Basic Mathematics for Biochemists

Athel Cornish-Bowden

Basic Mathematics for Biochemists

ATHEL CORNISH-BOWDEN

Lecturer in Biochemistry,
University of Birmingham

LONDON NEW YORK
CHAPMAN AND HALL

First published 1981 by
Chapman and Hall Ltd
11 New Fetter Lane, London EC4P 4EE
Published in the USA by
Chapman and Hall
in association with Methuen, Inc.
733 Third Avenue, New York NY 10017

Printed in Great Britain at the
University Press, Cambridge

ISBN 0 412 23000 3 (cased)
ISBN 0 412 23010 0 (paperback)

British Library Cataloguing in Publication Data

Cornish-Bowden, Athel
Basic mathematics for biochemists.
1. Biological chemistry 2. Mathematics
I. Title
510'.2454 QD42

ISBN 0–412–23000–3
ISBN 0–412–23010–0 Pbk

To Joy and Jenny

Basic Mathematics for Biochemists

Preface

Some teachers of biochemistry think it positively beneficial for students to struggle with difficult mathematics. I do not number myself among these people, although I have derived much personal pleasure from the study of mathematics and from applying it to problems that interest me in biochemistry. On the contrary, I think that students choose courses in biochemistry out of interest in biochemistry and that they should not be encumbered with more mathematics than is absolutely required for a proper understanding of biochemistry. This of course includes physical chemistry, because a biochemist ignorant of physical chemistry is no biochemist. I have been guided by these beliefs in writing this book. I have laid heavy emphasis on those topics, such as the use of logarithms, that play an important role in biochemistry and often cause problems in teaching; I have ignored others, such as trigonometry, that one can manage without. The proper treatment of statistics has been more difficult to decide. Although it clearly plays an important part in all experimental sciences, it is usually preferable to treat it as a subject in its own right and not to try to incorporate it into a course of elementary mathematics. In this book, therefore, I have used a few examples from statistics to illustrate more general points, but I have not discussed it for its own sake. To summarize, the book is directed primarily towards students taking compulsory courses in mathematics in the early stages of their training as biochemists, but I hope it will also prove useful as a short revision text at later stages in the study of biochemistry.

I should like to thank my wife Mary Ann for her encouragement and support during the writing of this book, and for pointing out various ways in which it could be made more comprehensible. I am also grateful to my colleagues Geoffrey Bray, Stuart Ferguson, Baz Jackson, John Teale and Chris Wharton for reading the manuscript and suggesting many improvements to it.

January, 1981 Athel Cornish-Bowden

Contents

1 The Language of Mathematics

1.1 Introduction

Lord Kelvin once remarked that all of science could be divided into physics and stamp collecting. Although this rather patronizing comment of a physicist is not one that will appeal to all biochemists, it has an element of truth in it. A science can hardly claim to be a science as long as it remains no more than a catalogue of unrelated observations. Only when general laws can be proposed and tested by experiment can it be said to have passed from mere description into science. In chemistry the transformation from stamp collecting into science corresponded with the development of thermodynamics and, later, the atomic theory and theories of chemical bonding; in biochemistry, the gradual realization that understanding of life processes requires a foundation of physical chemistry and not just a list of metabolic reactions has played a corresponding role. It is no coincidence that mathematics has been central in all of these developments, and it is now almost impossible to comprehend even elementary biochemistry without a grasp of elementary mathematics. Fortunately for non-mathematically minded biochemists, however, the mathematics necessary for an undergraduate course in biochemistry *is* nearly all elementary, and nearly all of it has been touched on in every science student's previous education. Little more is required, therefore, than to identify the parts of elementary mathematics that are important in biochemistry and to reinforce them with appropriate examples.

