A computer program for the algebraic determination of control coefficients in Metabolic Control Analysis

Simon THOMAS and David A. FELL

School of Biological and Molecular Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, U.K.

A computer program (MetaCon) is described for the evaluation of flux control, concentration control and branch-point distribution control coefficients of a metabolic pathway. Requiring only the reaction scheme as input, the program produces algebraic expressions for the control coefficients in terms of elasticity coefficients, metabolite concentrations and pathway fluxes. Any of these variables can be substituted by numeric or simple algebraic expressions; the expressions will then be automatically rearranged in terms of the remaining unknown variables. When all variables have been substituted, numeric values will be

obtained for the control coefficients. The program is a computerized implementation of the matrix method for the determination of control coefficients. The features of MetaCon are compared with those of other programs available to workers in Metabolic Control Analysis. Potential benefits of, and methods of using, MetaCon are discussed. The mathematical background and validity of the matrix method rules are discussed, and the algorithm used by MetaCon is described. The matrix method is shown to be a specific case of a previously described general formalism for calculating control coefficients.

INTRODUCTION

In the study of metabolic regulation, the use of Metabolic Control Analysis is becoming more and more widespread (Fell, 1992; Cornish-Bowden and Cárdenas, 1990). Metabolic Control Analysis (Kacser and Burns, 1973; Heinrich and Rapoport, 1974) is a methodology by which the control properties of a metabolic system at steady state can be related to the properties of its component parts. Central to the theory are the control coefficients, which quantify the effect of a change in a system parameter on the value of a system variable. Evaluation or estimation of control coefficients is the usual aim of carrying out a control analysis on a system. The parameter of interest is typically an amount of enzyme, the variable usually a metabolic flux or the concentration of a metabolite.

Despite its growing popularity, application of Metabolic Control Analysis to anything other than a trivial system is both time-consuming and error-prone. A number of approaches have been described to reduce the effort and the possibility of error which is inherent in carrying out an analysis of a realistic system. These approaches include diagrammatic (e.g. Hofmeyr, 1989), graph-theoretical (Sen, 1990), and matrix-based methods (e.g. Westerhoff and Chen, 1984; Fell and Sauro, 1985; Westerhoff and Kell, 1987; Sauro et al., 1987; Cascante et al., 1989a,b; Reder 1988). Control Analysis is also an obvious area of application of computers, but until now there has been very little available to help with the application of Control Analysis to experimental systems. Most computer programs currently available which can carry out Metabolic Control Analysis (e.g. Sauro and Fell, 1991; Cornish-Bowden and Hofmeyr, 1991) are primarily numerical simulators, and are therefore of more interest to mathematical modellers than to experimentalists. Letellier et al. (1991) and Schulz (1991) have written programs which are of more use to the experimentalist. MetaCon, the program described in the present paper, fully automates all stages of carrying out a control analysis to give values for the control coefficients. All calculations are carried out algebraically rather than numerically, so that if the values of one or more variables are unknown, the output will be polynomials in terms of the unknown(s). For this reason, MetaCon should be of interest to the theoretician as well

as to the experimentalist, as a tool for investigating the effect of any variable(s) on the values of the control coefficients.

THE MATRIX METHOD OF CONTROL ANALYSIS

MetaCon is, in essence, an automation of the matrix method (Fell and Sauro, 1985; Sauro et al., 1987; Small and Fell, 1989). This is a set of rules which can in principle be applied to a pathway of any complexity to determine the flux control, concentration control and branch-point distribution control coefficients (the last are defined in Sauro et al., 1987). Application of the matrix method produces the following matrix equation which can be solved for these control coefficients:

$$\mathbf{E} \cdot \mathbf{C} = \mathbf{M} \quad (1)$$

$\mathbf{E}$ is called the *elasticity matrix*; $\mathbf{C}$, the *control matrix*, is the matrix of the control coefficients that are to be evaluated; $\mathbf{M}$ is a matrix which defines the relationships between the elements of $\mathbf{E}$ and the control coefficients in $\mathbf{C}$. The forms of $\mathbf{E}$, $\mathbf{C}$ and $\mathbf{M}$ are described in the following subsections, with reference to the pathway shown in Scheme 1, which will be used throughout the present paper to illustrate aspects of both the matrix method and MetaCon. The equation is solved by inverting the elasticity matrix:

$$\mathbf{C} = \mathbf{E}^{-1} \cdot \mathbf{M} \quad (2)$$

The rows and columns of the three matrices are based on the various summation and connectivity theorems for the different control coefficients; the matrix method rules simply show how to generate enough relationships to enable the coefficients to be calculated. The first requirement is that, for the elasticity matrix to be inverted, it must be square. Therefore the matrix method must provide n relationships for the n rows of the matrix, where n is the number of steps in the pathway.

Fell and Sauro (1985) used the flux control coefficient connectivity and summation theorems (Kacser and Burns, 1973) for the initial development of the matrix method. For a linear pathway there is one summation theorem, and $n - 1$ connectivity theorems, enough to produce a square elasticity matrix. They showed how to modify the connectivity theorem to cover

Scheme 1 Hypothetical pathway showing features typical of a real biochemical system

R1 to R6 are step names; S1 to S4, NAD and NADH are variable metabolites; NAD/NADH are a conserved pair; glucose, glycogen and lactate are pool metabolites. Steps R1, R2 and R3 form a branch-point meeting at metabolite S1.

pathways with conserved metabolites. Kacser (1983) had shown how an additional constraint on flux control coefficients arose in a simple branched pathway; this idea was expanded by Fell and Sauro (1985) to a general branch-point theorem which could be applied to realistic biochemical systems. The branch-point theorem is necessary because each step meeting at a branch-point increases the number of steps by 1 without a concomitant increase in the number of variable metabolites. Therefore there are no longer enough connectivity equations to complete a square matrix. Each branch-point supplies one branch-point equation, so the required number of equations (n) is restored.

$$\begin{pmatrix}
 1 & 1 & 1 & 1 \\
 e_{S1}^{R1} & e_{S1}^{R2} & e_{S1}^{R3} & 0 \\
 0 & 0 & e_{S2}^{R3} & e_{S2}^{R4} \\
 0 & 0 & 0 & e_{NAD}^{R4} - e_{NADH}^{R4} \frac{NAD}{NADH} \\
 0 & 0 & 0 & e_{S3}^{R4} \\
 0 & -0.5 & 0.5 & 1.0 \\
 \hline
 & J_{R2} & J_{R3} & J_{R4}
 \end{pmatrix}
 \quad
 \begin{pmatrix}
 1 & 1 \\
 0 & 0 \\
 0 & 0 \\
 -e_{S4}^{R5} \frac{NAD}{S4} & e_{NAD}^{R6} - e_{NADH}^{R6} \frac{NAD}{NADH} - e_{S4}^{R6} \frac{NAD}{S4} \\
 e_{S3}^{R5} - e_{S4}^{R5} \frac{S3}{S4} & -e_{S4}^{R6} \frac{S3}{S4} \\
 \hline
 1.0 & 1.0 \\
 J_{R5} & J_{R6}
 \end{pmatrix}$$

