert and may not be used for further reactions involving nucleic acids. However, the presence of an exposed overhang or toehold in one of the strands forming X (pink domain in Fig. 12.1b), allows other strands to initiate interaction with the toehold region of X and displace a single-stranded component through branch migration process [63], schematically represented in Fig. 12.1c. The resulting dynamics are systematically derived using mass action kinetics.

The principle of toehold-mediated branch migration is used in a variety of recent circuit design methods, where multi-input and multi-output DNA complexes or “gates” can implement arbitrary chemical reaction dynamics after proper scaling of reaction rates [41, 53].

To expand the range of dynamics implementable with biochemical substrates, we can combine hybridization and gating techniques with enzymatic reactions. As previously mentioned, several off-the-shelf proteins allow us to reproduce routinely a variety of biochemical processes such as transcription, replication, nicking and ligation. The classical Michaelis-Menten approach is generally a good approximation of the detailed enzymatic reaction dynamics, when the substrate concentration is large relative to the enzyme, and the binding and unbinding reactions of substrate and enzyme are fast. For example, consider a transcription process where the substrate S is a synthetic gene, the enzyme E is an RNA polymerase, and the product P is an RNA species:



Assuming that the dynamics of the species \widehat{ES} reach a steady state fast:

$$\begin{aligned} \frac{d}{dt} \widehat{ES} &= k^+ E \cdot S - k^- \widehat{ES} - k_{cat} \widehat{ES} = 0, \\ \widehat{ES} &= \frac{k^+ E \cdot S}{k^- + k_{cat}} = \frac{E \cdot S}{K_M}, \end{aligned}$$

where K_M is the Michaelis constant. In the last equation, we may further apply mass conservation:

$$\begin{aligned} E^{tot} &= E + \widehat{ES}, \\ S &\approx S^{tot}. \end{aligned}$$

Then, the Michaelis-Menten approximation yields:

$$\frac{d}{dt} P = k_{cat} \widehat{ES} = \frac{k_{cat}}{K_M} S \left(\frac{E^{tot}}{1 + \frac{S}{K_M}} \right).$$

Depending on the amount of substrate relative to K_M , the dynamics of product P range from linear (first order) to saturated (zeroth order) as a function of the substrate. If the substrate amount varies over time, then nonlinear dynamics can be generated.

Transcriptional circuits build on the examples we just overviewed. The elementary unit of synthetic transcriptional networks is a DNA switch, which serves the role of simplified gene in a genetic regulatory circuit. The basic requirements for a synthetic

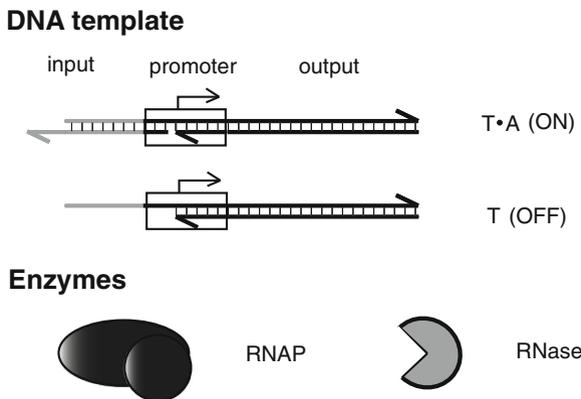


Fig. 12.2 Components of synthetic transcriptional circuits. The synthetic transcriptional circuits consist of switchable DNA templates and a few enzymes. The switchable DNA templates have a modular architecture consisting of an upstream input domain, a promoter for RNA polymerase (*boxed region with arrows to indicate direction of transcription*), and a downstream output domain. In its OFF state, the promoter domain of switch template is incomplete, i.e., partially single-stranded; in its ON state, an activator strand complements the missing promoter region such that the promoter domain is complete except for a nick. RNA polymerase drives the production of RNA signals using rNTP as fuels, while RNase removes transient RNA signals

DNA switch are to have a modular domain architecture and to have two distinct conformations with different transcriptional efficiency. Because the thermodynamics of DNA structures are well-known, it is straightforward to predict the conformational state from sequence information alone. To achieve modularity, the input domain of a DNA switch is placed upstream of the promoter region for RNAP; the output domain is placed downstream of the promoter region.

A possible mechanism of activation is the complementation of an incomplete promoter region within the OFF-state DNA switch. Figure 12.2 illustrates the design for DNA transcriptional switches and the mechanism for activation. Despite the nick, a break in the DNA backbone, within the promoter region, the ON-state DNA switch was found to transcribe well, approximately half as efficiently as the complete duplex promoter [27]. The ON-state switch can be converted back to an OFF-state switch using a perfect complement of the activator strand through a toehold-mediated strand displacement reaction [63]. Figure 12.3 illustrates the mechanism for inhibition. This set of binding reactions provides a means to choose the thresholds of the sigmoidal inhibition and activation functions, as described below.

The network computation is powered by rNTPs that drive the synthesis of RNA signals by RNAP, while RNases force transient RNA signals to decay. Since many of the RNA signals exert their effects through stoichiometric binding to target DNA strands, *E. coli* RNase H is used to selectively degrade RNA signals within RNA-DNA hybrid substrates [15, 27, 28]. Another RNase specific to single-stranded RNA,