In mathematics itself, the transformation from description into science is paralleled by the development from *arithmetic*, which is concerned with numbers and their manipulation, into *algebra*. Arithmetic is very useful, but it is much too limited to satisfy all of the needs of science. In arithmetic, every problem is a new and separate problem, and it is difficult to make useful generalizations and hence to express scientific laws. Let us consider the simple biochemical example in Table 1.1, which shows a set of rates of a reaction measured at the substrate concentrations given. As it stands the table is no more than a *description* of the results of a particular set of measurements and as long as we treat the numbers just as numbers we cannot make a general or useful statement about the enzyme to which they refer. The table tells us the rates observed at substrate concentrations of 5 mM and 10 mM but offers no guidance about what rate to expect at 8 mM; it does not tell us whether the system studied was behaving in accordance with some general law; it offers no clue as to what general law there might be. To remedy these omissions we must move beyond arithmetic into algebra, because only then will we be able to recognize a pattern or regularity in the numbers, and express it so that it can be recognized again if it occurs with another system. If the rates in Table 1.1 are represented as v and the substrate concentrations as s, then the following equation expresses a *law* that defines all of the numbers in the table:

$$v = \frac{10s}{4+s}$$

This equation has two advantages over the table: first, it allows a *summary* of all of the information while occupying much less space; secondly, it *predicts* what v values we might expect to observe at s values that are not included in the table. For example, it answers the question above by predicting that $v = 6.67\ \mu\mathrm{M\,s^{-1}}$ when $s = 8$ mM. This is clearly more useful than just listing a set of numbers recorded on a particular occasion.

One can proceed one stage further with this example by noting that the equation is typical of what is reported for many enzymes, and so if

Table 1.1 Observations from a kinetic experiment

Substrate concentration (mM)	*Rate* ($\mu\mathrm{M\,s^{-1}}$)
1	2.00
2	3.33
5	5.56
10	7.14
20	8.33

we replace the numbers 10 ($\mu M\ s^{-1}$) and 4 (mM) by V and K_m, respectively, we have an equation that expresses a *generalization* about enzymes:

$$v = \frac{Vs}{K_m + s}$$

Again, replacing numbers with *symbols* has increased the generality of what we want to say. In addition to the symbols v, V, s and K_m, which represent numbers, either particular ones or generalized ones, the equation contains three *operators*: one is represented by the addition sign +; one is shown by the horizontal line between Vs and $K_m + s$; and the third is implied by the juxtaposition of V and s, but could have been made explicit by writing $V \cdot s$ instead of Vs. Each operator specifies something to be done to the numbers or symbols operated on: the + sign requires K_m and s to be added together; the horizontal line requires Vs to be divided by $K_m + s$; the juxtaposition of V and s (or a dot between them) requires them to be multiplied together.

As long as no ambiguity is possible mere juxtaposition is sufficient to indicate multiplication, but if several numbers are to be multiplied together, or if we allow algebraic symbols consisting of more than one letter each (as in Fortran and many other computer languages), or if ambiguity is possible for some other reason, multiplication can be indicated by a dot or a cross. The dot is more common for pairs of symbols (where no confusion with the decimal point is possible) and the cross is more common for pairs of numbers, but the two symbols have the same meaning in most contexts, i.e. $V \times s$ means the same as $V \cdot s$. (In some specialized applications, such as in *vector algebra*, it is convenient to assign distinct meanings to · and ×, but these need not concern us in elementary biochemistry). In current usage the dot should be written *above* the line and the decimal point *on* the line, e.g. 5.1 · 8.7 = 44.37 *not* 5·5 . 8·7 = 44·37 but this is rather a recent convention so far as British books are concerned: in older British (but not American) work one is likely to encounter exactly the opposite convention.

1.2 Priority rules for operators

To avoid ambiguity it is important to realize that operators have to be obeyed in a proper order. Unlike ordinary language, equations are not read from left to right but in accordance with *priority rules* that require certain operators to be obeyed before others. Thus, the value of $5 \times 3 + 2 \times 4 - 3$ is 20, not 25 or 65, because multiplication must be done before addition or subtraction. In general, the rule is as follows:

(1) expressions within brackets must be evaluated first;

(2) if brackets are 'nested' (brackets within brackets), 'inner' brackets must be evaluated before 'outer';

(3) exponentiation must be carried out before multiplication and division;

(4) multiplication and division must be carried out before addition and subtraction.