Matrix 1

The scope of the method was extended by Sauro et al. (1987) and Westerhoff and Kell (1987) to allow determination of the concentration control coefficients by incorporating the concentration control coefficient summation theorem (Heinrich and Rapoport, 1974) and connectivity theorem (Heinrich and Rapoport, 1975; Westerhoff and Chen, 1984; Hofmeyr et al., 1986). The branch-point theorem was extended to include concentration control coefficients, and a further coefficient determining the distribution of flux at a branch-point was defined which could be calculated by the matrix method (Sauro et al., 1987; Westerhoff and Kell, 1987; Small and Fell, 1989). Finally, Reder (1988) generalized the proofs and showed that the matrix equations could be written for any pathway. The relationship between the matrix method and the method of Reder is discussed in Appendix 1.

As mentioned above, the use of MetaCon is illustrated by describing its application to determine the control coefficients for the pathway depicted in Scheme 1, a hypothetical six-step

pathway roughly modelled on glycolysis. This pathway has been chosen because it contains examples of the type of structural complexity found in real biochemical systems: non-unitary (though integral) stoichiometries; a pair of conserved metabolites, NAD and NADH; and a branch-point at S1. Solution of the control coefficients by hand for such a system is a non-trivial problem, yet the system is of the size to which Metabolic Control Analysis could, and should, be applied if important advances are to be made in our understanding of the control of real biochemical pathways.

Glucose, glycogen, lactate and NAD/NADH are named to help in understanding the meaning of and application of the terms *pool* metabolites (glucose, glycogen and lactate) and *conserved* metabolites (NAD and NADH), which are used in this and other papers on Control Analysis. However, simple representations are used for the names of the steps (R1–R6) and most of the metabolites (S1–S4). This is to keep Matrices 1 and 2 (the elasticity and control coefficient matrices respectively), and the expressions for the control coefficients as compact as possible.

The elasticity matrix

For pathways containing conserved metabolites or branch-points, there is no unique elasticity matrix; alternative matrices must have the same numbers of connectivity rows and branch-point rows, but the contents of the rows can differ, depending on how the matrix-method rules are applied. One possible elasticity matrix for the example pathway, the one which would be produced by MetaCon, is shown in Matrix 1 (for the determination of the control coefficients from the component properties, the elasticities, concentrations and fluxes appearing in the elasticity matrix represent the set that must be measured):

e_M^{Rn} is the elasticity of metabolite M towards the velocity of step Rn; S1–S4, NAD and NADH are the steady-state concentrations of the variable metabolites, and J_{Rn} represents the steady-state flux through step Rn.

Thus the six rows of the elasticity matrix represent the summation theory (row 1); connectivity rows for the independent metabolites (rows 2 and 3); connectivity rows showing the effects of conserved metabolites (rows 4 and 5); and a row for the branch-point at metabolite S1 (row 6). The form of row 6 is different from the form of the branch-point relationships, as they are derived by Fell and Sauro (1985) and Sauro et al. (1987). However, both ways of expressing branch-point rows are mathematically equivalent.

The control matrix

The full control coefficient matrix, C, as defined by Sauro et al. (1987), is a square matrix of dimension n . The first column of C is always a column of flux control coefficients for one chosen flux

in the pathway, the *reference flux*. For a pathway with m_0 independent metabolites and b branch points, this is followed by m_0 columns of concentration control coefficients for the independent metabolites, and b columns of branch-point distribution control coefficients. The branch-point distribution control coefficient C_{Rn}^a describes the effect of step Rn on the fraction, a , of the reference flux that enters a given branch of the pathway. Thus, in Scheme 1, if the flux through step $R1$, J_{R1} is chosen as the reference flux, the *branch flux* can be chosen to represent the proportion which passes out either through $R2$ (in which case a would be J_{R2}/J_{R1}), or through $R3$ (a would be J_{R3}/J_{R1}). The proportion passing through the other arm would be $(1-a)$, i.e. J_{R3}/J_{R1} and J_{R2}/J_{R1} respectively for the two choices of branch flux. The choice of flux ratios to represent a determines the contents of the corresponding columns of M , as described in the following subsection. Each of the n rows corresponds to all the control coefficients for a given step. Thus for the example pathway with J_{R1} as the reference flux, C is:

$$\begin{pmatrix} C_{R1}^{J_{R1}} & C_{R1}^{S1} & C_{R1}^{S2} & C_{R1}^{NAD} & C_{R1}^{S3} & C_{R1}^a \\ C_{R2}^{J_{R1}} & C_{R2}^{S1} & C_{R2}^{S2} & C_{R2}^{NAD} & C_{R2}^{S3} & C_{R2}^a \\ C_{R3}^{J_{R1}} & C_{R3}^{S1} & C_{R3}^{S2} & C_{R3}^{NAD} & C_{R3}^{S3} & C_{R3}^a \\ C_{R4}^{J_{R1}} & C_{R4}^{S1} & C_{R4}^{S2} & C_{R4}^{NAD} & C_{R4}^{S3} & C_{R4}^a \\ C_{R5}^{J_{R1}} & C_{R5}^{S1} & C_{R5}^{S2} & C_{R5}^{NAD} & C_{R5}^{S3} & C_{R5}^a \\ C_{R6}^{J_{R1}} & C_{R6}^{S1} & C_{R6}^{S2} & C_{R6}^{NAD} & C_{R6}^{S3} & C_{R6}^a \end{pmatrix}$$

Matrix 2

$C_{Rm}^{J_{Rn}}$ is the control coefficient of step m on the flux through step Rn ; C_{Rm}^X is the concentration control coefficient of step m on metabolite X , and C_{Rm}^a is the branch-point control coefficient of step m on the proportion of the reference flux which passes out through whichever of $R2$ and $R3$ was chosen to represent the branch flux.

As the flux control coefficients are only calculated for the reference flux, in a branched pathway there will be one or more fluxes whose control coefficients have not been directly calculated. However, to complete the analysis, the flux control coefficients for all other pathway fluxes can be calculated from those for the reference flux and the branch-point control coefficients.

