Preface

Entering a new millennium invites predictions. What will daily life be like, and where shall we be scientifically, philosophically, and socially, a hundred years from now? A thousand years from now? If the development of scientific knowledge during the twentieth century is any indication, we can expect a lot of excitement.

Three trends profoundly shaped biological thinking during the past half century and will continue unabated into the next millennium. First, the hard work and ingenuity of many researchers led to an explosion in factual knowledge about biology. Advanced methods of molecular biology brought forth an unprecedented wealth of detailed information about the most intricate component parts of organisms, and this information is now begging to be evaluated and integrated. Ten thousand strings of genetic code can be analyzed simultaneously with modern chip and microarray technologies; about three million genes have already been sequenced; and the complete human, mouse, and rat genomes will be known within the next few years. Second, the ready availability of personal computers made data management and computation accessible to most scientists. Huge amounts of information can be stored and retrieved with an efficiency that was unbelievable just a decade ago. Where it took many long minutes of mainframe computing in the 1970s to integrate even three or four moderately simple differential equations, systems of dozens of equations can now be evaluated very quickly with off-the-shelf personal computers. The third important development occurred in the scientific community at large, when the realization set in that the paradigm of reductionism is reaching its limits. It was recognized that even the complete hierarchical reduction of organisms to their fundamental components cannot always answer questions about function and purpose. An appreciation emerged for a new need of reconstructing organismic systems from the bottom up. It became evident that this reconstruction is an essential prerequisite for understanding the rationale for natural designs, for predicting their responses under untested conditions, and, eventually, for producing new pathways of medical or biotechnological significance with some confidence and reliability. It was realized that this reconstruction is only possible with the new mathematics of nonlinear systems analysis and modeling. The integration of knowledge with mathematical and computational means became a new frontier in biology.

The tasks at this new frontier require thinking beyond linear chains of causes and effects: thinking in terms of integrated functional entities; thinking in systems, networks, models. Over the past thirty years, a mathematical framework has evolved that deals with these issues. It is based on a number of theorems and on rigorous analytical and numerical methods, yet it has the advantage of being accessible also to the scientist who does not have an advanced mathematical background. The framework is called *canonical modeling*. This uncommon designation reflects the fact that the approach is based on a small number of principles that apply irrespective of the subject area. Canonical models can be constructed with a minimal amount of information, and the design processes and evaluation procedures follow a handful of guidelines. Canonical models always have the same structure, and this has facilitated the development of specific, very efficient computer algorithms, which are now part of an entirely interactive and intuitive software package, PLAS, that does not require any programming skills.

Throughout the thirty-year history of canonical modeling, great effort was expended to document and publicize new developments in the peer-reviewed literature, in dissertations, and at conferences. In spite of these efforts and of the relative ease with which canonical modeling can be learned, the new approach was embraced only by a limited group of scientists, most of whom had significant mathematical training. At a meeting a few years ago, some of the canonical modelers pondered this discrepancy and tried to pinpoint the possible hurdles that kept a larger audience from experiencing the fascination of canonical modeling. The main reason emerging from this soul searching was the lack of an up-to-date introductory text that would facilitate access to the technical literature and correct the wrong perception that canonical modeling was only for the mathematical whiz. I was asked to spearhead the project of writing such a text and accepted the challenge. This is the result.

As agreed, the book is written with the biochemist and molecular biologist in mind. It does not assume mathematical expertise beyond high school, but it requires an open mind and a willingness to learn new mathematical concepts. For those feeling uneasy about the underlying math, a crash course on some background material is provided in the Appendix. Techniques germane to the canonical modeling approach are introduced step by step in the first seven chapters. These technical chapters are followed by four detailed case studies, most of which are excerpted from the literature, reformulated, adapted, and annotated. It is my strong belief that studying these cases in detail and re-executing some of the analyses conducted by experts in the field will greatly benefit the reader's understanding. Beyond general strategies, modeling is not always straightforward in its nuances and sometimes has an almost artistic flavor. It is therefore of great value to retrace step by step the rationale for setting up a model and analyzing it. It is said that Beethoven copied entire Mozart symphonies note by note, just to see how the master had worked the details.

