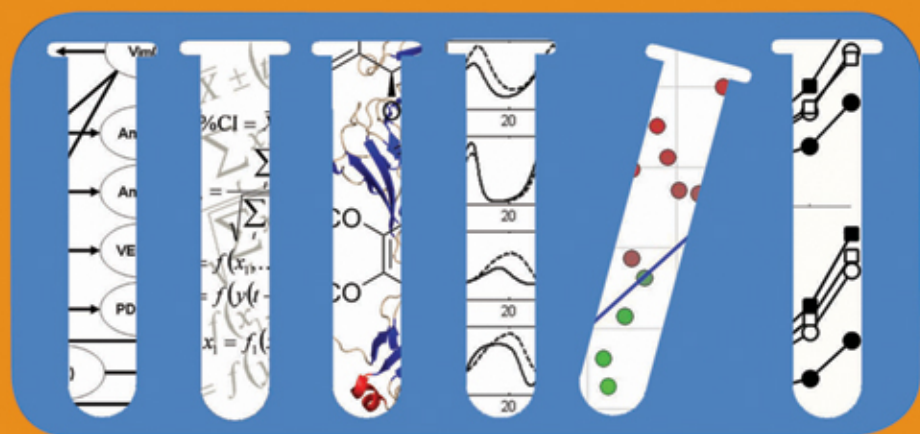


Wiley Series on Technologies for the Pharmaceutical Industry
Sean Ekins, Series Editor

Systems Biology in Drug Discovery and Development



Edited by
Daniel L. Young and Seth Michelson

SYSTEMS BIOLOGY IN DRUG DISCOVERY AND DEVELOPMENT

Wiley Series on Technologies for the Pharmaceutical Industry

Sean Ekins, Series Editor

Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals

Edited by Sean Ekins

Pharmaceutical Applications of Raman Spectroscopy

Edited by Slobodan Sasic

Pathway Analysis for Drug Discovery: Computational Infrastructure and Applications

Edited by Anton Yuryev

Drug Efficacy, Safety, and Biologics Discovery: Emerging Technologies and Tools

Edited by Sean Ekins and Jinghai J. Xu

The Engines of Hippocrates: From the Dawn of Medicine to Medical and Pharmaceutical Informatics

Barry Robson and O. K. Baek

Pharmaceutical Data Mining: Approaches and Applications for Drug Discovery

Edited by Konstantin V. Balakin

The Agile Approach to Adaptive Research: Optimizing Efficiency in Clinical Development

Michael J. Rosenberg

Pharmaceutical and Biomedical Project Management in a Changing Global Environment

Edited by Scott D. Babler

Systems Biology in Drug Discovery and Development

Edited by Daniel L. Young and Seth Michelson

Editorial Advisory Board

Dr. Renee Arnold (ACT LLC, USA)

Dr. David D. Christ (SNC Partners LLC, USA)

Dr. Michael J. Curtis (Rayne Institute, St Thomas' Hospital, UK)

Dr. James H. Harwood (Delphi BioMedical Consultants, USA)

Dr. Maggie A.Z. Hupcey (PA Consulting, USA)

Dr. Dale Johnson (Emiliem, USA)

Prof. Tsuguchika Kaminuma (Tokyo Medical and Dental University, Japan)

Dr. Mark Murcko (Vertex, USA)

Dr. Peter W. Swaan (University of Maryland, USA)

Dr. Ana Szarfman (Food and Drug Administration, USA)

Dr. David Wild (Indiana University, USA)

SYSTEMS BIOLOGY IN DRUG DISCOVERY AND DEVELOPMENT

Edited by

Daniel L. Young
Seth Michelson



A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2012 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Systems biology in drug discovery and development / edited by Daniel L. Young, Seth, Michelson.

p. ; cm.

Includes bibliographical references.