The M matrix

Matrix M is also of dimension n . The square submatrix formed by the first $m_0 + 1$ rows and columns is always a diagonal matrix (a diagonal matrix has non-zero elements on the leading diagonal, zeros everywhere else, i.e. $m_{ij} \neq 0$ for $i = j$; $m_{ij} = 0$ for $i \neq j$), and always has the following values on the leading diagonal: the top left element is one (i.e. $m_{1,1} = 1$); the elements in the next m_0 columns are -1 (i.e. $m_{j,j} = -1$, $1 < j \leq m_0 + 1$). The entries in each of the last b columns are defined by Small and Fell (1989) for pathways of any complexity: each column has $n - b$ zeros, followed by the last b elements of the j th column of E , where j is the step through which the chosen branch flux passes. Thus, in the example pathway, if the proportion of the reference flux passing through $R2$ is chosen as a , then the bottom-right element in M would be $-0.5/J_{R2}$, whereas if the proportion passing through $R3$ were chosen, it would be $0.5/J_{R3}$ (see Matrix 1). The contents of E , C and M for an example pathway with two branch-points are described by Small and Fell (1989).

OUTLINE DESCRIPTION OF THE PROGRAM

The vast majority of MetCon's input and output is from and to simple text (ASCII) files. The only exceptions are: messages written to the computer screen to inform the user of the state of

processing; a minimal amount of prompting, to which the user can respond from the keyboard, e.g. in case conflicting processing operations have been specified, or the reference flux has been omitted from the input file.

MetaCon uses several output files, but only one input file that contains two required, and two optional, sections, each starting with a relevant keyword (a compulsory 'heading' word).

1. The control section: instructions defining which processing options are to be carried out, including the choice of reference flux, and, in a branched pathway, the branch flux(es). Most commands are written in what amounts to plain English.

2. The reactions section: the balanced chemical equations of the system to be analysed. The form used for entering the reaction scheme to be read by MetaCon is similar to that used by other programs designed for modelling biochemical networks e.g. SCAMP (Sauro and Fell, 1991).

3. The additional elasticities section: contains elasticity terms for any reaction, other than those due to its substrate(s) and product(s).

4. The equations section: equations specifying values of, or expressions for, any variable defined in the system.

The first two sections are the required sections, and their contents are self-explanatory; the third and fourth are the optional sections, and their contents will be described in more detail below, in connection with the processing stages carried out by MetaCon. These stages are as follows.

1. Analyse the pathway, and create the corresponding elasticity matrix, which can be written to an output file. This step is the actual Metabolic Control Analysis. The algorithm is described in Appendix 2.

2. Solve the matrix equation to produce algebraic expressions (they are actually polynomials in several variables) for the control coefficients. The flux control coefficients of all steps in the pathway on the reference flux are calculated. The user can also request the control coefficients of all steps on all variable metabolite concentrations, and in a branched pathway, the branch-point distribution coefficients over the branch flux(es).

3. Modify these expressions, by substituting equations or numerical values from the equations section of the input file, for any of the variables contained in the elasticity matrix. These equations can contain additional 'user-defined' variables, which themselves can be further substituted. An obvious example of the use of this facility is to enter equations for elasticities of reactions which are derived from their kinetic expressions.

4. Carry out a sensitivity analysis to determine the sensitivity of all control coefficients, C , to each variable v , which appears in the elasticity matrix, in terms of both unscaled sensitivities, $\partial C/\partial v$, and scaled sensitivities $(\partial C/\partial v) \cdot (v/C)$. The sensitivity analysis is also carried out algebraically, to produce expressions of the same form as those for the control coefficients. To achieve this we have developed an algorithm based on that of Small and Fell (1990) (S. Thomas and D. A. Fell, unpublished work).

The only limitations of the form of network which MetaCon can analyse are: (1) stoichiometries must be integral; (2) the network must be fully connected by the flow of mass, that is, all metabolites and steps must form a continuous network. Systems composed of physically disconnected networks which are only connected by regulatory effects (e.g. cascades, gene expression, hormone action) cannot be analysed. MetaCon checks that the network is continuous, as described in Appendix 3. At the moment, this is not a particularly severe limitation; most of the theory and nearly all applications of Control Analysis have been carried out on fully linked networks of this type. However, some theoretical approaches to disconnected systems have been

(a) $C_{R3}^{R1} =$

$$\begin{aligned}
& - 0.5 * J_{R3} * J_{R4} * J_{R5} * J_{R6} * [NAD] * [S3] * [NADH]^{-1} * [S4]^{-1} * e_{R1, S1 * e, R4, S2 *} \\
& e_{R5, S4 * e, R6, NADH} + 0.5 * J_{R3} * J_{R4} * J_{R5} * J_{R6} * [NAD] * [NADH]^{-1} * e_{R1, S1 * e, R4, S2} \\
& * e_{R5, S3 * e, R6, NADH} + 0.5 * J_{R3} * J_{R4} * J_{R5} * J_{R6} * [NAD] * [S4]^{-1} * e_{R1, S1 * e, R4, S2 *} \\
& e_{R5, S3 * e, R6, S4} + 0.5 * J_{R3} * J_{R4} * J_{R5} * J_{R6} * [S3] * [S4]^{-1} * e_{R1, S1 * e, R4, S2 *} \\
& e_{R5, S4 * e, R6, NAD} - 0.5 * J_{R3} * J_{R4} * J_{R5} * J_{R6} * e_{R1, S1 * e, R4, S2 * e, R5, S3 * e, R6, NAD};
\end{aligned}$$

(b) $C_{R3}^{R1} = - 0.5 \frac{J_{R3} J_{R4} J_{R5} J_{R6} [NAD] [S3] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NADH}{[NADH] [S4]} + 0.5 \frac{J_{R3} J_{R4} J_{R5} J_{R6} [NAD] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NADH}{[NADH]} +$

$$0.5 \frac{J_{R3} J_{R4} J_{R5} J_{R6} [NAD] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6}}{[S4]} + 0.5 \frac{J_{R3} J_{R4} J_{R5} J_{R6} [S3] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NAD}{[S4]} - 0.5 J_{R3} J_{R4} J_{R5} J_{R6} e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NAD$$

(c) $C_{R3}^{R1} =$

$$\begin{aligned}
& - 0.01458 * [S3] * [S4]^{-1} * e_{R1, S1 * e, R4, S2 * e, R6, NAD} + 0.1458 * [S3] * [S4]^{-1} * \\
& e_{R1, S1 * e, R4, S2 * e, R6, NADH} + 0.2268 * [S4]^{-1} * e_{R1, S1 * e, R4, S2 * e, R6, S4} - \\
& 0.02268 * e_{R1, S1 * e, R4, S2 * e, R6, NAD} + 0.2268 * e_{R1, S1 * e, R4, S2 * e, R6, NADH};
\end{aligned}$$