'Exponentiation' is the raising of a number or expression to a power, as in x^2, $(x + y)^a$, etc. If a power is itself raised to a power we work down from the top, i.e.

$$\mathrm{e}^{-2x^2} \text{ means } \mathrm{e}^{(-2x^2)} \textit{ not } (\mathrm{e}^{-2x})^2$$

As in this example, it is always permissible and often desirable to use brackets to clarify an expression that might otherwise be misinterpreted. This is true even if the expression without brackets is strictly unambiguous.

There are no rules of priority between addition and subtraction among themselves, because the result of a sequence of additions and subtractions is independent of the order in which they are done. This is normally a matter of convenience only, although occasionally numerical considerations may make one order better than another. In principle, the same applies to multiplication and division, but greater care is needed because thoughtless use of the slash / to indicate division often results in expressions with meanings that are either unclear or clear but different from what the writer intended. It is wisest therefore to use the slash in moderation and to check carefully that expressions have the meanings intended. Consider for example the following equation:

$$v = \frac{Vs}{K_\mathrm{m}(1 + i/K_\mathrm{i}) + s}$$

This is unambiguous, and the priority rule should prevent the bracketed expression $(1 + i/K_\mathrm{i})$ from being misread as $[(1 + i)/K_\mathrm{i}]$. When there is more than a single term after the slash, however, as in $(i/K_\mathrm{i} + 1)$, misunderstanding is more likely because it is not always clear whether the slash indicates simple division or whether it is used to avoid the typographical inconvenience of a cumbersome fraction such as $\frac{i}{K_i + 1}$. Double slashes are so confusing that they should *never be used*: this applies not only to ordinary algebraic expressions but also the units of physical quantities, as in $R = 8.314$ J/mol/K. Here it is not clear whether the K belongs in the numerator of the unit with the J or in the denominator with the mol. To avoid this uncertainty the definition should be written as follows: $R = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$. For reasons that will become clear in Chapter 2, mol^{-1} and K^{-1} have the meanings (1/mol) and (1/K), respectively. In general, slashes should only be used in expressing units when there is only a single term in the denominator.

Certain computer languages do not obey precisely the same priority rules as conventional mathematics. This fact has generated rather more confusion about the conventions than existed before computers became widespread, and the appearance of cheap electronic calculators has made matters much worse in this regard because those that use so-called 'algebraic notation' commonly ignore mathematical conventions altogether and use a 'left-to-right' system. An expression such as

$$\text{A * A/B/C/D + A/B * C}$$

would be unambiguous in a Fortran program and would have the meaning

$$\frac{AA}{BCD} + \frac{AC}{B}$$

(The multiplication sign in Fortran is expressed as * and must be explicit). This unambiguous meaning, which may nonetheless be different from what the programmer intended, does *not* imply that such expressions are acceptable in ordinary mathematics. Similarly, the fact that simple calculators often disregard priority rules does not mean that they are obsolete. For example, the expression $5 \times 3 + 2 \times 4 - 3$ must be interpreted as $(5 \times 3) + (2 \times 4) - 3$ if the proper conventions are followed, even though most simple calculators using so-called 'algebraic notation' execute instructions as they are entered and consequently interpret the above expression as $\{[(5 \times 3) + 2] \times 4\} - 3 = 65$ (this is indeed what I get if I key in $5 \times 3 + 2 \times 4 - 3 =$ on my pocket calculator).

1.3 The summation sign

It often happens, especially in statistical calculations, that we need to add together a large number of terms of the same kind. For example, if we have a set of values $x_1, x_2, x_3 \ldots x_n$, their arithmetic mean $\bar{x}$ is given by

$$\bar{x} = (x_1 + x_2 + x_3 + \ldots + x_n)/n$$

In more complex examples explicit representation of the summation becomes cumbersome and unnecessary, and it is more convenient to use a special operator called the *summation sign* $\sum$ (a capital Greek sigma) instead:

$$\sum_{i=1}^{n} x_i \equiv x_1 + x_2 + x_3 + \ldots + x_n$$

The *limits* $i = 1$ and n written above and below the sign mean 'start adding at $i = 1$ and continue until $i = n$'. If the limits are obvious, as for example in statistical calculations where one often has to sum over all of a set of observations, they can be omitted.