ISBN 978-0-470-26123-1 (cloth)

1. Drug development. 2. Systems biology. I. Young, Daniel L., editor.
II. Michelson, Seth., editor.

[DNLM: 1. Drug Discovery—methods. 2. Models, Biological. 3. Pharmacokinetics. 4. Pharmacological Processes. 5. Systems Biology—methods. QV 744]

RM301.25.S97 2011

615'19—dc22

2010043475

Printed in the United States of America

oBook ISBN: 978-1-118-01643-5

ePDF ISBN: 978-1-118-01641-1

ePub ISBN: 978-1-118-01642-8

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE	xi
CONTRIBUTORS	xv
PART I INTRODUCTION TO SYSTEMS BIOLOGY IN APPROACH	
1. Introduction to Systems Biology in Drug Discovery and Development	3
<i>Seth Michelson and Daniel L. Young</i>	
Systems Biology in Pharmacology	3
References	5
2. Methods for <i>In Silico</i> Biology: Model Construction and Analysis	7
<i>Theresa Yuraszeck, Peter Chang, Kalyan Gayen, Eric Kwei, Henry Mirsky, and Francis J. Doyle III</i>	
2.1 Introduction	7
2.2 Model Building	8
2.3 Parameter Estimation	21
2.4 Model Analysis	28
2.5 Conclusions	32
References	32
3. Methods in <i>In Silico</i> Biology: Modeling Feedback Dynamics in Pathways	37
<i>Peter Wellstead and Olaf Wolkenhauer</i>	
3.1 Introduction	37
3.2 Statistical Modeling	39
3.3 Mathematical Modeling	46
3.4 Feedback and Feedforward	49
3.5 Conclusions	56
References	56
	v

4. Simulation of Population Variability in Pharmacokinetics	59
<i>Jiansong Yang</i>	
4.1 Introduction	59
4.2 PBPK Modeling	60
4.3 Simulation of Pharmacokinetic Variability	61
4.4 Conclusions and Future Directions	79
References	80

PART II APPLICATIONS TO DRUG DISCOVERY

5. Applications of Systems Biology Approaches to Target Identification and Validation in Drug Discovery	95
<i>Bart S. Hendriks</i>	
5.1 Introduction	95
5.2 Typical Drug Discovery Paradigm	97
5.3 Integrated Drug Discovery	99
5.4 Drivers of the Disease Phenotype: Clinical Endpoints and Hypotheses	100
5.5 Extracellular Disease Drivers: Mechanistic Biotherapeutic Models	106
5.6 Relevant Cell Models for Clinical Endpoints	109
5.7 Intracellular Disease Drivers: Signaling Pathway Quantification	110
5.8 Target Selection: Dynamic Pathway Modeling	117
5.9 Conclusions	123
References	125
6. Lead Identification and Optimization	135
<i>Seth Michelson</i>	
6.1 Introduction	135
6.2 The Systems Biology Tool Kit	139
6.3 Conclusions	142
References	143
7. Role of Core Biological Motifs in Dose–Response Modeling: An Example with Switchlike Circuits	147
<i>Sudin Bhattacharya, Qiang Zhang, and Melvin E. Andersen</i>	
7.1 Introduction: Systems Perspectives in Drug Discovery	147
7.2 Systems Biology and Toxicology	148

7.3	Mechanistic and Computational Concepts in a Molecular or Cellular Context	151
7.4	Response Motifs in Cell Signaling and Their Role in Dose Response	152
7.5	Discussion and Conclusions	165
	References	169

8. Mechanism-Based Pharmacokinetic–Pharmacodynamic Modeling During Discovery and Early Development **175**

Hans Peter Grimm, Ying Ou, Micaela Reddy, Pascale David-Pierson, and Thierry Lavé

8.1	Introduction	175
8.2	Challenges in Drug Discovery and Development	176
8.3	Methodological Aspects and Concepts	179
8.4	Use of PK–PD Models in Lead Optimization	183
8.5	Use of PK–PD Models in Clinical Candidate Selection	188
8.6	Entry-into-Human Preparation and Translational PK–PD Modeling	189
8.7	Use of PK–PD Models in Toxicology Study Design and Evaluation	189
8.8	Justification of Starting Dose, Calculation of Safety Margins, and Support of Phase I Design	191
8.9	Phase I and Beyond	193
8.10	Support of Early Formulation Development	195
8.11	Outlook and Conclusions	196
	References	197

PART III APPLICATIONS TO DRUG DEVELOPMENT

9. Developing Oncology Drugs Using Virtual Patients of Vascular Tumor Diseases **203**

Zvia Agur, Naamah Bloch, Boris Gorelik, Marina Kleiman, Yuri Kogan, Yael Sagi, D. Sidransky, and Yael Ronen

