Stoicheiometric analysis in studies of metabolism

Athel Cornish-Bowden*[1], Jan-Hendrik S. Hofmeyr† and María Luz Cárdenas*
*Bioénergétique et Ingénierie des Protéines, Institut Fédératif "Biologie Structurale et Microbiologie", Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France, and †Department of Biochemistry, University of Stellenbosch, Stellenbosch 7600, South Africa

Abstract

Stoicheiometric analysis of metabolic pathways provides a systematic way of determining which metabolite concentrations are subject to constraints, information that may otherwise be very difficult to recognize in a large branched pathway. The procedure involves representing the pathway structure in the form of a matrix and then carrying out row operations to convert the matrix into 'row echelon form': this is a form in which as many as possible of the elements on the main diagonal are non-zero, and all of the elements below the main diagonal are zero. If exactly the same operations are carried out on a unit matrix of order equal to the number of intermediate metabolites in the pathway, the resulting matrix allows the stoicheiometric constraints to be read off directly.

Introduction

Since its introduction by Reder [1], stoicheiometric analysis has become an essential tool for analysing metabolic systems prior to modelling their behaviour in the computer. It is incorporated into modern metabolic modelling programs such as Scamp [2], Gepasi [3,4] and Jarnac [5]. The sort of problems that could not easily be analysed without knowledge of stoicheiometric constraints are exemplified by various general applications in biotechnology (e.g. [6–8]) and by powerful new methods of pathway analysis [9] that are helping to rationalize the results from genome sequencing. Nonetheless, the intensely algebraic nature of stoicheiometric analysis, which makes it barely feasible to attempt in other than a matrix representation, has caused it to remain apparently obscure and difficult, so that one may wonder whether the information that it provides is sufficiently useful to justify the effort needed to understand it. Here we illustrate how to detect a stoicheiometric constraint that could easily be overlooked in a casual inspection, in the context of as simple a model as we can devise without making it completely trivial and transparent.

Key words: computer analysis of metabolism, metabolic modelling, pathway analysis, *Trypanosoma brucei*.

[1]To whom correspondence should be addressed (e-mail athel@ibsm.cnrs-mrs.fr).

Metabolic model

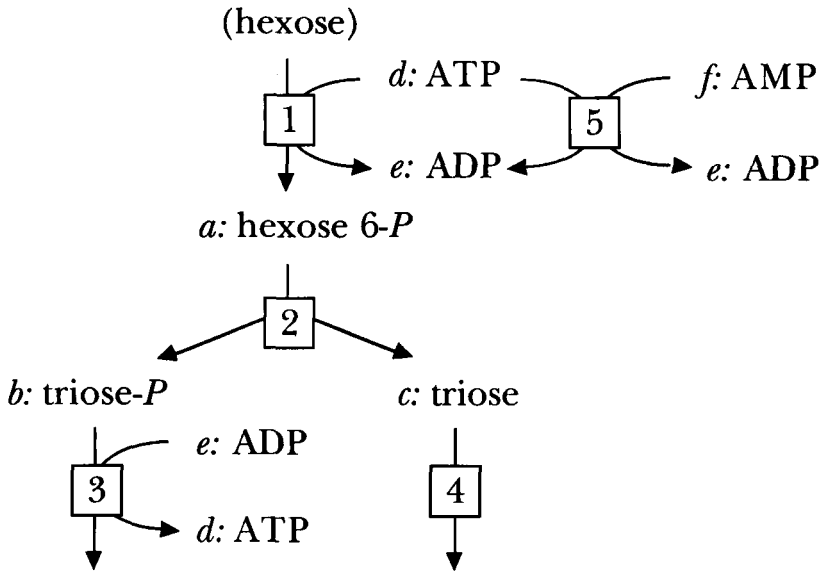
The model to be analysed is shown in Figure 1. It is based on a model of glycolysis in *Trypanosoma brucei* that was found to obey a complicated stoicheiometric constraint [10,11] that was unexpected and difficult to rationalize even after it had been recognized. For the present purpose we have simplified the original model by eliminating all characteristics that do not appear necessary for a non-intuitive stoicheiometric constraint to apply. The reservoir of hexose that acts as starting material for reaction 1 is shown in parentheses because its concentration is treated as being fixed independently of the metabolic system and does not enter into the stoicheiometric analysis, and the sinks that act as products of reactions 3 and 4 are left unspecified for the same reason (though the product of reaction 3 may be assumed to be different from the substrate of reaction 4).