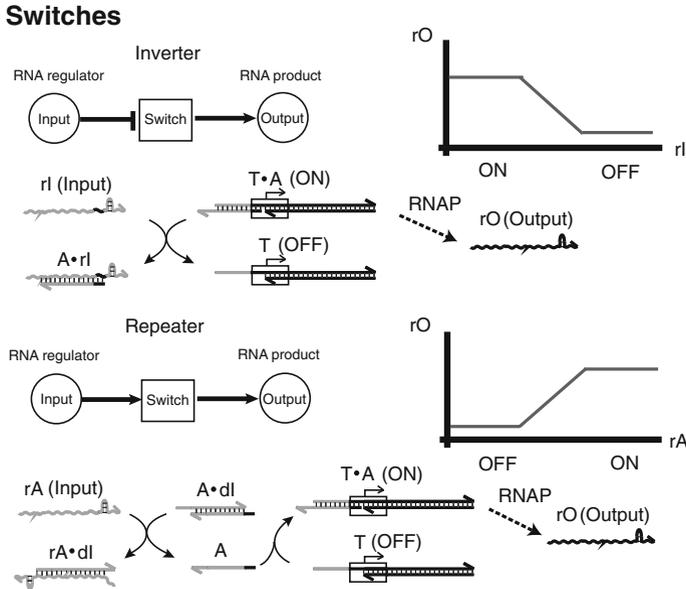


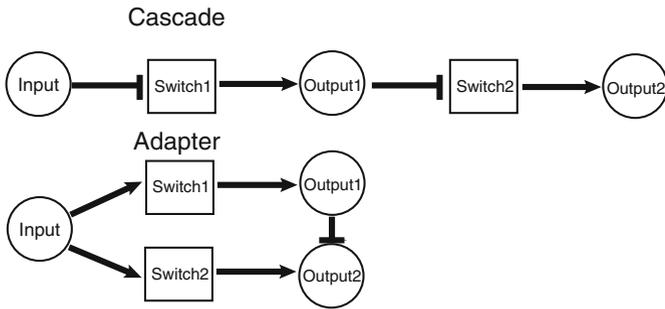
Fig. 12.3 Molecular implementation of inverter and repeater switches. For an inverter switch, an inhibiting RNA strand rI displaces DNA activator strand A from the promoter region of the active template $T \cdot A$, leaving the template in its transcriptionally inactive state T . The excess concentration of activator strand A can be used to set the inhibition threshold ($K_I = [A^{\text{tot}}] - \frac{1}{2}[T^{\text{tot}}]$). For a repeater switch, the activation of switch template is achieved indirectly: an activating RNA strand rA first displaces the DNA inhibitor dI from the inhibiting complex $A \cdot dI$; then, the released activator A binds to its target template T such that a transcriptionally active $T \cdot A$ complex forms. Here the excess of dI over A strands is used to set the activation threshold ($K_A = [dI^{\text{tot}}] - [A^{\text{tot}}] + \frac{1}{2}[T^{\text{tot}}]$)

E. coli RNase R, can be utilized for circuit functions and/or to clean up short RNA fragments [55].

12.4 Synthetic Biochemical Devices

The modular and flexible architecture of synthetic transcriptional network allows scaling up of network complexity. In principle, the synthetic transcriptional network is equivalent to neural networks capable of an associative memory task and a winner-take-all computation [25, 42]. In this section, we focus on the experimental implementation of elementary circuits and quantitative understanding thereof. Figure 12.4 illustrates examples of elementary circuits where the simple reaction primitives are implemented for functional circuits.

Feedforward Circuits



Feedback Circuits

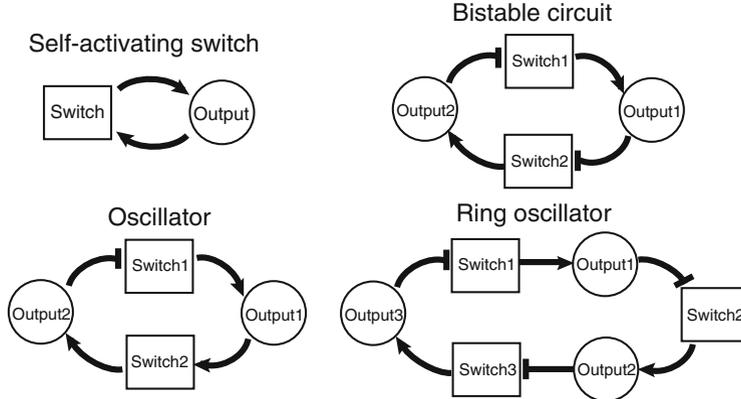


Fig. 12.4 Systematic construction of transcriptional circuits composed of modular switch motifs. Feedforward circuits include a cascade (as a part of oscillator circuit) and an adapter proposed in [26]. A bistable autoregulatory switch in [55] and a bistable circuit in [27] are examples of positive-feedback circuits; oscillators constructed in [28] are examples of negative-feedback circuits

12.4.1 Oscillator

A negative-feedback loop with delays has been utilized as the mechanism to generate temporal oscillations [17]. Using the inverter and repeater transcriptional switch motifs introduced earlier (Fig. 12.3), we constructed a two-switch negative feedback oscillator with the connectivity shown in Fig. 12.5a (cf. Fig. 12.4). A total of seven DNA strands are used, in addition to the two enzymes, RNAP and RNase H. RNA activator rA1 activates the production of RNA inhibitor rI2 by modulating switch Sw21, whereas RNA inhibitor rI2, in turn, inhibits the production of RNA activator rA1 by modulating switch Sw12.

First, we introduce a simplified 4-dimensional model to illustrate dynamics of a two-switch negative feedback oscillator. We take the two RNA species as dynamic signals and assume that their production rates are solely determined by the ON-

state switch concentrations and their degradation rates depend only on their own concentrations.

$$\frac{d[\text{rA1}]}{dt} = k_p \cdot [\text{T12A2}] - k_d \cdot [\text{rA1}], \quad (12.1)$$

$$\frac{d[\text{rI2}]}{dt} = k_p \cdot [\text{T21A1}] - k_d \cdot [\text{rI2}], \quad (12.2)$$

where k_p is the production rate from an ON-state switch and k_d is the degradation rate for RNA species. Experimentally, we observed that the steady-state switch response to RNA inputs can be reasonably approximated by Hill functions with thresholds set by the concentration of the strands competing for the activator. Thus, the two switch states are modeled to follow RNA input changes with relaxation time τ as follows:

$$\frac{d[\text{T12A2}]}{dt} = \frac{1}{\tau} \left([\text{T12}^{\text{tot}}] \frac{1}{1 + \left(\frac{[\text{rI2}]}{K_I}\right)^n} - [\text{T12A2}] \right), \quad (12.3)$$

$$\frac{d[\text{T21A1}]}{dt} = \frac{1}{\tau} \left([\text{T21}^{\text{tot}}] \left(1 - \frac{1}{1 + \left(\frac{[\text{rA1}]}{K_A}\right)^m} \right) - [\text{T21A1}] \right), \quad (12.4)$$

where K_A and K_I are the activation and inhibition thresholds for RNA activator rA1 and RNA inhibitor rI2. Reasonable approximations for the thresholds are given in terms of concentrations of molecular species as follows: $K_I = [\text{A2}^{\text{tot}}] - \frac{1}{2}[\text{T12}^{\text{tot}}]$ and $K_A = [\text{dI1}^{\text{tot}}] - [\text{A1}^{\text{tot}}] + \frac{1}{2}[\text{T21}^{\text{tot}}]$ (See [28] for more details on approximations for m , n , and τ). Simulated traces for the RNA species showed oscillation with reasonable parameter choices (Fig. 12.5b).

To explore the phase space of the simple model, we introduce non-dimensional variables as follows:

$$x = \frac{[\text{rA1}]}{K_A}, \quad y = \frac{[\text{rI2}]}{K_I}, \quad u = \frac{[\text{T12A2}]}{[\text{T12}^{\text{tot}}]}, \quad v = \frac{[\text{T21A1}]}{[\text{T21}^{\text{tot}}]}, \quad s = \frac{t}{\tau}. \quad (12.5)$$

Replacing the non-dimensional variables to Eqs. (12.1)–(12.4), we obtain:

$$\gamma \frac{dx}{ds} = \alpha \cdot u - x, \quad \gamma \frac{dy}{ds} = \beta \cdot v - y, \quad (12.6)$$

$$\frac{du}{ds} = \frac{1}{1 + y^n} - u, \quad \frac{dv}{ds} = 1 - \frac{1}{1 + x^m} - v, \quad (12.7)$$

where $\alpha = \frac{1}{K_A} \frac{k_p}{k_d} [\text{T12}^{\text{tot}}]$, $\beta = \frac{1}{K_I} \frac{k_p}{k_d} [\text{T21}^{\text{tot}}]$, and $\gamma = \frac{1}{k_d \cdot \tau}$.