9.1	Introduction	203
9.2	Modeling Angiogenesis	205
9.3	Use of Rigorous Mathematical Analysis to Gain Insight into Drug Development	213
9.4	Use of Angiogenesis Models in Theranostics	220
9.5	Use of Angiogenesis Models in Drug Salvage	226
9.6	Summary and Conclusions	230
	References	231

10. Systems Modeling Applied to Candidate Biomarker Identification	239
<i>Ananth Kadambi, Daniel L. Young, and Kapil Gadkar</i>	
10.1 Introduction	239
10.2 Biomarker Discovery Approaches	245
10.3 Examples of Systems Modeling Approaches for Identification of Candidate Biomarkers	252
10.4 Conclusions	260
References	260
11. Simulating Clinical Trials	265
<i>Tom Parke</i>	
11.1 Introduction	265
11.2 Types of Models Used in Clinical Trial Design	272
11.3 Sources of Prior Information for Designing Clinical Trials	276
11.4 Aspects of a Trial to Be Designed and Optimized	277
11.5 Trial Simulation	279
11.6 Optimizing Designs	281
11.7 Real-World Examples	283
11.8 Conclusions	284
References	284
 PART IV SYNERGIES WITH OTHER TECHNOLOGIES	
12. Pathway Analysis in Drug Discovery	289
<i>Anton Yuryev</i>	
12.1 Introduction: Pathway Analysis, Dynamic Modeling, and Network Analysis	289
12.2 Software Systems for Pathway Analysis	292
12.3 Pathway Analysis in the Modern Drug Development Pipeline	293
12.4 Conclusions	298
References	299
13. Functional Mapping for Predicting Drug Response and Enabling Personalized Medicine	303
<i>Yao Li, Wei Hou, Wei Zhao, Kwangmi Ahn, and Rongling Wu</i>	
13.1 Introduction	304
13.2 Functional Mapping	306
13.3 Predictive Model	311
13.4 Future Directions	315
References	318

14. Future Outlook for Systems Biology	323
<i>Daniel L. Young and Seth Michelson</i>	
14.1 Introduction	323
14.2 System Complexity in Biological Systems	324
14.3 Models for Quantitative Integration of Data	325
14.4 Changing Requirements for Systems Approaches During Drug Discovery and Development	328
14.5 Better Models for Better Decisions	330
14.6 Advancing Personalized Medicine	334
14.7 Improving Clinical Trials and Enabling More Complex Treatment Approaches	337
14.8 Collaboration and Training for Systems Biologists	340
14.9 Conclusions	342
References	343
INDEX	349

PREFACE

Despite the wealth of data describing mechanisms underlying health and disease in living systems, health care costs continue to rise, and there is a growing need for improved and more affordable treatments. Efficient drug discovery and development requires methods for integrating preclinical data with patient data into a unified framework to project both efficacy and safety outcomes for new compounds and treatment approaches.

In this book we present the foundations of systems biology, a growing multidisciplinary field, applied specifically to drug discovery and development. Systems biology formally integrates knowledge and information from multiple biological sources into a coherent whole by employing proven engineering and mathematical modeling approaches. The integrated system allows rapid analysis and simulation that can inform and optimize the drug research and development processes, by formalizing, and testing, the set of acceptable hypotheses *in silico*, thereby reducing development time and costs and ultimately improving the efficacy of novel treatments.

This book is the first systems biology text to focus on how systems biology can be specifically applied to enhance drug discovery and development, with particular emphasis on real-world examples. Other texts on systems biology to date have focused on particular subdisciplines of systems biology (such as cellular networks) and have not specifically addressed drug discovery and development. This book introduces key methodologies and technical approaches for helping to solve many of the current challenges facing the pharmaceutical and biotechnology industries.

The target audience for the book includes those training or currently involved in all phases of drug discovery and development. Specific examples include life scientists, pharmacologists, computational and systems biology modelers, bioinformaticians, clinicians, and pharmaceutical/biotech management. The methods and case studies presented here will help researchers understand the diverse applications of the systems approach and integrate these technologies into their drug discovery and development programs. Those who incorporate these approaches successfully should increase their organization's competitiveness to address unmet market needs as well as more complex diseases and therapies.