For such analysis the model must be represented as a stoicheiometric matrix, as follows, where each column represents a reaction and each row initially represents the rate of change of an intermediate concentration:

Figure 1

Metabolic scheme to illustrate stoicheiometric analysis

The starting material hexose is shown in parentheses because its concentration is treated as fixed and it does not enter into the stoicheiometric analysis, and the products of reactions 3 and 4 are left anonymous for the same reason. The labels *a* to *f* shown for the metabolites are used in the text as compact symbols for the corresponding concentrations.



$$\begin{array}{l} \frac{\mathrm{d[hexose\text{-}6}P]}{\mathrm{d}t} \\ \frac{\mathrm{d[triose\text{-}}P]}{\mathrm{d}t} \\ \frac{\mathrm{d[triose]}}{\mathrm{d}t} \\ \frac{\mathrm{d[ATP]}}{\mathrm{d}t} \\ \frac{\mathrm{d[ADP]}}{\mathrm{d}t} \\ \frac{\mathrm{d[AMP]}}{\mathrm{d}t} \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & v_5 \end{array} \\ \left[\begin{array}{ccccc} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right] \end{array} \qquad (1)$$

Just by examining Figure 1 one can immediately guess that the sum of adenine nucleotide concentrations must be conserved in this system, but it is less obvious that there is a second stoicheiometric constraint.

Stoicheiometric analysis

To analyse matrix (1) we need to apply Gaussian elimination to it. This consists of a series of row operations (addition, subtraction or exchange of rows) designed to make the main diagonal consist, as far as possible, of non-zero elements and the triangle below this main diagonal to consist entirely of zero elements. In the process the rows lose their one-to-one identification with particular metabolites, so we need to remove the row labels from matrix (1), and we need to adopt a different labelling system that allows us to keep track of the metabolites. The columns, however, retain their one-to-one identification with the five reactions, so the column labels can remain. The best solution is to augment the 5×6 matrix by writing a 6×6 unit matrix next to it and labelling the six new columns of the resulting 11×6 matrix with symbols for the rates of change of the metabolite concentrations. For compactness we use the one-letter symbols *a* to *f* defined in Figure 1 for these concentrations, and *t* for time:

$$\left[\begin{array}{ccccc:cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & \frac{\mathrm{d}a}{\mathrm{d}t} & \frac{\mathrm{d}b}{\mathrm{d}t} & \frac{\mathrm{d}c}{\mathrm{d}t} & \frac{\mathrm{d}d}{\mathrm{d}t} & \frac{\mathrm{d}e}{\mathrm{d}t} & \frac{\mathrm{d}f}{\mathrm{d}t} \\ 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \qquad (2)$$

We now eliminate non-zero elements below the main diagonal, starting with column 1. The element of -1 in row 4 can be eliminated by replacing row 4 with the sum of the original rows 1 and 4:

$$\left[\begin{array}{ccccc:cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & \frac{\mathrm{d}a}{\mathrm{d}t} & \frac{\mathrm{d}b}{\mathrm{d}t} & \frac{\mathrm{d}c}{\mathrm{d}t} & \frac{\mathrm{d}d}{\mathrm{d}t} & \frac{\mathrm{d}e}{\mathrm{d}t} & \frac{\mathrm{d}f}{\mathrm{d}t} \\ 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \qquad (3)$$