The nullclines for x and y are calculated as follows: $u = \frac{1}{1 + y^n}$ and $v = 1 - \frac{1}{1 + x^m}$ at steady-state, and thus, the nullcline for x (i.e., $\frac{dx}{ds} = 0$) becomes $x = \alpha u = \frac{\alpha}{1 + y^n}$

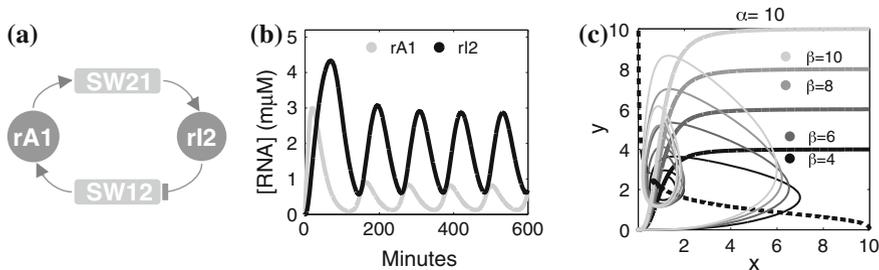


Fig. 12.5 **a** Scheme of the two-node synthetic oscillator. **b** Example trajectories for the RNA species in the four-dimensional model. The parameters are chosen as follows: $k_p = 0.05 /s$, $k_d = 0.001 /s$, $K_A = K_I = 0.5 \mu\text{M}$, $[\text{T21}^{\text{tot}}] = [\text{T12}^{\text{tot}}] = 100 \text{ nM}$, $m = n = 3$, $\tau = 1000 \text{ s}$ (this choice guarantees $\gamma = 1$). The initial conditions are the following: $[\text{rA1}](0) = 0 \mu\text{M}$, $[\text{T21A1}](0) = 0 \text{ nM}$, $[\text{rI2}](0) = 0 \mu\text{M}$, $[\text{T12A2}](0) = 100 \text{ nM}$. **c** Nullclines for x and y (non-dimensional variables corresponding to switch states) of the non-dimensional model where the parameter β was varied. The trajectories of x and y admit oscillation. The parameters are chosen as $\alpha = 10$, $m = n = 3$, $\gamma = 1$

and the nullcline for y (i.e., $\frac{dy}{ds} = 0$) becomes $y = \beta \left(1 - \frac{1}{1+x^m}\right)$. Because the nullcline for x as a function of y is monotonically decreasing while the nullcline for y as a function of x is monotonically increasing, the system admits a single fixed point. Nullclines are numerically computed in Fig. 12.5c for a wide range of β . The stability of the fixed point as calculated by the Jacobian matrix determines whether the system admits oscillations. Similarly, an analysis using control theoretic tools reveal that there are oscillating and non-oscillating domains determined by the gain of feedback system; this can be easily seen by linearizing the system around the unique equilibrium, and deriving the closed loop transfer function. Note that the system cannot oscillate when the time-scale separation for RNA dynamics and switch dynamics is valid (i.e., $\gamma \ll 1$ or $\gamma \gg 1$) because the reduced two-dimensional systems cannot admit oscillations. (See [28] for more details.) On the other hand, a two-dimensional system can admit oscillations if augmented by a positive feedback loop [2, 28].

While insightful for understanding basic requirements for oscillations and elucidating the space of dynamic potentials, the simple model neglects important details of the experimental system, such as asymmetries between switches and the complexities of enzyme and hybridization reactions. Therefore, a detailed model built from first principles was utilized to determine experimental conditions conducive to oscillations (See [28] for more details.). An important deviation from the simple model was that the Michaelis–Menten enzyme reaction was used instead of first-order approximations. The RNAP needs to support transcription from both switch templates and the RNase H works on both RNA signals within the RNA-DNA hybrid complexes. When the enzymes are saturated as the substrate concentrations increase, the enzyme reaction dynamics are no longer first-order but approach constant rates (zeroth order) determined by the maximum capacity of enzymes. The coupling of

reactions through common enzymes or resources have significant ramifications in the dynamics of systems, leading to non-intuitive results at times [44, 62]. On the other hand, the saturation of degradation enzymes may help increase the domain of oscillations [61]. Using a straightforward random sampling technique inspired by stochastic high-dimensional sensitivity analysis [12], the initial experimental conditions were determined which showed a damped oscillation; further optimization lead to sustained oscillations.

However, even the detailed model failed to capture certain aspects of experimental results, i.e., the continued accumulation of RNA signals after each cycle. When an initial hypothesis of reduced enzyme activity was implemented in the augmented model, the simulation result was not able to generate such phenomena. An alternative hypothesis that the incomplete degradation product interferes with the intended signal propagation among switches was implemented: this augmented model qualitatively captured the experimental observation. It is interesting to note that the synthetic oscillator was able to mount an increased signal response to compensate for the reduced signal transmission efficiency due to the presence of interfering species, showing robustness to such disturbance. The waste management is a different type of crosstalk not often discussed in synthetic biology. Although it is plausible that biological systems are equipped to handle waste management issues efficiently, introduction of novel molecular species to biological systems may cause unexpected difficulties. As an analogous example, the bacteria *E. coli* cannot survive without oligoRNase that specifically degrades short single-stranded RNA species [13], indicating the importance of waste management for biological circuits.

12.4.2 Adapter

One common feature found in many sensory systems of living organisms is the exact adaptation in which the output upon change of input to a new constant level gradually returns to a steady level independent of the input [52, 58]. A classic example is the bacterial chemotaxis in which bacteria responds to the gradient rather than the absolute level of nutrients. A type-1 incoherent feedforward loop (IFFL) is a regulatory pattern in which an input u serving as an activator controls a target gene y and also activates a repressor of that target gene, x [36] (cf. Fig. 12.4).

A generalized model of ‘sniffer’ [59] can take the following form:

$$\dot{x} = \alpha_1 \cdot u - \beta_1 \cdot x - k \cdot x \cdot y, \quad (12.8)$$

$$\dot{y} = \alpha_2 \cdot u - \beta_2 \cdot y - k \cdot x \cdot y. \quad (12.9)$$

Here, the α_1 and α_2 terms reflect a generic asymmetry in the effectiveness of u as an activator for x and y . The β_1 and β_2 are generic degradation terms for x and y . One notable characteristic is the existence of the $k \cdot x \cdot y$ term for dynamics of both x and y ; this type of accelerated degradation term can be chosen for the case in which x and