The book is divided into four complementary parts. Providing a foundation for the techniques of systems biology, Part I provides an introduction to

engineering and mathematical methods employed to characterize biological systems. In particular, Chapter 2 overviews model construction and analysis, focusing on model building, parameter estimation, model validation, and sensitivity analysis. Chapter 3 presents general statistical modeling approaches as well as methods for representing and analyzing nonlinear dynamical biochemical networks, of which feedback and feedforward loops are central players. In addition to modeling fundamental biological interactions and dynamics, an essential element of the systems biology approach is the study and simulation of population-level variability. To this end, Chapter 4 presents how drug pharmacokinetics is affected by variations in drug absorption, distribution, metabolism, and excretion, illustrating methods for predicting interindividual variability essential for rationale compound evaluation.

Part II highlights systems biology techniques aimed at enhancing the drug discovery process. An essential component of drug discovery is target identification and validation. To tackle many of the challenges inherent in these processes, Chapter 5 introduces a variety of complementary systems approaches, including text-mining, disease and therapeutics modeling, large multicontext data sets, regression modeling, and network and dynamic pathway modeling. In Chapter 6, systems biology approaches are applied to lead identification and optimization disciplines. In particular, systems approaches are shown to enable building bridges between compounds' chemical and biological activities. In this way, lead identification and optimization are enhanced by the systematic quantification of the optimal pharmacokinetic and pharmacodynamic compound profiles, defined potentially for specific patient populations. Chapter 7 addresses drug safety by exploring the role of biological motifs, in particular switchlike circuits, critical for dose–response models. Such models help uncover complex emergent behaviors and reveal factors driving variable patient responses to drugs that could limit efficacy or even lead to low-incidence adverse responses. Finally, Chapter 8 presents the use of mechanistic systems models for the study of pharmacokinetics and pharmacodynamics during discovery and early development. These models integrate a mechanistic understanding of biology and disease processes into a framework to aid in the selection of lead compounds, evaluation of dosing regimens, and support of optimal study design for specific patient populations.

Part III addresses particular applications of systems biology to drug development. Illustrating practical drug development challenges, Chapter 9 details the development and validation of a multiscale mathematical model for angiogenesis, integrating molecular and tissue-level processes. Here the exemplary model is applied for treatment personalization, and results suggest that an arrested drug candidate can be efficacious if applied in combination with current standards of care. Chapter 10 presents methods for applying systems biology to candidate biomarker identification. In particular, the chapter highlights the biomarker discovery process, its application to drug development, and the utility of mechanistic systems modeling to biomarker development in cardiovascular disease and rheumatoid arthritis. Finally, to aid in the design

and execution of costly clinical programs, essential aspects of clinical trial simulations are presented in Chapter 11, where both clinical efficacy and safety are essential considerations.

In the final section of the book, Part IV, we address how systems biology technologies can synergize with other approaches. To this end, Chapter 12 presents how biological pathway analysis can be integrated into drug discovery systems approaches. Chapter 13 addresses aspects of personalized medicine and how functional mapping aimed at understanding genes and genetic networks can be used to help predict drug responses in patients. The book concludes in Chapter 14 with a broad overview of opportunities and challenges in systems biology that should ultimately help to extend both its reach and its acceptance, thereby further enhancing pharmaceutical productivity and the success of drug discovery and development for the benefit of patients.

In addition to the contributing authors of this book, we would like to thank our collaborators and colleagues throughout the years who have helped develop and apply systems biology approaches to drug discovery and development. We look forward to future advances and successes in the coming years as these approaches are applied and extended by dedicated researchers for enhanced drug discovery and development and ultimately, better care for patients.

Palo Alto, California

DANIEL L. YOUNG

Redwood City, California

SETH MICHELSON

CONTRIBUTORS

Zvia Agur, Institute for Medical Biomathematics, Bene-Ataroth, Israel

Kwangmi Ahn, Department of Public Health Sciences, Pennsylvania State College of Medicine, Hershey, Pennsylvania

Melvin E. Andersen, Division of Computational Biology, The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina

Sudin Bhattacharya, Division of Computational Biology, The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina

Naamah Bloch, Optimata Ltd., Ramat-Gan, Israel

Peter Chang, Chemical Engineering Department, University of California, Santa Barbara, California

Pascale David-Pierson, Modeling and Simulation Group, Drug Metabolism and Pharmacokinetics Department, F. Hoffmann–La Roche Ltd., Basel, Switzerland