(d) $C_{R3}^{R1} = - 0.01458 \frac{[S3] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NAD}{[S4]} + 0.1458 \frac{[S3] e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NADH}{[S4]} + 0.2268 \frac{e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6}}{[S4]} -$

$$0.02268 e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NAD + 0.2268 e_{S1}^{R1} e_{S2}^{R4} e_{S3}^{R5} e_{S4}^{R6} NADH$$

Figure 1 MetaCon output: various forms of the dividend of the expression for the flux control coefficient C_{R3}^{R1} from the example pathway in Scheme 1

(a) Unsubstituted expression; (b) T_EX output for the unsubstituted control coefficient in (a), using a MetaCon-produced file; (c) expression following substitution of 11 of the 22 variables by numeric values; (d) T_EX output for the substituted control coefficient in (c), using a MetaCon-produced file. The numerator of the flux control coefficient C_{R3}^{R1} is represented in (a) and (c) by 'C, R1, R3', and in (b) and (d) by C_{R3}^{R1} . The flux through step Rn is represented in (a) and (c) by 'J, Rn', and in (b) and (d) by J_{Rn} . The elasticity of metabolite Sm towards step Rn is represented in (a) and (c) by 'e, Rn, Sm', and in (b) and (d) by e_{Sm}^{Rn} .

described (Savageau, 1971; Westerhoff and van Workum, 1990; Westerhoff et al., 1990; Kahn and Westerhoff, 1991), and an experimental determination of the control of histone expression in *Xenopus laevis* has been carried out (Koster et al., 1990). However, the theoretical approach of Kahn and Westerhoff (1991) could be readily incorporated into the capabilities of MetaCon. The bases of MetaCon's built-in checks for a pathway which is analysable are described in Appendix 3.

AN EXAMPLE OF USING METACON TO CARRY OUT A CONTROL ANALYSIS

In this section, important parts of the MetaCon input and output for the pathway in Scheme 1 are shown. The lines of the input file are written exactly as they would be entered in the file, to be read by MetaCon. Thus the control section for the reaction in Scheme 1 to (i) carry out the control analysis and (ii) calculate expressions

for all the control coefficients by inverting the elasticity matrix would be entered as follows:

```

CONTROL
ANALYSE
INVERT
PRINT ELASTICITY MATRIX
CONCENTRATION CONTROL COEFFICIENTS
BRANCH CONTROL COEFFICIENTS
REFERENCE FLUX = 1
BRANCH FLUX = 2

```

The keyword 'CONTROL' starts the section; 'ANALYSE' and 'INVERT' specify respectively that the elasticity matrix is to be produced, then inverted; 'PRINT ELASTICITY MATRIX' is self-explanatory; the flux control coefficients for the reference flux are calculated automatically whenever 'INVERT' is speci-

fied, but calculation of the concentration and branch-point distribution coefficients have to be specifically requested if they are required (lines 5 and 6 above). Lines 7 and 8 specify that the reference flux and the branch flux are to be those of the steps 1 and 2 in the reaction scheme, in the order in which they occur in the file. The reaction scheme is in the second compulsory section, starting with the keyword 'REACTIONS':

REACTIONS

```
[R1] $glucose = S1
[R2] S1 = $glycogen
[R3] S1 = 2 S2
[R4] S2 + NAD = NADH + S3
[R5] S3 = S4
[R6] S4 + NADH = NAD + $lactate
```

The step names are optional; if used they must be enclosed in square brackets at the start of each line. If a step is not named, it is automatically numbered according to its position in the scheme. Names of any length are allowed in the input file. Thus the third step could be entered as:

```
[Ald] FBP = 2 triose_phosphate
```

Pool metabolites are specified in the input file by preceding their names with a '\$'. To complete the analysis, effects due to other interactions such as feedback inhibition or activation, if any, are entered into the optional additional elasticities section; e.g., if metabolites S3 and S4 in Scheme 1 had feedback effects on step 1 (R1), they would be represented as:

ADDITIONAL ELASTICITIES

```
[R1] S3, S4
```

The only limitation is that effectors must be variable metabolites of the reaction scheme. Effects of pool metabolites, or species not defined in the reaction scheme, are not permitted. This is simply a limitation of the connectivity relationships, which do not include elasticities due to external metabolites (called *kappa* elasticities, ϵ). The signs and magnitudes are not specified at this stage, but if they are known, they can be entered along with other variable values in the equations section.

The contents of the file so far described are sufficient to carry out both the Control Analysis and the calculation of the control coefficients. The elasticity matrix is printed to file in a form readily recognizable as Matrix 1 (output not shown). To solve the equation for the control coefficients involves inverting the elasticity matrix. A discussion of the methods which can be employed to achieve this is outside the scope of the present paper. We have chosen to use a straightforward minor expansion, which can be modified to achieve greater efficiency if necessary [see Griss (1976), for example]. Using this method it is natural to express the inverse of a matrix as the transpose of the matrix of its cofactors divided by the determinant of the matrix. In other words, each element in the inverse can be expressed as a dividend unique for each element, and a divisor (the determinant) which is common to all elements. As the control coefficients are calculated by multiplying the corresponding element in the inverse of the elasticity matrix by a constant (1 or -1), they can also be expressed in the same way. This is the representation used by MetaCon when they are output; the dividends of all the control coefficients are printed, followed by the determinant of

the elasticity matrix, which is their common divisor. The expressions (dividends and divisor) will, of course, be functions of the elasticities, metabolite concentrations and fluxes in the elasticity matrix.

The dividend of one of the control coefficients, C_{R3}^{S1} , from the example pathway, as output by MetaCon, is shown in Figure 1(a). The expression is admittedly difficult to read, and could be improved in a number of ways. However, the form is very easy to re-read by computer, a necessary feature of MetaCon's design (see also the Discussion section). Also, MetaCon can produce an output file which can be read by the typesetting program T_EX. T_EX (Knuth, 1981) is in the public domain, and is available on a number of different systems, including IBM-PC compatible, and UNIX systems. Input is an ASCII file containing the typesetting instructions and the actual text, from which formatted output can be produced on a wide range of printers. Figure 1(b) shows the expression for C_{R3}^{S1} after T_EX processing of the MetaCon-produced T_EX file.

