Preface

(Molecular) Systems biology has developed over roughly the past 10 years. Its emergence has led to the development of broad genome-wide or network-wide viewpoints of organism functions that have developed against the context of whole genome sequences. Bottom-up approaches to network reconstruction have resulted in organism-specific networks that have a direct genetic and genomic basis. Such networks are now available for a growing number of organisms.

Genome-scale networks have been used to develop constraint-based reconstruction and analysis (COBRA) procedures that treat structural properties of networks, their physiological capabilities, optimal functional states of organisms, and studies of adaptive and long-term evolution. These topics are treated in the companion book that emphasizes that while biology is dynamic, it still functions under the constraints of the topological structure of the molecular networks that underlie its functions.

Events over the time scales associated with distal causation in biology, i.e., over multiple generations, can be studied within the COBRA framework. However, analysis of proximal or immediate dynamic responses of organisms is limited. The recent development of high-throughput technologies and the availability of omics data sets has opened up an alternative approach to building large-scale models that can compute the dynamic states of biological networks. Omics-based abundance measurements (i.e., for proteins, transcripts, and metabolites) can now be mapped onto network reconstructions. In addition, functional states can be determined from fluxomic, exo-metabolomic, and various physiological data types.

The combination of omics data sets and network reconstructions allows the generation of Mass Action Stoichiometric Simulation (MASS) models. Such models can, at this point in time, be formulated for metabolism and associated enzymes and other protein molecules. MASS models will be condition specific, as they use particular data sets. In principle, MASS models can be formulated for any cellular phenomena for which reconstructions and omics data sets are available. Although the procedure is now established, some of the practical issues associated with its broad implementation will need additional experience that will call on further research in this field.

This book is focused on the process and the issues associated with the generation of MASS models. Their foundational concepts are described and they are applied to specific cases. Once the reader has mastered these concepts and gone through the details of their application to familiar cellular processes, you should be able to build MASS models for cellular phenomena of interest.

One should be aware of the fact that dynamic models have been constructed to describe biological phenomena for many decades. At the biochemical level, such models have been largely based on biophysical principles, heavily focused in particular on the use of *in-vitro*-derived rate laws. Given the scarcity of such rate laws, this approach to building kinetic models has limited the scope and size of dynamic models built in this fashion. The omics data-driven MASS procedure provides an alternative condition-dependent approach that is scalable.

This book, in a sense, brings my career full circle. My first love in graduate school was building complex dynamic models in biology based on the contents of the graduate curriculum in chemical engineering. However, as stated above, the application of these methods to biology was necessarily limited due to data availability and due to the "absolute" characteristics of biophysical models. The path through stoichiometric models from the biochemical to the genome scale based on full genome sequences, to large-scale dynamic models based on omics data sets has been an interesting one. Given the impending onslaught of genetic data and associated potential for biological variation, this field might be just in its infancy.

This text is constructed to teach how to build complex dynamic models of biochemical networks and how to simulate their responses. The material has been taught both at the undergraduate and graduate level at UC San Diego since 2008. Teaching the material at these two levels has led to the development of a set of homework problems (Appendix B) and a collection of Mathematica workbooks. It is my intent to make these available through an on-line source, initially on http://systemsbiology.ucsd.edu. I hope both will be helpful to instructors.

The path to this book has had many influences. Reich and Selkov's 1982 book, *Energy Metabolism of the Cell*, certainly contains many foundational and influential concepts. The *Color Atlas of Biochemistry* by Koolman and Roehm provides succinct representation of biochemical knowledge that has been useful in developing the material. All the computations in the text were done in Mathematica. Throughout my entire career, LaTeX has been an essential resource, as it was for writing this book.

There are special thanks due to two individuals. Neema Jamshidi has been an MD/PhD student in my lab over the past 6 or 7 years. He has been a fantastic colleague and friend. He educated me about the use of Mathematica and tirelessly answered my repeated and often naive questions. He has also been a source of great intellectual stimulation and discussions. He was a major influence in completing this book. As with the companion book, Marc Abrams made the writing, preparation, editing, and production of this book possible. He supervised, coordinated, and implemented the construction of the LaTeX document and the preparation of many figures in the text. Special thanks to these two gentlemen.

In addition, three PhD students in my lab were of invaluable help in getting this book to the state of completion that it has reached. Aarash Bordbar helped me with the formulation of the complicated Mathematica workbooks for Part IV of the text. In addition, he has played a notable role in developing the work flow for MASS models. Daniel Zielinski not only helped build the Mathematica workbooks for Part III, but proofread the text with his impeccable eye for detail and logical flow of material. Addiel U. de Alba Solis helped with the Mathematica workbooks for Parts I and II of the text. All three were very helpful in reviewing, correcting, and providing solutions to the homework sets given in Appendix B.

Others have helped with this text either indirectly or directly through thoughtful comments or the preparation of illustrations. For their assistance, I am grateful: Kenyon Applebee, Tom Conrad, Markus Herrgard, Joshua Lerman, Vasiliy Portnoy, Jan Schellenberger, Paolo Vicini and Michael Zager.

This book is dedicated to my parents, who enabled, allowed, supported, and encouraged me to pursue my studies and interests in integrated biological processes. Without them I would not have reached this level of professional development and would not have written this book. Kærar þakkir.

Bernhard Palsson
La Jolla, CA
April 2010