Francis J. Doyle III, Chemical Engineering Department, University of California, Santa Barbara, California

Kapil Gadkar, Theranos, Inc., Palo Alto, California

Kalyan Gayen, Chemical Engineering Department, University of California, Santa Barbara, California

Boris Gorelik, Optimata Ltd., Ramat-Gan, Israel

Hans Peter Grimm, Modeling and Simulation Group, Drug Metabolism and Pharmacokinetics Department, F. Hoffmann–La Roche Ltd., Basel, Switzerland

Bart S. Hendriks, Merrimack Pharmaceuticals, Cambridge, Massachusetts

Wei Hou, Department of Epidemiology and Health Policy Research,
University of Florida, Gainesville, Florida

Ananth Kadambi, Entelos Inc., Foster City, California

Marina Kleiman, Optimata Ltd., Ramat-Gan, Israel

Yuri Kogan, Institute for Medical Biomathematics, Bene-Ataroth, Israel

Eric Kwei, Chemical Engineering Department, University of California,
Santa Barbara, California

Thierry Lavé, Modeling and Simulation Group, Drug Metabolism and
Pharmacokinetics Department, F. Hoffmann–La Roche Ltd., Basel,
Switzerland

Yao Li, Quantitative Genetic Epidemiology, Fred Hutchinson Cancer
Research Center, Seattle, Washington

Seth Michelson, Genomic Health Inc., Redwood City, California

Henry Mirsky, Chemical Engineering Department, University of California,
Santa Barbara, California

Ying Ou, Modeling and Simulation Group, Drug Metabolism and
Pharmacokinetics Department, Roche Palo Alto LLC, Palo Alto, California

Tom Parke, Tessella plc, Oxfordshire, UK

Micaela Reddy, Modeling and Simulation Group, Drug Metabolism and
Pharmacokinetics Department, Roche Palo Alto LLC, Palo Alto, California

Yael Ronen, Optimata Ltd., Ramat-Gan, Israel

Yael Sagi, Optimata Ltd., Ramat-Gan, Israel

D. Sidransky, The Johns Hopkins University School of Medicine, Baltimore,
Maryland

Peter Wellstead, The Hamilton Institute, NUIM, Maynooth, Republic of
Ireland

Olaf Wolkenhauer, Systems Biology and Bioinformatics, University of
Rostock, Rostock, Germany

Rongling Wu, Center for Statistical Genetics, Pennsylvania State University,
Hershey, Pennsylvania

Jiansong Yang, Simcyp Ltd., Sheffield, UK

Daniel L. Young, Theranos Inc., Palo Alto, California

Theresa Yuraszeck, Chemical Engineering Department, University of
California, Santa Barbara, California

Anton Yuryev, Ariadne Genomics Inc, Rockville, Maryland

Qiang Zhang, Division of Computational Biology, The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina

Wei Zhao, Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, Tennessee

PART I

INTRODUCTION TO SYSTEMS BIOLOGY IN APPROACH

Introduction to Systems Biology in Drug Discovery and Development

SETH MICHELSON

Genomic Health Inc., Redwood City, California

DANIEL L. YOUNG

Theranos Inc., Palo Alto, California

Summary

Over the last several decades, medical and biological research has opened vast windows into the mechanisms underlying health and disease in living systems. Integrating this knowledge into a unified framework to enhance understanding and decision making is a significant challenge for the research community. Efficient drug discovery and development requires methods for bridging pre-clinical data with patient data to project both efficacy and safety outcomes for new compounds and treatment approaches. In this book we present the foundations of systems biology, a growing multidisciplinary field applied specifically to drug discovery and development. These methods promise to accelerate time lines, to reduce costs, to decrease portfolio failure rates, and most significantly, to improve treatment by enhancing the workflow, and thus the competitiveness, of pharmaceutical and biotechnology organizations. Ultimately, these improvements will improve overall health care and its delivery.

SYSTEMS BIOLOGY IN PHARMACOLOGY

Discovering a new medicine is a multistep process that requires one to:

- Identify a biochemically based cause–effect pathway (or pathways) inherent in a disease and its pathophysiology

Systems Biology in Drug Discovery and Development, First Edition.

Edited by Daniel L. Young, Seth Michelson.

© 2012 John Wiley & Sons, Inc. Published 2012 by John Wiley & Sons, Inc.