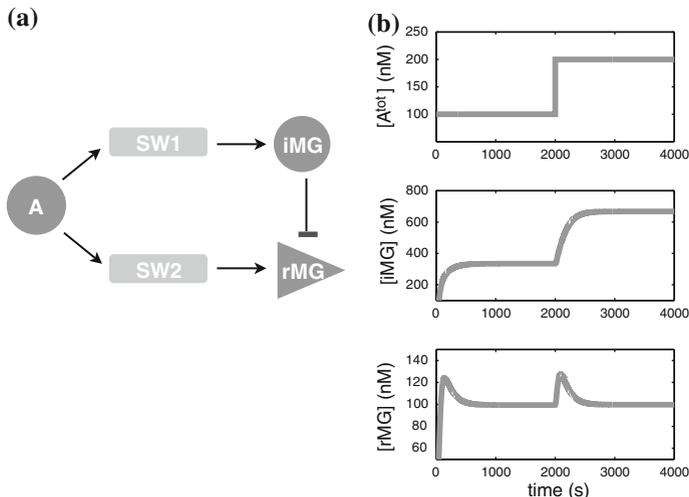


Fig. 12.6 **a** Scheme of the synthetic adapter. **b** Example trajectories for the input and RNA species. The parameters are chosen as follows: $[T1^{tot}] = 200$ nM, $[T2^{tot}] = 100$ nM, $k_+ = 10^5$ /M/s, $k_p = 0.1$ /s, $\beta_1 = 0.01$ /s, $\beta_2 = 0.0001$ /s, and $k = 10^5$ /M/s

y stoichiometrically react to annihilate each other as in the small RNA regulatory circuits in biological systems [37].

Using the transcriptional switch as the regulatory motif and the aptamer for chromophore malachite green (MG) as the output signal, we can construct an IFFL circuit and monitor its output state real-time (Fig. 12.6a). It is straightforward to design two transcriptional switches that share common input domains such that they are activated by the same DNA activator A and transcribe different outputs: MG aptamer (“rMG”) and its inhibitor (“iMG”). The MG aptamer, rMG, consists of a short RNA sequence whose central loop region serves as the binding pocket of MG [18]; upon binding to MG aptamer, rMG, the fluorescence signal of MG greatly increases. The inhibitor for rMG consists of domains that complement part of rMG such that it opens up the central loop region of rMG upon binding. Therefore, rMG bound to iMG loses the binding pocket for MG and the fluorescence signal decreases.

Table 12.1 shows the list of hybridization and branch migration reactions and enzyme reactions approximated as first-order reactions. The dynamics of this in vitro circuit (Fig. 12.6a) can be described by the following four ordinary differential equations:

$$[T1 \cdot A] = k_+[T1][A], \tag{12.10}$$

$$[T2 \cdot A] = k_+[T2][A], \tag{12.11}$$

$$[i\dot{M}G] = k_p[T1 \cdot A] - \beta_1[iMG] - k[iMG][rMG], \tag{12.12}$$

$$[r\dot{M}G] = k_p[T2 \cdot A] - \beta_2[rMG] - k[iMG][rMG]. \tag{12.13}$$

Table 12.1 Reaction pathways for the transcriptional adapter circuit

Reaction type	Reaction		
Activation	$T1 + A$	$\xrightarrow{k_+}$	$T1 \cdot A$
Activation	$T2 + A$	$\xrightarrow{k_+}$	$T2 \cdot A$
Production	$T1 \cdot A$	$\xrightarrow{k_p}$	$T1 \cdot A + iMG$
Production	$T2 \cdot A$	$\xrightarrow{k_p}$	$T2 \cdot A + rMG$
Degradation	iMG	$\xrightarrow{\beta_1}$	ϕ
Degradation	rMG	$\xrightarrow{\beta_2}$	ϕ
Inhibition	$rMG + iMG$	\xrightarrow{k}	$rMG \cdot iMG$

The system preserves the conservation relations, $[Ti^{tot}] = [Ti] + [Ti \cdot A]$, and similarly for $[A^{tot}]$. Because the binding of A to either of switches is practically irreversible, we expect that $[A] \simeq 0$ and $[T1 \cdot A] + [T2 \cdot A] \simeq [A^{tot}]$ at steady-state given $[A^{tot}] < [T1^{tot}] + [T2^{tot}]$. Let $u = [A^{tot}]$, $x = [iMG]$, and $y = [rMG]$. Then, we obtain the same set of ODEs as in Eqs. (12.8) and (12.9) for (12.12) and (12.13) as follows:

$$\dot{x} = \alpha_1 \cdot u - \beta_1 \cdot x - k \cdot x \cdot y, \quad (12.14)$$

$$\dot{y} = \alpha_2 \cdot u - \beta_2 \cdot y - k \cdot x \cdot y, \quad (12.15)$$

where α_1 and α_2 are functions of k_p (RNAP concentration), $[T1^{tot}]$, and $[T2^{tot}]$, β_1 and β_2 are functions of RNase R concentration and the secondary structure of RNA molecules, and k is a function of the length of exposed toehold of rMG. Therefore, all of the rate constants are amenable to tuning (See [26] for details.).

If β_2 is zero with other constants being positive, the steady-state of x is proportional to the input u and the steady-state of y is independent of u , satisfying the requirement of exact adaptation (if $\alpha_1 > \alpha_2$). The steady-state solutions are as follows:

$$x_s = \frac{(\alpha_1 - \alpha_2)u}{\beta_1}, \quad y_s = \frac{\alpha_2 \cdot \beta_1}{k(\alpha_1 - \alpha_2)}. \quad (12.16)$$

Through mathematical analysis, we can demonstrate that the transcriptional circuit designed to have an IFFL motif can provide exact adaptation behavior as long as β_2 is orders of magnitude smaller than β_1 (Fig. 12.6b). Differential regulation of the two degradation rate constants, β_1 and β_2 , is possible by utilizing the substrate recognition property of RNase R: RNase R can tightly bind to a single-stranded RNA that has no secondary structure with a K_d of 2 nM, while it binds weakly to a duplex RNA having less than 4nt 3' overhang with a K_d greater than 5 μ M [60]. Therefore, iMG designed to have no significant secondary structure degrades fast, while rMG designed to have a strong secondary structure with no 3' overhang degrades slow, satisfying the requirement $\beta_1 \gg \beta_2$. Employing different secondary structures for RNA substrates at their 3' ends offers an opportunity to tune the degradation rates by RNase R.

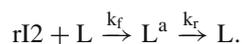
The preliminary experimental data support the simplified mathematical model used to describe the dynamics of transcriptional adapter circuit. It is interesting to note that the simplified model could capture the experimental data effectively without resorting to more detailed models which was necessary to capture the behavior of oscillator circuit. One of the important distinctions between the adapter circuit and the oscillator circuit is the mechanism of degradation. The degradation mechanism used for the oscillator circuit relies on RNase H, an endonuclease that cleaves substrates in the middle. While mostly processive in its mode of action, RNase H cannot process 5–7 bases at the 5' end of RNA within the RNA-DNA substrate, leaving a relatively big chunk of RNA that can potentially interfere with the intended hybridization reactions. On the other hand, the degradation for the adapter circuit relies solely on RNase R, an exonuclease that cleaves RNA substrates from its 3' end. RNase R is known to process highly structured RNA down to di- and trinucleotides [5]: the resulting dimer or trimer would be too short to interfere with the designed hybridization reactions. Therefore, the successful application of simple model for the adapter circuit can in part be ascribed to the efficient waste management through RNase R.

