Biochemical Systems Analysis

A Study of Function and Design in Molecular Biology

Michael A. Savageau

Addison-Wesley Publishing Company
Advanced Book Program

Biochemical Systems Analysis

A Study of Function and Design in Molecular Biology

Michael A. Savageau

The University of Michigan

With a Foreword by Robert Rosen

Dalhousie University

1976

Addison-Wesley Publishing Company

Advanced Book Program
Reading, Massachusetts

London • Amsterdam • Don Mills, Ontario • Sydney • Tokyo

Library of Congress Cataloging in Publication Data

Savageau, Michael A
Biochemical systems analysis.

Includes indexes.
1. Biological control systems. 2. Molecular biology.
I. Title.
QH508.S26 574.1'8 76-42469
ISBN 0-201-06738-2
ISBN 0-201-06739-0 pbk.

Printed in the United States of America

ABCDEFGHIJ-MA-79876

Foreword

Some fifteen years ago (in 1961), a conference entitled "Mathematical Problems in the Biological Sciences" was organized by the American Mathematical Society. A number of prominent biochemists participated in that conference; one of them, Dr. A. B. Pardee, concluded: "Eventually, we biochemists would hope to describe many of the essentials of biochemistry—the network of life—in mathematical terms. But I find it inconceivable that this mathematical treatment will be central to the thinking and language of biochemistry in the near future."

Dr. Savageau's book, which I believe represents a landmark contribution to biochemistry and to theoretical biology, provides ample evidence that the time has indeed come in which mathematical representation has become an integral part of biochemistry; a tool for the understanding of life processes no less important than the spectrophotometer or the ultracentrifuge. The familiar experimental tools of biochemistry and molecular biology reveal to us the properties of the individual molecules which enter into biochemical reactions. But only mathematical arguments based on these empirical data can allow us to infer the properties of complex reaction pathways, and most important, to understand how such pathways can be controlled, regulated, and integrated to give rise to the global behaviors characteristic of a living organism. It is this aspect with which Dr. Savageau's book is concerned, and which he terms biochemical systems analysis.

To a zeroth approximation, a functional cell can be regarded simply as a population of interacting chemical species. New species are produced at definite rates through chemical reaction, and pre-existing species are consumed. Such

populations possess definite regulatory properties governed by the familiar rate laws of chemistry, and much significant insight into the properties of cells can be obtained through the study of such systems. Especially when such a system is regarded as distributed in space, so that diffusion of chemical species enters into the picture, many remarkable dynamical properties emerge. There is a rich literature on reaction-diffusion systems, beginning with the earliest work of Rashevsky in the early 1930's, and culminating most recently in the mathematical analyses of René Thom.†

A better approximation, however, is obtained by regarding the cell not simply as a population of interacting chemical species, but rather one in which the chemical reactions which occur are each mediated through a specific catalyst, or enzyme. A suitable family of such enzymes converts an otherwise anarchic population of interacting chemical species into a reaction *network*. Such a reaction network is uniquely determined by the specificities of the family of enzymes, in the sense that the products of any enzyme in the system become substrates for other enzymes; this sequential relation is rigidly maintained through the enzyme specificities in the same way that the pattern of a neural net is maintained through axonal connectivities. In a certain sense, our initial population of chemical species is converted into a "society" by the presence of specific catalysts.

One of the great triumphs of classical biochemistry has been the elucidation of the specific pathways which comprise the reaction networks in cells. It was found that these pathways, or sub-networks, can take a variety of forms; they may be linear, branched, convergent, or cyclic. Again, using the basic rate laws of enzyme kinetics, the dynamic properties of such networks can be studied by mathematical means, and their properties assessed. Again, definite regulatory properties appear in such networks; these have been studied by numerous investigators, including Chance, Garfinkel, and Heinmets.

A still better approximation arises from the realization that at least some of the catalysts involved in biochemical reaction networks are not simply catalysts, but are also *control elements*. That is, the rates at which such catalysts operate can be modulated by the concentrations of chemical species which are neither substrate nor product of the catalyst, and which are involved either in the same pathway or in remote pathways in the overall reaction network. The presence of such control elements can be regarded as converting a reaction network into an "information-processing" network, and provides a crucial insight into the manner in which the behavior of specific reaction pathways, or sub-networks, can be integrated into a functional, adaptive biological system.