- Identify those cells and molecular entities (e.g., receptors, cytokines, genes) involved in the control of those pathways (typically termed *targets*)
- Identify an exogenous entity that can manipulate a molecular target to therapeutic advantage (typically termed a *drug*)
- Identify, with some level of specificity, how manipulation modulates the disease effects (termed the *mechanism of action* of the drug)
- Identify that segment of the patient population most likely to respond to manipulation (typically through the use of appropriate surrogates termed *biomarkers*)

Given these challenges, pharmaceutical drug discovery and development is an extremely complex and risky endeavor. Despite growing industry investment in research and development, only one in every 5000 new drug candidates is likely to be approved for therapeutic use in the United States (PhRMA, 2006). In fact, approximately 53% of compounds that progress to phase II trials are likely to fail, resulting in amortized costs of between \$800 million and \$1.7 billion per approved drug (DiMasi et al., 2003; Gilbert et al., 2003; Pharmaceutical Research and Manufacturers of America, 2006). Clearly, the crux of the problem is the failure rate of compounds, especially those in late-stage clinical development. To solve this problem, one must clearly identify the most appropriate compound for the most appropriate target in the most appropriate subpopulation of patients, and then dose those patients as optimally as possible. This philosophy forms the cornerstone of the “learn and confirm” model of drug development suggested by Sheiner in 1997.

For example, to address these three issues specifically, the Center for Drug Development Science at the University of California–San Francisco has developed a set of guidelines for applying one particular *in silico* technology, biosimulation, to the drug development process (Holford et al., 1999).

These guidelines define a three-step process. During step 1, the most relevant underlying biology describing the pathophysiology of the disease is characterized, as are the pharmacokinetics of any candidate compound aimed at its treatment. In step 2, the various clinical subpopulations expected to receive the compound are identified and characterized, including measures of interpatient variability in drug absorption, distribution, metabolism, and excretion, and compound-specific pharmacodynamics are established. Once steps 1 and 2 are complete, this information is used in step 3 to simulate and thus design the most efficient clinical trial possible.

We believe that the general principles outlined above should not be restricted to only one methodology (i.e., biosimulation) but should be extended to the entire spectrum of *in silico* technologies that make up the generic discipline called *systems biology*. Systems biology is a rapidly developing suite of technologies that captures the complexity and dynamics of disease progression and response to therapy within the context of *in silico* models. Whether these models and their incumbent analytical methodologies represent explicit physi-

ological models and dynamics, statistical associations, or a mix thereof, *en suite* they provide the pharmaceutical researcher with access to the most pertinent information available. By definition, that information must be composed of those data that best characterize the disease and its pathophysiology, the compound and its mechanism of action, and the patient populations in which the compound is most likely to work. With the advance of newer and faster assay technologies, the gathering of those data is no longer the rate-limiting process it once was. Rather, technologies capable of sampling the highly complex spaces underlying biological phenomena have made the interpretation of those data in the most medically and biologically reasonable context the next great hurdle in pharmaceutical drug discovery and development.

To address these challenges adequately, the pharmaceutical or clinical researcher must be able to understand and characterize the effects of diverse chemical entities on the pathways of interest *in the context of the biology they are meant to affect*. To accomplish that, research scientists and clinicians must have at their disposal the means to acquire the most pertinent and predictive information possible. We believe that systems biology is a particularly attractive solution to this problem. It formally integrates knowledge and information from multiple biological sources into a coherent whole by subjecting them to proven engineering, mathematical, and statistical methodologies. The integrated nature of the systems biology approach allows for rapid analysis, simulation, and interpretation of the data at hand. Thus, it informs and optimizes the pharmaceutical discovery and development processes, by formalizing, and testing, the most biologically relevant family of acceptable hypotheses *in silico*, thereby enabling one to reduce development time and costs and improve the efficacy of novel treatments.

REFERENCES

- DiMasi, J.A., Hansen, R.W., and Grabowski, H.G. (2003). The price of innovation: new estimates of drug development costs. *J Health Econ* 22, 151–185.
- Gilbert, J., Henske, P., and Singh, A. (2003). Rebuilding big pharma's business model. *In Vivo* 21, 1–10.
- Holford, N.H.G., Hale, M., Ko, H.C., Steimer, J.-L., Sheiner, L.B., and Peck, C.C. (1999). Simulation in drug development: good practices. <http://bts.ucsf.edu/cdds/research/sddgpreport.php>.
- PhRMA (2006). *Pharmaceutical Industry Profile 2006*. Pharmaceutical Research and Manufacturers of America, Washington, DC.
- Sheiner, L.B. (1997). Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 61, 275–291.