12.4.3 Signal Transmission

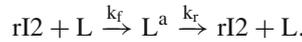
The devices we just described have an important role in the development of integrated biochemical systems for molecular computation, robotics, and nanomedicine. For example, synthetic circuitry capable of site recognition, timed operations, and sampling of environment could lead to smart drug delivery systems: oscillators, memory elements and nucleic acid sensors (e.g., aptamers) would have to be combined in larger integrated circuits.

Thus, the usefulness of all these modules is based on their ability to be interconnected easily and reliably, transmitting and receiving molecular signals like electrical circuits. However, a major challenge highlighted in the recent literature is the intrinsic inability of many biological pathways to perform modular signal transmission [6, 48]. This challenge is best exemplified through the problem of propagating oscillations generated by our molecular clock in Eqs. (12.1)–(12.4) to a downstream 'load', L. Without loss of generality, we will assume that we can couple rI2 to L (the same analysis can be easily carried out for the switch elements SW21, SW12 and rA1). A schematic representation of this model problem is shown in Fig. 12.7a. We distinguish two cases:

1. rI2 is consumed by the load. We call this 'consumptive' coupling. The corresponding chemical reactions are:



2. rI2 is not consumed by the load. We call this 'non-consumptive' coupling. The corresponding chemical reactions are:



These additional reactions introduce some modifications in the oscillator Eqs. (12.1)–(12.4). In particular, the dynamics of $rI2$ are perturbed, and we must now augment our model to include the dynamics of the load:

$$\frac{d[rI2]}{dt} = k_p \cdot [SW21] - k_d \cdot [rI2] + \underbrace{k_r \cdot [L^a]}_{\text{non-consumptive}} \underbrace{-k_f \cdot [L][rI2]}_{\text{consumptive}}, \quad (12.17)$$

$$\frac{d[L^a]}{dt} = -k_r \cdot [L^a] + k_f \cdot [L][rI2]. \quad (12.18)$$

For illustrative purposes, we choose $k_r = 0.003 /s$ and $k_f = 1.9 \cdot 10^3 /M/s$ for the numerical simulations (other parameters are unchanged from Fig. 12.5). Here we do not derive the full non-dimensional model for the oscillator with load; we refer the reader to the supplement of [15] for further details. For simplicity we focus on the consumptive coupling example. The non-consumptive case has been considered in [6], where the authors derive an analytical expression for the retroactivity induced by the load (Note that the derivation in [6] relied on a time scale separation argument which requires arbitrarily fast rates for k_r and k_f ; this may not be applicable depending on the biomolecular system under consideration.). Proceeding with our example, we assume that the load dynamics are fast relative to the source; thus we have a stationary expression:

$$[\widehat{L^a}](t) = [L^{tot}] \left(1 - \frac{k_r}{k_r + k_f[\widehat{rI2}](t)} \right), \quad (12.19)$$

where $[\widehat{rI2}](t)$ is an approximated solution, due to our stationary assumption. If we plug the load stationary solution into the consumptive dynamics of $[rI2]$ (approximated), we find:

$$\frac{d[\widehat{rI2}]}{dt} = k_p \cdot [SW21] - k_d \cdot [\widehat{rI2}] - k_f \cdot [\widehat{rI2}][L^{tot}] \left(\frac{k_r}{k_r + k_f[\widehat{rI2}]} \right), \quad (12.20)$$

where the box highlights the quasi-steady state approximated perturbation term. Loosely speaking, the total amount of load linearly modulates an additional, bounded degradation term (In fact, the perturbation term converges to $k_r \cdot [L^{tot}]$ for high values of $[rI2]$.). The differential equations above were solved for varying amounts of $[L^{tot}]$ numerically. As shown in Fig. 12.7b and c, a large load significantly deteriorates the oscillation.

We can make some analytical considerations on the load dynamics. Assume that $[rI2](t)$ is a sinusoidal signal: then, using the so called ‘contractivity’ theory [47] we can elegantly show that also $[L^a](t)$ is a periodic signal. In short, it is sufficient to

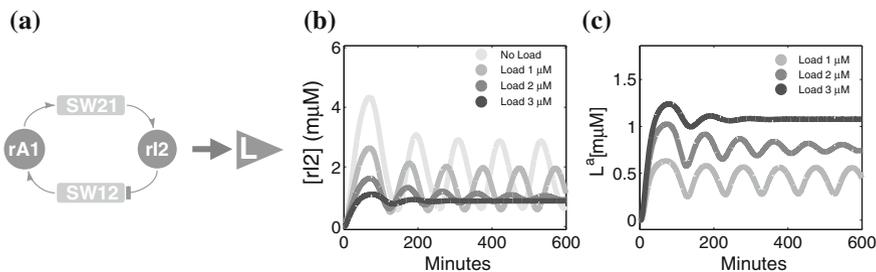


Fig. 12.7 **a** Scheme of the two-node synthetic oscillator connected to a load L . **b** Trajectories of $[rI2]$ for variable, increasing load, highlighting the deterioration of the oscillation amplitude and frequency. **c** Load traces:

verify that the linearization of the differential equation for the load is bounded by a negative constant, satisfying the definition of contractivity. Since our system evolves on a compact and convex set, such property is global inside such set, and for any initial condition the system will converge to the periodic solution. In particular, we can show that the dynamics of $[L^a]$ converge exponentially to a stationary periodic solution, and the convergence speed is larger than k_r .

While the stationary behavior of the load is periodic, the efficiency of the signal transmission is compromised. This loss of efficiency can be evaluated analytically using the static load approximation of Eq. (12.19). In particular, we can compute the amplitude of the load as a function of the oscillator amplitude. We will assume that $[rI2](t) = A_0 + A_1 \sin \omega t$, where $A_0, A_1 > 0$ and $A_0 > A_1$. Define $\kappa = k_r/k_f$. The amplitude of the load oscillations is then given by:

$$A_L = \frac{[L^{\text{tot}}]}{2} \left(\frac{\kappa}{\kappa + (A_0 - A_1)} - \frac{\kappa}{\kappa + (A_0 + A_1)} \right).$$

By taking the derivative of A_L with respect to κ , and setting the derivative to zero, we can calculate the value of κ that maximizes A_L :

$$\kappa_{\text{max}} = \sqrt{A_0^2 - A_1^2}.$$

For instance, take $A_0 \approx 1.9 \mu\text{M}$ and $A_1 \approx 1.12 \mu\text{M}$, the mean and amplitude of the nominal $[rI2]$ oscillations. Then, if we assume $k_r = 0.003/\text{s}$, the value of k_f that maximizes the load amplitude is $k_f \approx 1.9 \cdot 10^3/\text{M}\cdot\text{s}$.

Now let us examine the back-action of the load on the oscillator components, highlighted in the box of Eq. (12.20). How can the perturbation on the oscillator be reduced?