From an experimental point of view, the most deeply studied of such control elements are those involved in the control of unbranched biosynthetic pathways,

---

† René Thom, Structural Stability and Morphogenesis, 1975, W. A. Benjamin, Inc., Advanced Book Program, Reading, Mass.

either through feedback inhibition, or through forward activation. One striking aspect of such controlled pathways is that they exhibit the capability for *oscillation*. Such biochemical oscillators are important because they can provide the basis for cellular clocks which may control biological cycles and periodicities. Another striking feature of these controlled pathways is that the control is always exerted in a constantly recurring pattern; for instance, feedback inhibition in a biosynthetic pathway is always exerted through the first enzyme in the pathway. It is thus suggested that there is something particularly efficient about such modes of control, and it is important to find out what it is. As we shall see, Dr. Savageau considers such problems at great length in his book.

Finally, we must recognize that a cell is not simply a reaction network controlled by enzymes, but also involves a crucial genetic constituent. The genes may be regarded as determining, at any time, the kinds and amounts of enzymes which are present in the cell. The genes thus represent another kind of reaction network superimposed on the enzymatic one, and subject to similar kinds of control; each gene is itself a control element, whose activities at any time are modulated by the products of enzymes present in the cell, and hence ultimately by the activities of other genes. It is a remarkable fact that the properties of this genetic network can be represented in a fashion completely analogous to that appropriate for the enzymatically controlled reaction networks, so that for instance the repression of a gene and the inhibition of an enzyme can be treated in a formally identical fashion.

We thus come to view the cell as a collection of interacting biochemical pathways, integrated into an overall reaction network through both enzymatic control elements and genetic controls. Dr. Savageau's work is concerned with analyzing the dynamical properties of such networks. His point of departure is that of a biochemist and microbiologist, and his emphasis is always directed towards the properties of specific pathways in specific organisms. Indeed, the amount of experimental data he brings to bear in the construction and interpretation of his mathematical arguments is most impressive. To overcome the nonlinearities inherent in the analysis of reaction pathways, he develops a novel power-law approximation, which as he shows, possesses many advantages over the conventional procedure of linearization.

Dr. Savageau's treatment of the design of actual biochemical and genetic control circuits in real cells is a most interesting and instructive combination of analytic and synthetic biology. Given a particular biosynthetic pathway, the problem to be solved is that of the most "efficient" placement of control elements. The nature and properties of the enzymes involved belong to analytic biology; but the design problem involves the consideration of enzymatic pathways which *could* exist, and comparing their properties with the pathways which *do* exist. Basically, the problem is one of *optimal design*; to assess the respects in which the existing designs are (or are not) superior to the other members in the class of possible designs. Dr. Savageau's results in this direction have obvious

evolutionary implications bearing on the selection pressures responsible for cellular control mechanisms, and also represent one of the few cases where it seems possible to quantitatively specify the criteria through which such selection takes place.

In all, then, Dr. Savageau's work represents a fundamental contribution to the analysis and synthesis of biochemical control systems. An understanding of these control mechanisms lies at the very heart of the fundamental problems of biology and medicine, and in my view such understanding can only arise through the imposition of imaginative mathematical analysis on the empirical foundations. This is precisely what Dr. Savageau has done; I hope that his book will be widely read, and that it will have the influence it deserves.

ROBERT ROSEN

Killam Professor of
Physiology and Biophysics
Dalhousie University

Preface

Biochemical systems analysis is a newly emerging field that promises to be of considerable importance in the coming years. Its focus is the integrated functioning of the intact system, rather than the chemical and physical properties of the isolated molecular components that comprise the system. Historically there always has been a great deal of interest in the problems of integration and organization in biology. However, most ideas derived from the early work would have to be termed metaphorical since it was not until recent years that the underlying biochemical basis for such complex biological phenomena was revealed. In the short span of a few decades essentially all of the basic mechanisms and most of the individual enzymes that comprise a living cell such as the bacterium *Escherichia coli* have been defined. The reductionist approach of molecular biology has brought about this revolution. This is not to say that we know all there is to know about the cell, even at the molecular level, but it does indicate that activity in this area of exploration may have reached its zenith. Completion of the molecular inventory of the cell is in view, but still we have no real understanding of the problems of integration and organization that were the original stimuli for much of the biochemical work in the past 25 years. From this perspective the time is right for the development of a more synthetic approach to these biological problems. As a result of the molecular revolution a complementary synthetic approach has become not only possible, but necessary if we are to understand the more complex phenomena typical of higher levels of organization. It is the objective of this book to provide such an approach for the study of integrated biochemical phenomena.