The element of 1 in row 5 can be similarly eliminated by replacing row 5 with the difference between the original rows 1 and 5:

v_1	v_2	v_3	v_4	v_5	$\frac{da}{dt}$	$\frac{db}{dt}$	$\frac{dc}{dt}$	$\frac{dd}{dt}$	$\frac{de}{dt}$	$\frac{df}{dt}$
1	−1	0	0	0	1	0	0	0	0	0
0	1	−1	0	0	0	1	0	0	0	0
0	1	0	−1	0	0	0	1	0	0	0
0	−1	1	0	−1	1	0	0	1	0	0
0	1	−1	0	2	−1	0	0	0	1	0
0	0	0	0	−1	0	0	0	0	0	1

(4)

Exactly similar manipulations now allow the non-zero elements in rows 3, 4 and 5 in column 2 to be eliminated by subtracting row 2 from or adding row 2 to these other rows:

v_1	v_2	v_3	v_4	v_5	$\frac{da}{dt}$	$\frac{db}{dt}$	$\frac{dc}{dt}$	$\frac{dd}{dt}$	$\frac{de}{dt}$	$\frac{df}{dt}$
1	−1	0	0	0	1	0	0	0	0	0
0	1	−1	0	0	0	1	0	0	0	0
0	0	1	−1	0	0	−1	1	0	0	0
0	0	0	0	−1	1	1	0	1	0	0
0	0	0	0	2	−1	−1	0	0	1	0
0	0	0	0	−1	0	0	0	0	0	1

(5)

Column 3 is already in the required state, with no non-zero elements below the main diagonal, so there are no further row operations involving row 3. However, matrix (5) contains a zero in column 4 on the main diagonal, but there are no row operations involving only rows 4 to 6 that could eliminate it, so we accept the zero and proceed to column 5, which can be dealt with as before, by adding row 4 twice to row 5, and subtracting it once from row 6:

v_1	v_2	v_3	v_4	v_5	$\frac{da}{dt}$	$\frac{db}{dt}$	$\frac{dc}{dt}$	$\frac{dd}{dt}$	$\frac{de}{dt}$	$\frac{df}{dt}$
1	−1	0	0	0	1	0	0	0	0	0
0	1	−1	0	0	0	1	0	0	0	0
0	0	1	−1	0	0	−1	1	0	0	0
0	0	0	0	−1	1	1	0	1	0	0
0	0	0	0	0	1	1	0	2	1	0
0	0	0	0	0	−1	−1	0	−1	0	1

(6)

The Gaussian elimination can now be regarded as finished, as the original objective is satisfied: the main diagonal consists, as far as possible, of non-zero elements, and only zero elements are found below it. However, row 6 is left in a more complicated state than necessary, and can be simplified by adding row 5 to it. Note that as this is equivalent to adding zero to zero in the left-hand part of the matrix it has no effect on the correctness of the result:

v_1	v_2	v_3	v_4	v_5	$\frac{da}{dt}$	$\frac{db}{dt}$	$\frac{dc}{dt}$	$\frac{dd}{dt}$	$\frac{de}{dt}$	$\frac{df}{dt}$
1	−1	0	0	0	1	0	0	0	0	0
0	1	−1	0	0	0	1	0	0	0	0
0	0	1	−1	0	0	−1	1	0	0	0
0	0	0	0	−1	1	1	0	1	0	0
0	0	0	0	0	1	1	0	2	1	0
0	0	0	0	0	0	0	0	1	1	1

(7)

The two blocks of zeroes in rows 5 and 6 of the left-hand part of this final matrix now represent the two stoicheiometric constraints that characterize the network in Figure 1. As row 6 is simpler and represents an intuitively obvious result, we consider it first. It means that the column labels multiplied by the coefficients in row 6 below them add up to zero:

$$\frac{dd}{dt}+\frac{de}{dt}+\frac{df}{dt}=0 \qquad (8)$$

This can be written in a more familiar way by replacing the one-letter column symbols with more recognizable metabolite symbols, and recalling that a zero rate implies a constant concentration:

$$[\text{ATP}]+[\text{ADP}]+[\text{AMP}]=\text{constant} \qquad (9)$$

This equation just expresses the relationship that anyone with basic biochemical knowledge would guess without doing the analysis. The second may be obtained from row 5 of matrix (7) in the same way, but is less obvious:

$$[\text{hexose-6}P]+[\text{triose-}P]+2[\text{ATP}]+[\text{ADP}]=\text{constant} \qquad (10)$$