Subsequent substitution of variable(s) by numerical values will result in simpler expressions, whereas substitution by expressions will, not surprisingly, tend to produce more complex expressions. For example, supposing the values of 11 of the 22 variables in Matrix 1 were as given below, the expressions for C_{R3}^{S1} in Figures 1(a) and 1(b) would be modified to produce Figures 1(c) and 1(d) respectively. The values would be entered into the input file as shown:

EQUATIONS

```
e, R3, S1 = 0.3
e, R5, S3 = 0.7
e, R5, S4 = -0.45
[NAD] = 10
[NADH] = 1
J, 6 = J, 5
J, 5 = J, 4
J, 4 = 2 * J, 3
J, 3 = J, 1 - J, 2
J, 2 = 0.1
J, 1 = 0.4
```

where the meanings of the symbols are as defined in the legend to Figure 1. This example shows instances of the assignment to variables of both numeric values and other variables. A feature currently being added to MetaCon is the automatic determination of the relationships between fluxes in the same branch of the pathway. When it is implemented, the lines $J, 6 = J, 5$ and the two following lines will no longer be necessary.

DISCUSSION

The past decade has seen a revolution in the microcomputer world, as both the gross power, and the power/unit cost available in desk-top computers has increased by orders of magnitude. The same period has seen an increased awareness of, and number of applications of, Metabolic Control Analysis throughout biochemistry and closely related fields (Fell, 1992). Not surprisingly, a number of microcomputer-based methods have been described for applying Metabolic Control Analysis, of which MetaCon is one of the latest. Examples include SCAMP (Sauro and Fell, 1991), MetaModel (Cornish-Bowden and Hofmeyr, 1991), CONTROL (Letellier et al., 1991) and a program developed by Schulz (1991).

Of these, SCAMP and MetaModel are slightly different from the others in that they are primarily metabolic simulators; their

main purpose is to simulate the time-dependence of metabolic systems or to analyse a metabolic steady state. These functions are carried out by numerical means. The control and elasticity coefficients of a system are calculated from the steady state, if one exists, by determining the response in the relevant reaction flux or metabolite concentration to a small variation in the controlling parameter (for control coefficients) or steady-state concentration (for elasticities).

CONTROL, like MetaCon, is much more specific for Metabolic Control Analysis applications. Control coefficients are calculated directly from the values of the system's elasticities, the values of which are entered by the user into a matrix. CONTROL can calculate both normalized and unnormalized control coefficients, and can produce the corresponding summation and connectivity theorems. However, like the numerical simulators described above, which require values for all parameters (kinetic constants, pool metabolite concentrations etc.) before an analysis can be carried out, CONTROL requires values for all elasticities before it can calculate the control coefficients. Furthermore, if normalized control coefficients are to be calculated, all the steady-state fluxes of the pathway under investigation must be known. In addition, the user must convert the reaction scheme to a stoichiometry matrix and type this in.

The program written and described by Schulz (1991) is the most similar to MetaCon in that it produces symbolic expressions for control coefficients. However, it is currently limited to a linear pathway, which can have branches arising from that pathway, and does not have any facility for analysing pathways with conserved cycles. Nor does it have the ability to carry out user-defined substitution into the calculations of the control coefficients, other than indicating the effect of each metabolite on each enzyme of the system. Finally, instructions are entered in simply as user responses to a series of prompts: there is no facility for saving input instructions, and if a mistake is made entering a scheme, or a scheme is to be modified for another analysis, all the relevant information must be re-entered.

The major feature which distinguishes MetaCon from other programs is the combination of symbolic (algebraic) and numeric information which may be incorporated into a model. Depending on the particular system being investigated, and the amount of data available, the control coefficients may evaluate to a number, or may be expressed as polynomials containing enzyme kinetic constants, equilibrium constants, reaction fluxes etc.

The decision as to whether to use a simulator or an algebraic program such as MetaCon can be influenced by a number of factors, and there are discussions in the computing literature for those who are interested (e.g. Brown and Hearn, 1979). Each approach has advantages and disadvantages, especially when carrying out theoretical investigations. For the analysis of experimental data, the advantages tend to lie more with the algebraic approach. Whilst we believe that MetaCon will prove to be a very useful tool for theoretical investigations, we feel that it is to the experimental community that MetaCon will be of greatest benefit. Whereas theoreticians have been very well served by the programs discussed above, CONTROL has been, up until now, the only program available to workers in Metabolic Control Analysis, specifically for the *analysis* of experimental data.

One further difference between MetaCon and metabolic simulators is that simulators calculate the steady-state fluxes from the values of the parameters of the model system, whereas in MetaCon the steady-state fluxes are treated like any other quantity in the system; values can be entered for them, or they can be left as variables in the polynomials, and their effects on the values of the control coefficients determined. This difference is potentially important, and it typifies the difference in approach

between the two methods: on the one hand a numerically calculated value is used, dependent on the fidelity of the model; on the other, a value which will probably depend on the accuracy of one or more experimentally measured flux values.

MetaCon should be useful to newcomers to Control Analysis who might be daunted by the amount of data processing normally required before they can obtain any control information whatever. It should also be beneficial to established workers who are interested in the effect on the pattern of control of altering one or more parameters in an experimental investigation or a model. A promising area is to use MetaCon's method for determining the sensitivity of control coefficients to the variables in the elasticity matrix. Using this method, those variables that have the largest effect on one or more control coefficients of interest can be identified. This narrows down the number of variables that need to be investigated. These variables can then be left as unknowns, either singly or together, in MetaCon; the resulting expressions for the control coefficients can then be plotted, using appropriate ranges of the unknowns, to see the potential effect of altering the variables' values. As discussed by Small and Fell (1990), the relative importance of experimental errors in the accuracy of the values of the control coefficients can then be assessed. To summarize MetaCon's uses, it can help in an investigation in three ways. First, to show which variables need to be measured, by producing the elasticity matrix. Second, to analyse experimental/model data and so produce expressions and/or values for the control coefficients. Third, to help determine the degree of accuracy which is required when ascertaining the values of variables. However, as we have stated, these aims do not preclude its use in theoretical investigations.

As mentioned above, although we have automated the complete analysis procedure, the program uses text files for both receiving and outputting information. This contrasts with many modern programs which use features such as graphical interfaces or menus for entering data and/or commands and for examining output. The main reason that this approach has been chosen is to try to minimize development time, at least up to the stage now reached, where a fully workable program exists; a secondary reason is that we wanted the program to be portable to different operating systems and machines. To help achieve this second aim the program has been written entirely in the language C complying to the ANSI draft standard (Kernighan and Ritchie, 1988). Use of graphics features (for example) is invariably compiler-dependent (and therefore machine-dependent). A version to run on a DEC workstation under the UNIX operating system is also under development; using standard C, the code is very nearly 100% compatible, reducing problems porting the program to the new system.