Chapters 5 and 7 through 16 constitute the core of this book, which is for the most part the product of my own research. I begin with a thorough consideration of the molecular components of biochemical systems, particularly the mathematical properties of their kinetic descriptions. From this foundation a novel power-law method of analysis is developed that provides a unified formalism specifically appropriate for biochemical systems and yet mathematically tractable (Chapters 5, 7, and 8). The validity of this method is established both by arguments from first principles and by direct comparison with experimental results. In Chapters 9 through 16 the previously developed methods are applied to several important classes of biochemical and genetic control systems. The organization within each of these latter chapters is much the same. An introduction to a biological problem and its literature is given in sufficient detail to make it understandable to the non-specialist. Then a model is formulated and analyzed in detail. Finally, the conclusions of the analysis are interpreted and discussed in terms of their biological implications. The discussions are written to be understood by those who have not gone through the analyses in detail. In each of these chapters, the behavior of the intact system is stressed and the implications of its design are brought out. Much of this information is presented for the first time.

By the addition of Chapters 1, 2, 3, 4, and 6, and the inclusion of problems after each chapter, this book has been made reasonably self-contained and suitable for use as a text in an advanced undergraduate/beginning graduate course. I first used this material for such a course at Stanford University in the summer of 1969 and have given it annually at The University of Michigan since 1971. The prerequisites are mathematics through calculus and a course in modern biology. The students drawn to this course have had quite diverse backgrounds, including all the basic biological sciences, chemistry, physics, mathematics, engineering, and even a few clinical specialties. This reflects the interdisciplinary nature of the material as well as an interest in its theoretical and practical uses. Students with a good background in mathematics are able to omit Chapter 6 and skim rather quickly Chapters 7 and 8, but they usually spend more time on the first four chapters if they have had no previous exposure to enzyme kinetics. On the other hand, many students with a good background in basic biology are able to skim or omit the first four chapters and spend a little extra time on the mathematical techniques in Chapters 6, 7, and 8 if necessary. This permits the instructor to concentrate on the core chapters mentioned above.

I am indebted to many associates who undoubtedly influenced me in the course of numerous discussions. In this regard I would especially like to acknowledge John P. Steward of the Stanford University School of Medicine, in whose laboratory these ideas germinated many years ago, and Stephen Cooper, Rowland H. Davis, David I. Friedman, Elliot Juni, and Frederick C. Neidhardt, my colleagues for the past several years at the University of Michigan. I also am grateful

to the National Science Foundation for their support of my research and the University of Michigan Medical School for providing funds for use of the computer.

Throughout the writing of this book I was fortunate in having numerous and expert colleagues who contributed many useful suggestions. I particularly appreciate the efforts of Ole Maaløe of the University of Copenhagen and Robert Rosen then of the State University of New York at Buffalo, and now at Dalhousie University, who critically read the entire manuscript and the following individuals who read rather lengthy sections: Rowland H. Davis, now of the University of California at Irvine; Robert F. Goldberger, National Institutes of Health; Ethel N. Jackson, The University of Michigan; Sidney J. Strickland III, currently at Rockefeller University; and H. E. Umbarger, Purdue University.

Special thanks also are due to Barry Bochner, Naoto Sakamoto, Bonnie Templeton and the other students who studied, criticized and often improved the preliminary drafts upon which this book is based. Finally, I must acknowledge my first critic, Ann Savageau, and my last critic, Ann Marie Kotre, both of whom read numerous drafts and considerably improved the exposition.

MICHAEL A. SAVAGEAU

Contents

Chapter 1

Introduction

They [atoms] move in the void and catching each other up jostle together, and some recoil in any direction that may chance, and others become entangled with one another in various degrees according to the symmetry of their shapes and sizes and positions and order, and they remain together and thus the coming into being of composite things is effected.