Although the summation sign is a considerable convenience when the underlying calculation is understood, it can sometimes obscure the meaning when it is not. Indeed, one of the main reasons why more advanced mathematics than one is familiar with can appear much more difficult than it actually is, is that it often uses special notation to express results more compactly. For example, the whole of *matrix algebra* is a way of expressing very complicated relationships in an extremely compact way: very convenient when one is familiar with the symbolism but baffling when one is not. Whenever obscurity threatens for this sort of reason it is often helpful to translate the compact expressions into a more long-winded form, and then their meanings are likely to become clearer.

A corresponding sign for *products* also exists, although it is much less often encountered than the summation sign. It is written as the Greek capital pi, $\prod$, and is used in an exactly analogous way, i.e.

$$\prod_{i=1}^{n} x_i \equiv x_1 x_2 x_3 \ldots x_n$$

1.4 Functions

A mathematical *function* can be regarded as a set of instructions to carry out a series of operations on a variable or set of variables. For example, if we define v in terms of s as

$$v = \frac{Vs}{K_m + s}$$

where V and K_m are constants, then we are defining v as a *function* of s, by defining what operations have to be carried out on s to obtain v. We can also have functions of more than one variable. For example, v may be determined not solely by a single concentraion s but may depend both on s and on another concentration i:

$$v = \frac{Vs}{K_m(1 + i/K_i) + s}$$

and now we say that v is a function of both s and i.

Sometimes we may wish to symbolize the existence of a dependence of one variable on another without specifying what the dependence is. We then often use the symbol $f(\;)$ or something similar, e.g.

$$v = f(s, i)$$

which states that v depends on s and i but does not indicate whether the dependence follows the equation given above or some other equation.

There are a number of functions that are so often required in mathematics that they are given special symbols. For example, if y is always given by multiplying x by itself we say that y is the *square* of x and symbolize the relationship as

$$y = x^2$$

where the superscript 2 indicates that 2 x's need to be multiplied together. Conversely, x in this example is the *square root* of y, which we may write as

$$x = \sqrt{y}$$

or, more commonly and for reasons that I shall discuss in Chapter 2, we may express the same relationship as

$$x = y^{\frac{1}{2}}$$

Other functions of great importance are the *logarithmic* and *exponential* functions, which I shall also consider in Chapter 2, the *derivative*, or result of differentiation (Chapter 3), and the *integral* (Chapter 4).

There are others, such as *trigonometric functions*, that are important in mathematics generally, but have little application in elementary biochemistry, and so I shall say little about them. On the other hand, there are certain functions that have little importance in mathematics as a whole but which are useful to define for biochemical purposes. For example, various properties of proteins can be related to the hydrogen-ion concentration $[H^+]$ in terms of the following kind of expression:

$$y = \frac{\tilde{y}}{1 + ([H^+]/K_1) + (K_2/[H^+])}$$

in which $\tilde{y}$, K_1 and K_2 are constants. This kind of function was first studied by Michaelis and it is consequently called a *Michaelis function.*

1.5 Constants, variables and parameters

Some quantities, such as the number 2.0, have a unique value under all circumstances and are called *constants*. Other numbers, such as the *gas constant* $R \simeq 8.3\,\mathrm{J\,mol^{-1}\,K^{-1}}$ are found by experiment to be constant also. Others, such as K_m in the Michaelis–Menten equation,

$$v = Vs/(K_m + s)$$

may be constant for a particular enzyme and substrate under well-defined and constant conditions, although they may vary with temperature, pH, etc. These quantities can be treated mathematically as constants only as long as the physical conditions that determine them are constant.

We are often interested in quantities that change when the conditions change. For example, we may find that an equilibrium 'constant' K varies with the temperature T according to the van't Hoff equation:

$$K = \exp\left(\frac{\Delta S^0}{R} - \frac{\Delta H^0}{RT}\right)$$

where exp( ) is the exponential function (Chapter 2), and ΔS^0, ΔH^0 and R are constants. Thus although K may be constant at constant T it varies with T and so if we are concerned with changes in temperature we must treat K as a *variable*.