Methods for *In Silico* Biology: Model Construction and Analysis

THERESA YURASZECK, PETER CHANG, KALYAN GAYEN, ERIC KWEI,
HENRY MIRSKY, and FRANCIS J. DOYLE III

University of California, Santa Barbara, California

2.1. INTRODUCTION

Despite increasing investment in research and development, the productivity of the pharmaceutical industry has been declining, and this unfortunate phenomenon necessitates novel approaches to drug discovery and development. Systems biology is an approach that shows great promise for identifying and validating new drug targets and may ultimately facilitate the introduction of personalized and preventive medicine. This interdisciplinary field integrates traditional experimental techniques from molecular biology and biochemistry with computational biology, modeling and simulation, and systems analysis to construct quantitative mathematical models of biological networks in order to investigate their behavior. The utility of such models depends on their predictive abilities. Although constructing models that can predict all phenotypes and perturbation responses is not feasible at present, it is tractable to develop models of sufficient detail and scope to predict behavioral responses to particular perturbations and to perform sensitivity analyses. Model building, validation, and analysis are usually iterative processes in which the model becomes successively closer to the reality of the biological network and its predictions become more accurate. In this chapter we introduce model building, parameter estimation, model validation, and sensitivity analysis and present case studies in each section to demonstrate these concepts.

Systems Biology in Drug Discovery and Development, First Edition.

Edited by Daniel L. Young, Seth Michelson.

© 2012 John Wiley & Sons, Inc. Published 2012 by John Wiley & Sons, Inc.

2.2. MODEL BUILDING

2.2.1. Types of Models

Systems biologists use a variety of models to describe biological data. These models can be categorized into interaction-, constraint-, or mechanism-based models (Stelling, 2004). *Interaction-based models* represent network topology without consideration for reaction stoichiometry and kinetics. Topology maps reveal the modular organization of biological networks, a property that facilitates the study of biological organisms because it suggests that subnetworks can be studied in isolation. These maps also reveal the principles by which cellular networks are organized. Such principles provide insight into network behaviors.

Constraint-based approaches utilize information about interaction partners, stoichiometry, and reaction reversibility but contain no dynamic information. Due to the availability of such data, metabolic networks are frequently analyzed using constraint-based approaches. This approach can elucidate the range of phenotypes and behaviors that a system can achieve given the stoichiometry, interaction, and reversibility constraints. It has also been used to predict the optimal distribution of metabolic fluxes within a system from the range of possible solutions, where the optimal distribution is that which maximizes or minimizes some assumed objective, such as biomass production (Famili et al., 2003). Such analyses give insight into the behavior of an organism not only as it currently exists, but also its evolution; if the *in silico* predictions are in agreement with the experimental data, the assumption that the organism evolved to produce the optimized function is consistent with the data.

The most detailed models, the *mechanism-based models*, capture reaction stoichiometry and kinetics, providing quantitative insights into the dynamic behavior of biological networks. These models require substantial amounts of information about network connectivity and kinetic parameters. These requirements have limited the application of these models, although there are several systems for which this type of model has been constructed successfully. Such models are advantageous because they generate testable experimental hypotheses about dynamic cellular behavior. They also facilitate *in silico* experiments designed to elucidate biological design principles. For example, a model of the heat shock response in *Escherichia coli* was analyzed to determine the role of the feedback and feedforward loops that characterize this system (El-Samad et al., 2005). The *heat shock response* (HSR) is a mechanism that compensates for stress in the cytoplasm. Stress leads to the accumulation of unfolded and misfolded proteins and subsequently triggers the HSR, which induces the expression of genes that relieve the accumulation of these denatured proteins in the cytoplasm. Induced genes include those that encode chaperone proteins, which facilitate the folding of unfolded and misfolded proteins, and proteases to eliminate denatured proteins from the system. The HSR is a tightly controlled process governed by a complex regulatory architecture consisting of

interconnected feedback and feedforward loops. Although simpler systems could in theory also prevent protein accumulation, evolution and natural selection led to this more complex design. *In silico* experiments in which the feedback and feedforward loops were removed from the system successively showed that this relatively complex design provides enhanced robustness compared to simpler systems (El-Samad et al., 2005). These insights would be difficult if not impossible to generate *in vivo*.