It represents the sum of all the transferable phospho groups in the system. Once its meaning is pointed out it is easy to confirm by inspecting Figure 1 that the relationship expressed by eqn (10) must hold, as there is transfer of phosphate between the system and its exterior, but this is because we have used as simple a case as possible to illustrate the method. With a more complicated example, such as the model for *T. brucei* that formed the starting point for this discussion, a stoicheiometric constraint can remain obscure even after it is known. In that example the equation that corresponded to eqn (10) involved most of the transferable phospho groups in the glycosomal compartment of the cell, but not all of them, and also involved two extraglycosomal metabolites [10,11].

Discussion

The above example illustrates how stoicheiometric analysis can provide a way of deducing constraints that are not obvious from inspection of a metabolic pathway. However, it does not explain the utility of the resulting knowledge. What is the practical value of knowing that trypanosomal glycolysis conserves certain phospho groups, for example? One answer is that this is a necessary piece of information for computer modelling of the pathway, but as modern programs [2–5] calculate the stoicheiometric constraints automatically they effectively absolve the user from needing to know about them. Away from a modelling context, the major importance of such constraints is that they have important implications for the design of pesticides and pharmacological agents. Such agents normally act by provoking changes in either metabolic fluxes or metabolite concentrations, and both of these are affected by stoicheiometric constraints. Varying fluxes in the cell is rendered much more difficult than varying enzyme rates *in vitro* by the fact that metabolite concentrations, unlike substrate, product and effector concentrations *in vitro*, are not fixed by the experimenter but can respond to external inhibitors in such a way as to restore the steady-state flux to a value close to what it would have in the absence of the inhibitor. However, a metabolite whose concentration is tightly constrained by stoicheiometric relationships may behave rather like a substrate *in vitro*, with the result that an enzyme that uses it as substrate may show a greater than average response to an external inhibitor. Conversely, the same enzyme is likely to prove a poor target for an inhibitor that is intended to provoke a large change in the concentration of its substrate.

References

1 Reder, C. (1988) J. Theor. Biol. **135**, 175–201
2 Sauro, H. M. (1993) Comput. Appl. Biosci. **9**, 441–450
3 Mendes, P. (1993) Comput. Appl. Biosci. **9**, 563–571
4 Mendes, P. (1997) Trends Biochem. Sci. **22**, 361–363
5 Sauro, H. M. (2000) in Animating the Cellular Map (Hofmeyr, J.-H. S., Rohwer, J. M. and Snoep, J. L., eds.), pp. 221–228, Stellenbosch University Press, Stellenbosch
6 Liao, J. C. (2000) in Technological and Medical Implications of Metabolic Control Analysis (Cornish-Bowden, A. and Cárdenas, M. L., eds), pp. 49–56, Kluwer Academic Publishers, Dordrecht
7 Rohwer, J. M. and Hofmeyr, J.-H. S. (2000) in Technological and Medical Implications of Metabolic Control Analysis (Cornish-Bowden, A. and Cárdenas, M. L., eds), pp. 73–79, Kluwer Academic Publishers, Dordrecht
8 Rohwer, J. M. and Botha, F. C. (2001) Biochem. J. **358**, 437–445
9 Schuster, S., Fell, D. and Dandekar, T. (2000) Nat. Biotechnol. **18**, 326–332
10 Bakker, B. M., Michels, P. A. M., Opperdoes, F. R. and Westerhoff, H. V. (1997) J. Biol. Chem. **272**, 3207–3215
11 Eisenthal, R. and Cornish-Bowden, A. (1998) J. Biol. Chem. **273**, 5500–5505

Received 19 November 2001