We start from Eqs. (12.19) and (12.20), where the quasi-steady state approximation for the dynamics of $[L^a](t)$ is substituted in the dynamics of $[rI2]$. Equation (12.20) clearly shows that when it is not practical to modify the binding rates that introduce the coupling, the only way to reduce perturbation is to use a

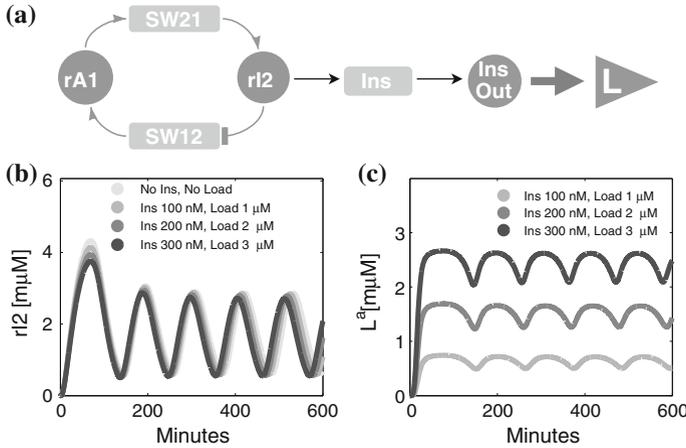


Fig. 12.8 **a** Scheme of the two-node synthetic oscillator with an insulated load. **b** Trajectories of [rI2] for variable insulator and load: insulation reduces retroactivity effect. **c** Load traces: insulation guarantees signal transmission

minimal amount of load. We can overcome this restriction by coupling the oscillatory signal to a small amount of another molecular device, whose output is capable of amplifying the oscillator signal and driving large amounts of load. We will call this device an insulator, inspired by the analysis proposed in [6, 14]. A schematic representation of this idea is shown in Fig. 12.8a.

An insulating device can be implemented easily as a small amount of a third switch, Ins, which is directly coupled to the oscillator. The RNA output from the insulating switch, InsOut, is used to drive the load. The differential equations describing the dynamics of rI2 and the insulated load are:

$$\begin{aligned}\frac{d[rI2]}{dt} &= k_p \cdot [SW21] - k_d \cdot [rI2] - k_f \cdot [Ins][rI2], \\ \frac{d[Ins]}{dt} &= k_r \cdot [Ins^a] - k_f \cdot [Ins][rI2], \\ \frac{d[InsOut]}{dt} &= k_p^i \cdot [Ins^a] - k_d^i \cdot [InsOut] - k_f^i \cdot [InsOut][L], \\ \frac{d[L]}{dt} &= k_r^i \cdot ([L^{tot}] - [L]) - k_f^i \cdot [InsOut][L].\end{aligned}$$

(The remaining differential equations for rA1, SW21, and SW12, are unchanged.) The additional parameters for insulator are chosen for numerical simulations as follows: $k_p^i = 0.15$ /s, $k_d^i = 0.003$ /s, $k_r^i = 0.003$ /s and $k_f^i = 3 \cdot 10^3$ /M/s (All the other parameters are the same as in Fig. 12.7.). The simulation results showed insulation, i.e., little retroactivity, from the load as well as proper signal transmission (Fig. 12.8b, c).

Now the direct load on the oscillator is given by the insulator switch, coupled consumptively to the oscillator. Following our previous reasonings, we can show that the active state of the switch can be approximated as:

$$[\widehat{\text{Ins}}^a] \approx [\text{Ins}^{\text{tot}}] \left(1 - \frac{k_r}{k_r + k_f[rI2](t)} \right).$$

Using this approximation, the perturbation on the dynamics of $rI2$ can be reduced arbitrarily by minimizing the amount of $[\text{Ins}^{\text{tot}}]$.

We can quantify the efficiency of insulated signal transmission starting from the expressions derived previously. As shown in [15], if the production and degradation rates of the insulator output are suitably chosen, we can ensure not only propagation but also amplification of the oscillatory signal.

This particular example, where we have described the challenges arising in propagating a molecular oscillatory signal to a downstream load, clearly shows that to scale up synthetic biological circuits we need systematic methods to ensure modular signal transmissions. Insulation devices, which decouple signal source and load while introducing amplification, are a possible solution to interconnect modules characterized in isolation, ensuring proper signal transmission and reduced back-action.

12.5 The Future of Biosynthetic Circuits

In this chapter we have introduced a variety of molecular circuit elements and devices that can be built using nucleic acids and proteins. The programmability of nucleic acids, which follows from their predictable, highly specific Watson-Crick base pairing interactions, allows us to design molecular interactions systematically with unprecedented easiness much as writing codes in programming languages. Although we merely scratched the surface of the possible range of programming methods for biomolecular reactions, we have described how general reaction dynamics can be synthesized using DNA, RNA, and proteins. These reactions underlie the basis of a variety of circuits with predictable input/output functionalities: in particular, we described synthetic transcriptional circuits and their application in constructing oscillators, bistable circuits, adapters, and a range of other architectures. Finally, we discussed design requirements to ensure scalability of molecular circuits.

Throughout this chapter, we have briefly mentioned some of the challenges affecting synthetic biochemical systems. It is worth recalling that enzymatic processes introduce phenomena such as abortive transcription and incomplete degradation pathways, which create waste management issues in existing circuits. Advances in rational enzyme design, a rapidly advancing research area, will possibly allow custom design options to improve performance. Undesired interactions among complexes (including unwanted binding and output leak in DNA gates) is another challenge in designing nucleic acid circuits: while circuits comprising over 100 species have been successfully demonstrated [41], systematic design methods to avoid spu-

rious interactions are still in demand. Universal toeholds and stringent stoichiometry requirements (enforced through gel purifications) have been proposed as possible solutions to this issue.

Nucleic acid-based circuits have thus far been used to demonstrate various computational programs, including logic circuits and neural networks [42, 49], and layered dynamic circuits [15]; once suitably adapted to operate in complex cellular environment, strand displacement cascades promise to yield a versatile synthetic platform for control of gene expression [3].

An emerging application for programmable biosynthetic circuits is the generation of advanced materials. We previously mentioned a range of techniques to build self-assembled nanostructures using DNA, including two and three dimensional origami, nanotubes, and molecular machines. Dynamic nucleic acid circuits could be used to power and direct assembly of nanostructures, creating one of the first generations of entirely synthetic, dynamic, programmable biomaterials [29].

Finally, the circuits we described are a promising toolkit to develop artificial cells [38]. Cells can be viewed as complex ensembles of molecular circuits, shaped by billions of years of evolutionary tinkering. Understanding the extent to which human-developed design principles can be implemented for synthetic biological compartments that recapitulate the main cellular functions (such as information processing, growth, replication, and motility) is of fundamental scientific interest. If successful, artificial cells could be useful as drug delivery, probing and sampling devices for nanomedicine.