– Simplicius

ORGANISMS AS BIOCHEMICAL SYSTEMS

Nothing so characterizes the living state as the ubiquitousness of chemical change–not chaotic flux, but integrated systematic change. Autotrophic organisms utilize solar energy and simple minerals to synthesize the complex organic molecules required for their maintenance and reproduction. Heterotrophic organisms degrade such complex molecules to simpler compounds and then utilize these compounds, as well as the energy released in the degradation process, for their biosynthetic needs. Despite extreme differences in structure and function among these major groups of organisms, there is a common chemical basis for all forms of life as we know it.

Although organisms must deal with many forms of energy, the basic unit of exchange is chemical. All other types of energy are interconvertible with the chemical form by means of specialized energy transduction processes. Electromagnetic energy is converted to chemical energy in photosynthesis, whereas chemical energy is converted to mechanical energy in muscular contraction. Aside from such specialized energy conversions, the overwhelming majority of cellular functions are of a strictly chemical nature.

The reactions performed in the cell are exceptionally difficult for the organic

chemist to reproduce, even under extreme conditions of temperature and acidity not available to the cell. This is because nearly all these reactions normally proceed by themselves at a very slow rate. However, organisms have numerous catalysts that increase the speed of chemical reactions. Thus, no matter what cellular function is considered, systematic chemical changes occur and these changes are catalyzed.

From these basic observations it can be concluded that the organism is a complex network of catalyzed reactions. For an indication of the complexity of these biochemical systems one need only examine a graphical summary of the known reactions in the intermediary metabolism of the cell (Fig. 1-1). Watson (1976) has estimated that one-sixth to one-third of all the reactions in the bacterium *Escherichia coli* are known, and thus that there are 3000 to 6000 different types of molecules present in this organism. The enormous complexity of the organism as a biochemical system is evident. Fortunately, as will be seen in later chapters, some basic simplicity in the general pattern of cellular metabolism is

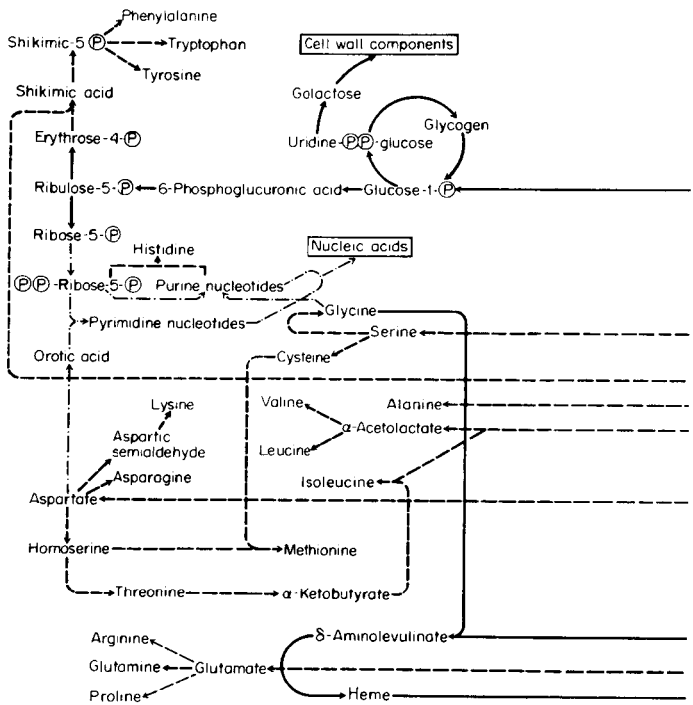


Fig. 1-1. See p. 3 for caption.

beginning to emerge, and certain other features of biological systems tend to limit the complexity of the biochemical network with which we must eventually deal. However, before the behavior of integrated biochemical systems can be considered, the basic components of these systems—i.e., the individual catalyzed reactions—must be understood.

ENZYMES AS PHYSICAL COMPONENTS

By definition a catalyst is an agent that increases the rate of a reaction without itself undergoing any net chemical change. Biological catalysts are called enzymes. In the past there was some question about the chemical composition of enzymes, but in all cases subsequently examined in depth they have proved to be composed primarily of protein. It is therefore of importance to examine the structure of proteins and determine how this relates to the properties of enzymes.