Whilst we realize that the lack of a sophisticated user interface is a potential drawback to people who are unfamiliar with computers, very little computing ability is required to use MetaCon. An understanding of the rudiments of text files is required, including the ability to edit and examine them, using either a text editor, or a word-processor in text mode: these are generally useful skills which, if not already possessed, are straightforward to learn. Furthermore, although menu-type interfaces are useful in the initial stage of familiarization with a program, once familiarity has been achieved, the type of system used by MetaCon can often be quicker and easier to use.

Also, the interface as it is, especially the input, is relatively easy to understand and use. This can be seen from the example input file; commands are very simple, and all biochemists understand how to write stoichiometric chemical equations. An input file can be copied and edited to produce an input file quickly for a different problem.

As discussed in the previous section, the expressions for the control coefficients which are output by MetaCon could be improved. An obvious way in which this could be achieved is to use a factored form of expression, rather than the fully expanded form. Minimization of development time is one reason we have used the fully expanded form. Another is that there are very powerful and sophisticated algebraic manipulators, such as REDUCE (Hearn, 1985) which are already available. A small amount of editing (such as removal of commas and semi-colons) quickly converts the MetaCon output into a form suitable for reading into an algebraic manipulator. Criticism could be levelled that this assumes additional computing ability on the part of the user, but we feel that use of a system such as REDUCE will only be necessary in the analysis of large systems with many unknowns, a situation unlikely to be encountered by the naive computer user.

The MS-DOS version of the program is available from the authors on the receipt of a 3½ in formatted floppy disk; the manual is included on the disk.

This work was supported by grant number GR/E89117 from the Science and Engineering Research Council. We thank Dr. Jeremy Bennett and Dr. John Fitch for their help in the design of MetaCon, and Joao Pedro Moniz Barreto and Dr. Rankin Small for their helpful comments on the manuscript.

REFERENCES

- Brown, W. S. and Hearn, A. C. (1979) *Comput. Phys. Commun.* **17**, 207–215
- Cascante, M., Franco, R. and Canela, E. I. (1989a) *Math. Biosci.* **94**, 271–288
- Cascante, M., Franco, R. and Canela, E. I. (1989b) *Math. Biosci.* **94**, 289–309
- Cornish-Bowden, A. and Cárdenas, M. L. (eds.) (1990) *Control of Metabolic Processes*, Plenum Press, New York
- Cornish-Bowden, A. and Hofmeyr, J.-H. S. (1991) *Comput. Appl. Biosci.* **7**, 89–93
- Fell, D. A. (1992) *Biochem. J.* **286**, 313–330
- Fell, D. A. and Sauro, H. M. (1985) *Eur. J. Biochem.* **148**, 555–561
- Griss, M. L. (1976) *ACM Trans. Math. Software* **2**, 31–49
- Hearn, A. C. (1985) *REDUCE User's Manual*, Rand Corporation, Santa Monica
- Heinrich, R. and Rapoport, T. A. (1974) *Eur. J. Biochem.* **42**, 89–95
- Heinrich, R. and Rapoport, T. A. (1975) *Biosystems* **7**, 130–136
- Hofmeyr, J.-H. S. (1989) *Eur. J. Biochem.* **186**, 343–354
- Hofmeyr, J.-H. S., Kacser, H. and van der Merwe, K. J. (1986) *Eur. J. Biochem.* **155**, 631–641
- Kacser, H. (1983) *Biochem. Soc. Trans.* **11**, 35–40
- Kacser, H. and Burns, J. A. (1973) *Symp. Soc. Exp. Biol.* **27**, 65–104
- Kahn, D. and Westerhoff, H. V. (1991) *J. Theor. Biol.* **153**, 255–285
- Kernighan, B. W. and Ritchie, D. M. (1988) *The C Programming Language*, Prentice-Hall, Englewood Cliffs, NJ
- Knuth, D. E. (1981) *The Art of Computer Programming*, vol. 2: Seminumerical Algorithms, 2nd edn., Addison-Wesley Publishing Co., Reading, MA
- Koster, J. G., Destree, O. H. J., Raat, N. J. H. and Westerhoff, H. V. (1990) *Biomed. Biochim. Acta* **49**, 855–877
- Letellier, T., Reder, C. and Mazat, J.-P. (1991) *Comput. Appl. Biosci.* **7**, 383–390
- Meléndez-Hevia, E., Torres, N. V., Sicilia, J. and Kacser, H. (1990) *Biochem. J.* **265**, 195–202
- Reder, C. (1988) *J. Theor. Biol.* **135**, 175–201
- Savageau, M. A. (1971) *Arch. Biochem. Biophys.* **145**, 612–621
- Sauro, H. M. and Fell, D. A. (1991) *Math. Comput. Modelling* **15**, 15–28
- Sauro, H. M., Small, J. R. and Fell, D. A. (1987) *Eur. J. Biochem.* **165**, 215–221
- Schulz, A. R. (1991) *Biochem. J.* **278**, 299–304
- Sen, A. K. (1990) *Biochem. J.* **186**, 343–354
- Small, J. R. and Fell, D. A. (1989) *J. Theor. Biol.* **136**, 181–197
- Small, J. R. and Fell, D. A. (1990) *Eur. J. Biochem.* **191**, 413–420
- Westerhoff, H. V. and Chen, Y. D. (1984) *Eur. J. Biochem.* **142**, 425–430
- Westerhoff, H. V. and Kell, D. B. (1987) *Biotechnol. Bioeng.* **30**, 101–107
- Westerhoff, H. V. and van Workum, M. (1990) *Biomed. Biochim. Acta* **49**, 839–853
- Westerhoff, H. V., Koster, J. G., van Workum, M. and Rudd, K. (1990) in *Control of Metabolic Processes* (Cornish-Bowden, A. and Cárdenas, M. L., eds.), pp. 399–412, Plenum Press, New York

Received 9 July 1992/18 November 1992; accepted 22 December 1992

APPENDIX 1

Correspondence between the Reder method and the matrix method

From the work of Reder (1988) it can be shown that the matrix method produces exactly the correct number and form of equations no matter what the complexity of the pathway. Reder proved that the following relationships hold.