2.2.2. Specification of Model Granularity and Scope

One of the design challenges a modeler faces is that of determining the appropriate granularity and scope of a model. These choices are made based on the intended purpose of the model and the available data. When designed prudently, models will yield useful testable predictions and provide insights to pertinent mechanisms underlying an observed behavior. *Granularity* defines the level of scale that a model encompasses for a given biological network. In modeling biological systems, granularity from the level of molecules to cells to organ systems is considered. Usually, a model encompasses several levels based on the available data, the current understanding of the biological components, the model complexity, and the intended model applications. The appropriate level of granularity is also determined by considering the biological properties and behaviors of interest.

On the other hand, *scope* describes the extent of mechanistic details represented in a model. For example, at the molecular level, one has to decide which molecular components and reactions to include, and when modeling tissue behavior, one may have to decide what cell types to include. Mechanism-based models are typically very granular but reduced in scope compared to less detailed but larger-scoped topology networks. Constraint-based models are intermediate in scope and detail. Regardless of the modeling approach, the appropriate level of abstraction, taking into consideration granularity and scope, will yield consistent links between biological levels without including every detail (Stelling, 2004). The case study presented in Section 2.2.6 illustrates the impact of granularity and scope on model predictions.

2.2.3. Approaches to Model Construction

Model construction can be approached in a top-down or bottom-up manner. Top-down approaches are essentially a reverse-engineering exercise and are not to be confused with the traditional reductionist approach frequently taken by biologists. The top-down approach to *in silico* model building starts with genome-wide data, such as microarray data, and attempts to infer the underlying networks leading to the observed behavior from these data. This type of approach is facilitated by the availability of high-throughput data and is advantageous when mechanistic details and connectivity, or the wiring diagram for a system, are not well known (Kholodenko et al., 2002). The building of more

empirical models in which the mechanistic details are “lumped” together is also considered a top-down approach; this results in a model that captures the relevant behavior although the mechanistic details are masked. Bottom-up approaches, on the other hand, combine connectivity and pathway information into a larger network. They start with the constitutive elements, such as genes or proteins, link them to their interaction partners, and identify the reaction-rate parameters associated with each interaction. Both top-down and bottom-up approaches can lead to detailed models able to predict dynamic response to perturbations.

A method that combines concepts from the top-down and bottom-up approaches has been proposed and applied with success to model protein folding (Hildebrandt et al., 2008). This top-down mechanistic modeling approach starts with the most basic mathematical model possible and successively expands the model scope. The impact of each model addition on the system’s performance is evaluated, elucidating the structural requirements of the system (Hildebrandt et al., 2008). In essence, this top-down approach starts with a model that captures limited mechanistic detail of the system and elucidates the most critical network interactions as it progressively adds detail to the wiring diagram, ultimately resulting in a highly detailed mechanistic model. A case study employing this method to study protein folding of a single-chain antibody is described in Section 2.2.8.

2.2.4. Metabolic Network Analysis

Metabolic behavior is closely associated with phenotype, and the sequencing of the human genome enables the possibility of metabolic network analysis (Cornish-Bowden and Cardenas, 2000; Oliveira et al., 2005; Schwartz et al., 2007). Metabolic networks are highly complex, formed by hundreds of densely interconnected chemical reactions. Powerful computational tools are required to characterize such complex metabolic systems (Famili et al., 2003; Klamt and Stelling, 2003; Nielsen, 1998; Palsson et al., 2003; Reed and Palsson, 2003; Schilling et al., 2000; Wiback et al., 2004).

Two basic approaches are available for metabolic network analysis. First, the kinetic approach is based on fundamental reaction engineering principles, but this approach generally suffers from a lack of detailed kinetic information. The Palsson group (University of California–San Diego) has developed a dynamic model for a human red blood cell, a system for which detailed kinetic information is available. Second, structural approaches require only the stoichiometry of the metabolic network. For a structure-based metabolic network analysis, four approaches are available:

1. Metabolic flux analysis
2. Flux balance analysis
3. Extreme pathway analysis
4. Elementary mode analysis