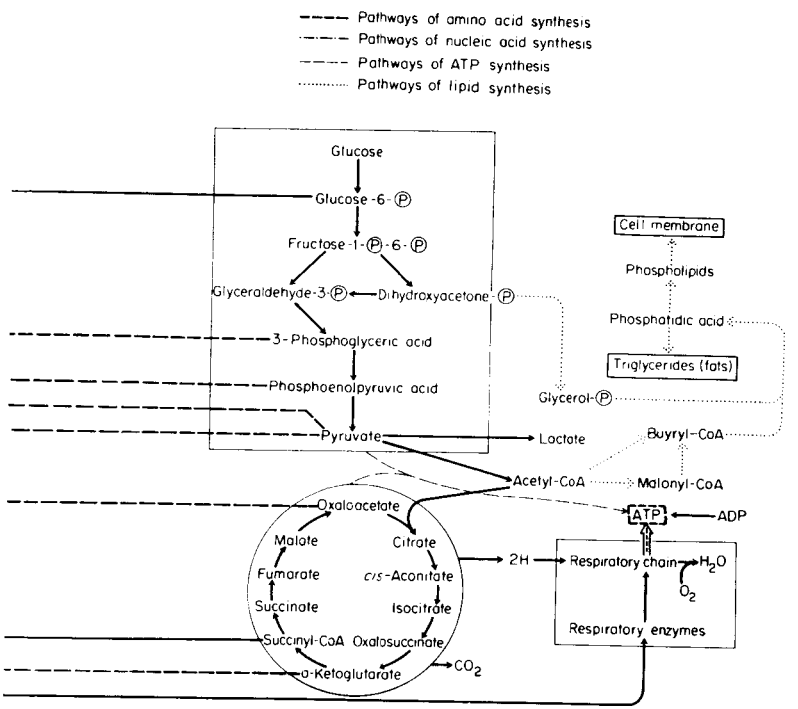
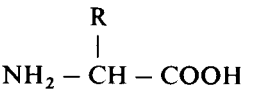


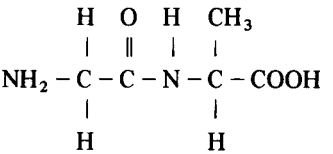
Fig. 1-1. The principal metabolic pathways in the bacterium *Escherichia coli.* Only selected intermediates and end products are shown. (From J. D. Watson, 1976, "Molecular Biology of the Gene," 3rd ed., Benjamin, Menlo Park, with permission.)

Proteins

Proteins are polymeric macromolecules. Their monomeric units are amino acids, of which there are about twenty different types. The natural amino acids, with one or two exceptions, consist of an amino group, a carboxyl group, and a specific side chain (Mahler and Cordes, 1971). Thus, their general structure is the following:



The carboxyl terminus of one amino acid is linked to the amino terminus of another by the removal of a water molecule to form the so-called peptide bond. Although the simplest polymer is a dipeptide–for example, of glycine and alanine



the proteins of enzymatic interest are much larger, containing as many as 1000 residues. Since each position in a linear chain of 1000 units can be occupied by any one of the twenty different amino acids, there are, mathematically, a total of 20^{1000} distinguishable sequences. This enormous variety of structural possibilities is thus able to account for the diversity of functions performed by cellular proteins.

Enzymatically active proteins are generally not linear but are folded into complex three-dimensional structures. The precise description of these structures is quite difficult. However, an adequate picture often can be given in terms of four hierarchical levels of organization called primary, secondary, tertiary, and quaternary structure.

Primary structure. The ordered sequence of amino acid residues in a polypeptide chain is defined as its primary structure. This sequence is ultimately specified, via the genetic code, by a corresponding sequence in the DNA of the organism. Changes in the genetic sequence, or a modification in the transcription-translation process, can result in a protein with an altered primary structure.

Determination of the amino acid sequence for a given protein, which is now almost routine, has been one of the major technological advances in the past decade. Sanger (1956) was the first to accomplish this feat by establishing the sequence for the peptide hormone insulin (Fig. 1-2). Since then numerous other