For flux control coefficients:

$$\mathbf{C} \cdot \mathbf{K} = \mathbf{K} \quad (\text{A1})$$

$$\mathbf{C} \cdot \mathbf{D}_x \mathbf{v} \cdot \mathbf{L} = \mathbf{0} \quad (\text{A2})$$

and for concentration control coefficients:

$$\mathbf{\Gamma} \cdot \mathbf{K} = \mathbf{0} \quad (\text{A3})$$

$$\mathbf{\Gamma} \cdot \mathbf{D}_x \mathbf{v} \cdot \mathbf{L} = -\mathbf{L} \quad (\text{A4})$$

where $\mathbf{\Gamma}$ is an $m \times n$ matrix of all possible unscaled concentration control coefficients; $\mathbf{C}$ is an $n \times n$ matrix of all possible unscaled flux control coefficients; and $\mathbf{D}_x \mathbf{v}$ is an $n \times m$ matrix of all possible unscaled elasticity coefficients. $\mathbf{K}$ is a basis of the nullspace (or kernel), of the stoichiometry matrix, and $\mathbf{L}$ is the link matrix of the stoichiometry matrix [these matrices are defined fully by Reder (1988)]. The relationship between these definitions of the control coefficients and the standard ones are explained in Schuster and Heinrich (1992).

The link matrix identifies the presence of conserved groups of metabolites (if any), and defines the relationship between the concentrations of the members of each conserved group. If there are conserved metabolites, there will be fewer independently variable metabolites than the total number, m , as alteration of the concentration of any metabolite involved in a conserved

group will necessarily entail alteration of the concentration of at least one of the other metabolites involved in the group. If there are no conserved metabolites, $\mathbf{L}$ is the identity matrix. If there are, $\mathbf{L}$ can be partitioned into (i) an identity matrix, $\mathbf{I}$, of dimension m_0 where m_0 equals m minus the number of conservation relationships, and (ii) a further matrix, $\mathbf{L}_0$, which defines the conservation relationships. The $\mathbf{L}_0$ partition has dimension $m - m_0$ by m (Reder, 1988). m_0 is the number of independent metabolites, $m - m_0$ the number of dependent ones. The independent metabolites are represented by the columns of $\mathbf{L}_0$, the dependent ones by the rows. The $\mathbf{L}_0$ partition for the example pathway in Scheme 1 of the main paper is described in Appendix 2.

Eqns. (A1)–(A4) differ from those of the matrix method in the following ways.

1. They use unnormalized rather than normalized coefficients.
 2. All the matrices of coefficients are transposed versions of those used in the matrix method.
 3. They produce values for all possible flux control and concentration control coefficients, whereas the matrix method provides a solution only for the flux control coefficients of each step on a chosen 'reference flux', and for the concentrations of the m_0 independent metabolites.
 4. There are no branch-point distribution control coefficients.
- Normalizing (scaling) the equations involves scaling the elasticity coefficients ($\mathbf{D}_x \mathbf{v}$), the control coefficients ($\mathbf{C}$ and $\mathbf{\Gamma}$), $\mathbf{L}$ and $\mathbf{K}$. The normalized forms of the control and elasticity coefficient matrices are (Reder, 1988):

$$C_{(\text{norm})} = (\partial J^0)^{-1} \cdot C \cdot (\partial J^0)$$

$$\Gamma_{(\text{norm})} = (\partial S^0)^{-1} \cdot \Gamma \cdot (\partial J^0)$$

$$D_x v_{(\text{norm})} = (\partial J^0)^{-1} \cdot D_x v \cdot (\partial S^0)$$

where (∂J^0) is a diagonal matrix of all the steady state fluxes, and (∂S^0) is a diagonal matrix of all the steady-state metabolite concentrations in the system.

$$K_{(\text{norm})} = (\partial J^0)^{-1} \cdot K$$

$$L_{(\text{norm})} = (\partial S^0)^{-1} \cdot L \cdot (\partial S^d)$$

(∂S^d) is a diagonal matrix of all the independent metabolite concentrations.

Depending on the structure of the pathway, the normalized matrix $D_x v_{(\text{norm})} \cdot L_{(\text{norm})}$ can take one of two forms.

1. In pathways with no conserved metabolites $(\partial S^0)^{-1} \cdot L \cdot (\partial S^d) = I$ (the identity matrix).

2. In pathways with conserved metabolites, $L \neq I$, and (∂S^0) and (∂S^d) are not equal. As a result, $(\partial S^0)^{-1} \cdot L \cdot (\partial S^d) \neq I$, and certain elements in $D_x v_{(\text{norm})} \cdot L_{(\text{norm})}$ contains ratios of the concentrations of the conserved metabolites.

This is a major difference between the use of normalized and unnormalized elasticity coefficients: no metabolite ratios occur in the latter case. The effect of conserved metabolites on the form of $D_x v_{(\text{norm})} \cdot L_{(\text{norm})}$ and consequently on the elasticity matrix, for the example pathway of Scheme 1 is described in Appendix 2.

Dropping the subscripts (norm) for clarity, and combining eqns. (A1) with (A2), and eqns. (A3) with (A4):

$$C[K|D_x v \cdot L] = [K|\theta] \quad (\text{A5})$$

$$\Gamma[K|D_x v \cdot L] = [\theta|-L] \quad (\text{A6})$$

Transposing the matrices, reversing the order of the matrix multiplications (which is necessary because the matrices have been transposed) and combining the two equations produces the following equation:

$$\begin{bmatrix} K^T \\ \varepsilon_m \end{bmatrix} [C^T|\Gamma^T] = \begin{bmatrix} K^T & 0 \\ 0 & -L^T \end{bmatrix} \quad (\text{A7})$$

APPENDIX 2

MetaCon's algorithm for Control Analysis

As shown in Appendix 1, the elasticity matrix and the M matrix can be derived from the matrices K , L and $D_x v$. Thus, all of the information for the Control Analysis can be obtained from the reaction scheme: $D_x v$ is obtained by automatically assigning elasticities to the substrate(s) and product(s) of each reaction, plus any additional elasticity effects which exist; K and L are obtainable from the stoichiometry matrix of the system, which is itself determined directly from the reactions in the network. Most of the matrices used in the course of the control analysis e.g. stoichiometry matrix, L and K , can be printed to an output file by MetaCon, by including the relevant command in the input file. In the present Appendix the representation of all the matrices is the same as that used by MetaCon, including labelling of rows and columns where applicable.

Creation of the stoichiometry matrix

MetaCon parses (reads) the input file and creates the corresponding stoichiometry matrix [as defined by Reder (1988)], based on the information in the reaction scheme. The rows and

where ε_m has been used for brevity in place of $[D_x v \cdot L]^T$.