References

1. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2:2006.0028
2. Atkinson MR, Savageau M, Myers J, Ninfa A (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* 113:597–607
3. Benenson Y, Shapiro E (2004) Molecular computing machines. Dekker Encyclopedia of Nanoscience and Nanotechnology. Springer, New York
4. Benner SA, Sismour AM (2005) Synthetic biology. *Nature Rev Genet* 6(7):533–543
5. Cheng ZF, Deutscher MP (2002) Purification and characterization of the *Escherichia coli* exoribonuclease RNase R: Comparison with RNase II. *J Biol Chem* 277:21,624–21,629
6. Del Vecchio D, Ninfa A, Sontag E (2008) Modular cell biology: retroactivity and insulation. *Mol Syst Biol* 4:161
7. Douglas SM, Bachelet I, Church GM (2012) A logic-gated nanorobot for targeted transport of molecular payloads. *Science* 335(6070):831–834
8. Douglas SM, Dietz H, Liedl T, Hgberg B, Graf F, Shih WM (2009) Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* 459:414–418
9. Dyson F (2007) Our biotech future. *The New York Review of Books* 54. <http://www.nybooks.com/articles/20370>. Accessed 19 July 2007
10. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346:818–822
11. Endy D (2005) Foundations for engineering biology. *Nature* 438(7067):449–453
12. Feng X, Hooshangi S, Chen D, Li G, Weiss R, Rabitz H (2004) Optimizing genetic circuits by global sensitivity analysis. *Biophys J* 87:2195–2202

13. Fiedler TJ, Vincent HA, Zuo Y, Gavrialov O, Malhotra A (2004) Purification and crystallization of *Escherichia coli* oligoribonuclease. *Acta Crystallogr D Biol Crystallogr* 60:736–739
14. Franco E, Del Vecchio D, Murray RM (2009) Design of insulating devices for in vitro synthetic circuits. In: Proceedings of the IEEE conference on decision and control, pp 4584–4589
15. Franco E, Friedrichs E, Kim J, Jungmann R, Murray R, Winfree E, Simmel FC (2011) Timing molecular motion and production with a synthetic transcriptional clock. *Proc Natl Acad Sci USA* 108(40):E784–E793
16. Gao Y, Wolf LK, Georgiadis RM (2006) Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res* 34(11):3370–3377
17. Goodwin BC (1965) Oscillatory behavior in enzymatic control processes. *Adv Enzyme Regul* 3:425–436
18. Grate D, Wilson C (1999) Laser-mediated, site-specific inactivation of RNA transcripts. *Proc Natl Acad Sci USA* 96:6131–6136
19. Hockenberry AJ, Jewett MC (2012) Synthetic in vitro circuits. *Curr Opin Chem Biol* 16:253–259
20. Hodgman CE, Jewett MC (2012) Cell-free synthetic biology: thinking outside the cell. *Metab Eng* 14(3):261–269
21. Jiang M, Rong M, Martin C, McAllister WT (2001) Interrupting the template strand of the T7 promoter facilitates translocation of the DNA during initiation reducing transcript slippage and the release of abortive products. *J Mol Biol* 310:509–522
22. Karkare S, Bhatnagar D (2006) Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA and morpholino. *Appl Microbiol Biotechnol* 71:575–586
23. Ke Y, Ong LL, Shih WM, Yin P (2012) Three-dimensional structures self-assembled from DNA bricks. *Science* 338(6111):1177–1183
24. Khalil AS, Collins JJ (2010) Synthetic biology: applications come of age. *Nat Rev Genet* 11(5):367–379
25. Kim J, Hopfield JJ, Winfree E (2004) Neural network computation by in vitro transcriptional circuits. *Adv Neural Inf Proc Syst* 17:681–688
26. Kim J, Murray R (2011) Analysis and design of a synthetic transcriptional network for exact adaptation. In: IEEE biomedical circuits and systems conference (BioCAS), pp 345–348
27. Kim J, White KS, Winfree E (2006) Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Mol Syst Biol* 2:68
28. Kim J, Winfree E (2011) Synthetic in vitro transcriptional oscillators. *Mol Syst Biol* 7:465
29. Krishnan Y, Simmel FC (2011) Nucleic acid based molecular devices. *Angew Chem Int Ed* 50(14):3124–3156
30. Kuzyk A, Schreiber R, Fan Z, Pardatscher G, Roller EM, Hogele A, Simmel FC, Govorov AO, Liedl T (2012) DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* 483:311–314
31. Lakin MR, Parker D, Cardelli L, Kwiatkowska M, Phillips A (2012) Design and analysis of DNA strand displacement devices using probabilistic model checking. *J R Soc Interface* 9(72):1470–1485
32. Lesnik EA, Freier SM (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry* 34(34):10,807–10,815
33. Mandell D, Kortemme T (2009) Computer-aided design of functional protein interactions. *Nat Chem Biol* 5(11):797–807
34. Markham NR, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* 33:577–581
35. Maune HT, Han SP, Barish RD, Bockrath M, III WAG, Rothmund PWK, Winfree E (2010) Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. *Nat Nanotechnol* 5:61–66
36. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. *Science* 298(5594):824–827

37. Mitarai N, Benjamin JAM, Krishna S, Semsey S, Csiszovszki Z, Masse E, Sneppen K (2009) Dynamic features of gene expression control by small regulatory RNAs. *Proc Natl Acad Sci USA* 106:10,655–10,659
38. Noireaux V, Maeda YT, Libchaber A (2011) Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc Natl Acad Sci USA* 108:3473–3480
39. Phillips A, Cardelli L (2009) A programming language for composable DNA circuits. *J R Soc Interface* 6:S419–S436
40. Pinheiro AV, Han D, Shih WM, Yan H (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat Nanotechnol* 6:763–772
41. Qian L, Winfree E (2011) Scaling up digital circuit computation with DNA strand displacement cascades. *Science* 332:1196–1201
42. Qian L, Winfree E, Bruck J (2011) Neural network computation with DNA strand displacement cascades. *Nature* 475:368–372
43. Rizzo J, Gifford LK, Zhang X, Gewirtz AM, Lu P (2002) Chimeric RNA–DNA molecular beacon assay for ribonuclease H activity. *Mol Cell Probes* 16:277–283
44. Rondelez Y (2012) Competition for catalytic resources alters biological network dynamics. *Phys Rev Lett* 108(1):018,102
45. Rothmund PWK (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440:297–302
46. Rothmund PWK, Ekani-Nkodo A, Papadakis N, Kumar A, Fygenson DK, Winfree E (2004) Design and characterization of programmable DNA nanotubes. *J Am Chem Soc* 126(50):16,344–16,352
47. Russo G, di Bernardo M, Sontag ED (2010) Global entrainment of transcriptional systems to periodic inputs. *PLoS Comput Biol* 6(4), e1000, 739
48. Saez-Rodriguez J, Kremling A, Gilles E (2005) Dissecting the puzzle of life: modularization of signal transduction networks. *Comput Chem Eng* 29(3):619–629
49. Seelig G, Soloveichik D, Zhang DY, Winfree E (2006) Enzyme-free nucleic acid logic circuits. *Science* 314(5805):1585–1588
50. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* 19(8):751–755
51. Shin J, Noireaux V (2012) An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth Biol* 1(1):29–41
52. Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M (1997) Adaptation of retinal processing to image contrast and spatial scale. *Nature* 386:69–73
53. Soloveichik D, Seelig G, Winfree E (2010) DNA as a universal substrate for chemical kinetics. *Proc Natl Acad Sci USA* 107:5393–5398
54. 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
55. Subsoontorn P, Kim J, Winfree E (2012) Ensemble bayesian analysis of bistability in a synthetic transcriptional switch. *ACS Synth Biol* 1:299–316
56. Tombelli S, Mascini M (2010) Aptamers biosensors for pharmaceutical compounds. *Comb Chem High Throughput Screening* 13(7):641–649
57. Triana-Alonso FJ, Dabrowski M, Wadzack J, Nierhaus KH (1995) Self-coded 3'-extension of run-off transcripts produces aberrant products during in vitro transcription with T7 RNA polymerase. *J Biol Chem* 270:6298–6307
58. Tu Y, Shimizu TS, Berg HC (2008) Modeling the chemotactic response of Escherichia coli to time-varying stimuli. *Proc Natl Acad Sci USA* 105:14,855–14,860
59. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* 15:221–231
60. Vincent HA, Deutscher MP (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. *J Biol Chem* 281:29,769–29,775
61. Wong WW, Tsai TY, Liao JC (2007) Single-cell zeroth-order protein degradation enhances the robustness of synthetic oscillator. *Mol Syst Biol* 3:130