Several chapters are augmented with *Complements*, which present material that expands the horizon of the chapter but is not essential for later chapters. This material, combined with numerous references to the pertinent literature, may well be useful for further independent study. Every chapter also contains a set of exercises,

which range from working numerical examples to thinking about a newly introduced concept and to related projects that are open-ended. In addition to the *Complements*, the *Epilogue* summarizes canonical modeling efforts that reach beyond biochemistry and molecular biology and may serve as an inspiration to branch out into other application areas.

Because modeling is a hands-on endeavor, the book comes with the computer program PLAS, which allows the reader to try out almost all of the concepts discussed in the book. PLAS is user-friendly and does not require any programming skills. The overhead of developing expertise in PLAS is minimal, and the learning curve has been steep for all who have worked with the program so far.

When I agreed to take on the book project, I did it under the condition of "communal help" from the leaders in canonical modeling. This help was solemnly pledged, and I have indeed received it generously throughout the process. Everyone involved in the early discussions contributed with published or unpublished material and was eager to help with input and feedback. It is very difficult to rank my appreciation to these individuals, so I will thank them in alphabetical order.

Marta Cascante was very helpful in the planning stage and discussed early versions of the material with me in great detail. She wrote the first draft on parameter estimation with methods developed by our colleagues in the related area of *metabolic control analysis*. Marta also provided the material for the last case study and for some of the smaller examples. Her contributions were of enormous help.

Raul Curto spent many long hours developing a series of models of purine metabolism. As is often the case, only the final model eventually appeared in the literature, thereby obscuring the hard work involved in the process of model development. I am very grateful to Raul for generously giving me all his background material and without hesitation allowing me to compose a case study from it.

António Ferreira, of course, is a main contributor to this project. His computer program PLAS may be the strongest part of it. It never ceases to amaze me with what fervor and patience António is willing to implement new ideas in PLAS or to transport the program to yet another platform. In addition to his computational efforts, António has given me uncounted comments and valuable feedback, especially for Chapter 4.

Rebecca Knapp gave me a lot of moral support and good suggestions for the subtitle.

Michael Savageau has always been and will always be "the one with the big picture." He has supported this project in multitudinous tangible and intangible ways. He impressed on me the importance of having the reader start with PLAS as early as possible and not following the standard procedure of doing all the theory first and finally getting to the fun (playing with PLAS) after many readers might already have put the book aside. He really liked my original 30-page introduction giving a historical perspective of biochemical endeavors beginning early B.C. but suggested throwing it out anyway in order to get going with the modeling. Only a good friend could have done that!

Albert Sorribas always reminded me that serious science and writing are not necessarily divorced from enjoying the process. I greatly appreciate his frank and

constructive criticism, his help with Chapter 7, and his good sense of humor in person and through the Internet.

Néstor Torres kept me on track, reminding me regularly that I had promised to help him with another project after this book was done. He expended great effort checking several versions of the last case study and re-computing the model with several sets of parameter values. I could not have done that chapter without him.

Zhen Zhang helped me uncounted times to outfox my computer when it didn't behave.

Several of my students helped me by proofreading various versions of various chapters, giving me good suggestions and fixing "unusual grammatical constructions" in my English. I would like to express my gratitude in particular to Keith Bangerter, Russell Goodman, and Nicole White for making insightful comments.

While they were not directly involved with this book project, I wish to thank Drs. Alan Gross (in memoriam), Clinton Miller, Philip Rust, and Peter Sands, who never failed to encourage me in my work on canonical models. Last but definitely not least I thank my family here in Charleston, in Germany, and in Michigan for their unconditional support.

The publication of a book constitutes the completion of a big project. I consider it also a beginning. It is my hope that the readers will include many newcomers who develop an interest in canonical models and biochemical systems analysis and who will give me honest feedback and new impulses. On occasion I have been criticized for being too verbose. That doesn't bother me much. I remember a friend who wrote a 50-page dissertation draft on a very complicated, obscure topic of abstract mathematical logic. His professor, himself a brilliant mathematician, pleaded with my friend *please* to put a few sentences between the theorems, proofs, and corollaries, indicating which direction the argument was going. Obligingly, my friend wrote up the same mathematical material in 150 pages. Pulling him aside at graduation, the professor congratulated him, saying, "Now that I understand your dissertation, I really like the short version." I hope that some of my audience, having read the book, will really like the concise technical papers that are plentiful in the literature.

Eberhard O. Voit
Charleston, SC
2000