62. Yeung E, Kim J, Yuan Y, Gonçalves J, Murray R (2012) Quantifying crosstalk in biochemical systems. In: Proceedings of the IEEE conference on decision and control, pp 5528–5535
63. Yurke B, Mills AP (2003) Using DNA to power nanostructures. *Genet Program Evolvable Mach* 4:111–122
64. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA (2011) NUPACK: analysis of nucleic acid systems. *J Comput Chem* 32:170–173

Index

A

Akaike's information criterion, 9, 24
Analysis of variance (ANOVA), 8, 12, 19
Automatic modeling, 202

B

Bifurcation, 154
Bifurcation point, 247, 248
Bistability, 142–145, 148, 149, 154, 156–159, 161–163
Bistable, 141–150, 155, 157, 160–164, 220–223, 225–238, 240–242, 248, 273, 274, 281, 291

C

Cellular signaling, 18
Chain termination, 32
Chemical Reaction Network Theory (CRNT), 145
Circadian entrainment, 59–61, 76, 80
Coding region, 44, 258, 259
Comparative genome hybridization (CGH), 32
Conjunctive normal form (CNF), 260

D

Designer sequences, 202
Digital, 257, 260, 262, 265, 267, 269, 270
Dissipativity, 117–120, 122
DNA, 32, 202, 204, 205, 208, 210
Domain of attraction (DA), 149
Double strand meta-DNA (dsmDNA), 42, 172, 174–180, 182, 183, 189, 190, 275, 276

E

Entrainment, 59–61, 65–67, 76–80
Enzyme, 18, 25, 40, 44, 50, 100, 143, 144, 147, 149, 171, 173, 176, 182, 197, 206, 207, 209, 210, 274, 277–279, 281, 283–285, 291
Equilibrium Point, 142, 143, 149–151, 154–157, 159–161, 163, 248
Escherichia coli (E. coli), 32, 34, 36, 38, 44–46, 129, 130, 142, 143, 145, 212, 213, 215, 277, 279, 280, 284
Eukaryotic, 32, 60, 85, 142, 143

G

Global sensitivity, 155

H

High osmolarity glycerol (HOG), 85, 87–90, 92–94, 98–100, 105, 107–112
Hill kinetics, 206
Homeostasis, 86, 146
Hybridization, 32

I

Incremental dissipativity, 119, 120
Incremental signal, 118, 128
Incremental supply rate, 119, 120
Incrementally output feedback passive (iOFP), 118, 120, 122, 124, 125, 127, 128, 137, 138, 297
Incrementally output feedback passive (iOFP), 129
Incrementally passive, 118–120, 132
Input-output, 60, 61, 64, 80
Integral feedback, 85, 101, 102, 106, 114

Inverse bifurcation analysis, 248
 Inverse bifurcation synthesis, 246
 Isothermal, 172, 173, 176, 189, 191

K

Kinetic law, 204, 206, 207

L

Linear temporal logic (LTL), 246
 Linearization, 247, 248, 253
 Local sensitivity, 153
 Lung cancer, 3–5, 7, 8, 11, 12, 14, 17–22
 Lung cancer investigation, 5

M

Mass action, 206
 Meta-DNA (mDNA), 42, 172, 176–180, 182, 183, 185, 189–193, 196–199
 Microarray, 32
 Microarray data, 5, 7, 8, 21
 Model definition language (MDL), 258, 267
 Molecular switches, 143, 164
 Monostable, 220–223, 225–231, 233–238, 240–242
 Multistable dynamics, 143

N

Network biomarker, 3, 5, 7–12, 14, 16, 18–22
 Next generation sequencing, 32
 Nucleic acid, 274–277, 287, 291, 292
 Nucleic acid-based circuits, 292
 Nucleotide, 275–277

O

Osmoadaptation, 85
 Osmoregulation, 85–87, 93, 99, 102, 106, 113, 114
 Osmosis, 84

P

Parameters, 246, 248, 249, 252, 253
 Passive, 118, 119, 132
 Passivity, 118–120, 124, 128, 134, 136–138
 Pathway, 144, 157
 Pattern, 203, 205, 209
 Phase plane, 150
 Phase response curve (PRC), 60, 61, 64, 65, 67, 72, 74, 75, 78
 ProMoT, 259, 267
 Promoter, 276, 277, 279, 280
 Protein, 274–276, 278, 291

Protein association network, 3, 5, 7–10, 12, 13, 20, 21, 24
 Protein-protein interaction (PPI), 4, 7–9, 11, 12, 21, 22

Q

Qualifier, 205, 208

R

Reaction rule, 202–206, 208, 209, 212, 215
 Regulatory network, 143–145, 157, 159, 163
 Ribosome binding site (RBS), 258, 259, 263–266, 269
 RNA, 202, 205
 RNA polymerase, 276, 278, 279
 Robustness analysis, 149
 Rule-based modeling, 201, 202, 204

S

Sanger sequencing, 32
 Sensitivity, 60, 61, 66–72, 74–80
 Sensitivity analysis, 152
 Sequence library, 32
 Signal transduction, 220, 224, 238, 240, 242
 Signaling pathway, 85, 89, 100, 102, 107
 Single strand meta-DNA (ssmDNA), 172, 177, 178, 180, 183, 185, 189–192, 196
 Small RNA, 258, 259, 263, 266, 267, 269
 Species, 202–206, 212, 213, 215
 Specification functionals, 247, 248, 252, 253
 SRNA, 259, 263, 264, 266, 267
 State space, 248
 Stoichiometric matrix, 248
 Storage function, 119, 120, 124, 126–128, 132, 133
 Strand displacement, 173, 174, 176, 177, 182–184, 191, 192
 Supply rate, 119, 120

T

Temporal logic, 246, 247
 Terminator, 258, 259, 263
 Transcriptional circuit, 274, 278, 279, 286, 291

W

Watson-Crick base pair interactions, 274
 Wildcard, 205

Y

Yeast, 84–86, 94, 97, 99, 106, 113, 114