If C^T is restricted to the coefficients referring to control over the reference flux (i.e. C^T has only one column), and Γ^T to the coefficients referring to control over the independent metabolites, it can be seen that this equation is equivalent to the equation produced by the matrix method, except for the lack of branch-point distribution control coefficients. The upper partition of eqn. (A7):

$$[K^T][C^T|\Gamma^T] = [K^T|\theta] \quad (\text{A8})$$

contains the combined summation and branch-point rows, whereas the lower partition contains the connectivity rows:

$$[\varepsilon_m][C^T|\Gamma^T] = [\theta|-L] \quad (\text{A9})$$

where only the first m_0 columns of L^T have been used, which correspond to the identity matrix, as described above.

For any pathway, this technique produces a lower partition that is unique unless there are rows involving conserved metabolites, when the result depends on the choice of dependent metabolites. However, K is not unique for a pathway with branch-points, so eqn. (A8) may have several possible equivalent forms. In these cases, the Fell and Sauro elasticity matrix puts two restrictions on the acceptable form of K . The required form can be produced by taking certain linear combinations of the vectors of any K for the pathway. The restrictions are as follows.

1. One row must correspond to the summation row (i.e. a row of 1s).

2. The remaining (*nullity*-1) rows correspond to the branch-point rows, where *nullity* is the number of vectors in the nullspace basis. They must not contain any entries in the column representing the reference flux. Furthermore, the right-hand side of each of the branch-point equations must be 0.

These restrictions have been incorporated into the MetaCon algorithm, as described in Appendix 2.

REFERENCES

- Reder, C. (1988) J. Theor. Biol. **135**, 175–201
 Schuster, S. and Heinrich, R. (1992) BioSystems **27**, 1–15

columns of the matrix, representing the metabolites and steps respectively, are ordered simply in the order in which they are read from the reaction scheme in the input file:

	R1	R2	R3	R4	R5	R6
S1	1.00	-1.00	-1.00	0.00	0.00	0.00
S2	0.00	0.00	2.00	-1.00	0.00	0.00
NAD	0.00	0.00	0.00	-1.00	0.00	1.00
NADH	0.00	0.00	0.00	1.00	0.00	-1.00
S3	0.00	0.00	0.00	1.00	-1.00	0.00
S4	0.00	0.00	0.00	0.00	1.00	-1.00

Creation of the connectivity rows of the elasticity matrix

The next step of processing involves examining the stoichiometry matrix to see if the pathway contains conserved metabolites, as described in Appendix 1. For the example pathway in Scheme 1 of the main paper the L_0 partition is:

	S1	S2	NAD	S3
NADH	0.00	0.00	-1.00	0.00
S4	0.00	0.00	-1.00	-1.00

This matrix represents two conservation relationships which exist between the metabolites of the example pathway, and is equivalent to defining their time-dependencies by the following equations:

$$\frac{d}{dt}NADH = -\frac{d}{dt}NAD$$

$$\text{i.e. } \frac{d}{dt}(NAD + NADH) = 0$$

and

$$\frac{d}{dt}S4 = -\frac{d}{dt}NAD - \frac{d}{dt}S3$$

$$\text{i.e. } \frac{d}{dt}(S4 + NAD + S3) = 0$$

In this relatively simple example the relationships found by the algorithm can also be determined by inspection of the stoichiometry matrix, by observing that not all the rows are linearly independent: $\text{row}(NAD) = -\text{row}(NADH)$, and $\text{row}(S4) = -[\text{row}(S3) + \text{row}(NAD)]$.

As stated in Appendix 1, the ratios of the concentrations of pairs of metabolites which occur together in conservation relations appear in the connectivity rows of the elasticity matrix; the numerator is the 'independent' member of the pair (i.e. the one which does not have a row in the L_0 partition), the denominator one of the 'dependent' members of the conserved group. Thus from the L_0 partition for the pathway in Scheme 1 of the main paper it can be seen that the ratios which will occur are $NAD/NADH$, $NAD/S4$ and $S3/S4$. The ratios occur wherever there is a non-zero elasticity of any of the 'dependent' metabolites; in the normalized matrix they are multiplied by that elasticity. Thus the denominator of the concentration ratio is the same metabolite whose elasticity it multiplies. The effect on the structure of the normalized elasticity coefficient matrix can be seen in rows 4 and 5 of Matrix 1 in the main paper.

During decomposition of the stoichiometry matrix its rows are reordered, the first m_0 rows containing the independent metabolites, the remaining $m - m_0$ rows the dependent ones:

	R1	R2	R3	R4	R5	R6
S1	1.00	-1.00	-1.00	0.00	0.00	0.00
S2	0.00	0.00	2.00	-1.00	0.00	0.00
NAD	0.00	0.00	0.00	-1.00	0.00	1.00
S3	0.00	0.00	0.00	1.00	-1.00	0.00
NADH	0.00	0.00	0.00	1.00	0.00	-1.00
S4	0.00	0.00	0.00	0.00	1.00	-1.00

From this analysis, the connectivity rows, ϵ_m , can be generated, by normalizing the L matrix as described in Appendix 1: $\epsilon_m = D_x v_{(\text{norm})} \cdot L_{(\text{norm})}$. In the example of Matrix 1 of the main paper, this generates rows 2-5.

Creation of the summation row and branch-point rows of the elasticity matrix

As shown in Appendix 1, Reder (1988) has proved that the summation and branch-points in the elasticity matrix are equivalent

to a basis of the nullspace of the stoichiometry matrix. The algorithm used in MetaCon for the determination of the nullspace basis is a modification of one described by Knuth (1981), which also determines the rank of the matrix.

Although Reder showed that any basis for the nullspace implies a valid set of summation and branch-point relationships for the control coefficients, the Fell and Sauro (1985) matrix method follows Kacser and Burns (1973) and Heinrich and Rapoport (1974) in explicitly using the summation theorem of the flux and concentration control coefficients as the first row of the matrix equation. To create the matrix equations in the correct forms, linear combinations of the nullspace vectors can be used, and as MetaCon produces expressions for the normalized control coefficients, the nullspace can be normalized as shown in Appendix 1.

Normalizing the control coefficients and K , and summing all the resulting equations, produces the summation row. An example of this procedure for the flux control coefficients of a simple branched pathway can be found in Mazat et al. (1990). This process corresponds to choosing one of the nullspace vectors. To produce the branch-point rows requires a set of nullspace vectors with all entries in the row corresponding to the reference flux being equal to zero. The set of such vectors will therefore be independent of the summation row, as this has a non-zero entry in the reference-flux column.

The required vectors could be produced by subtracting multiples of any vector with a non-zero entry in the row (of which there must be at least one to satisfy the steady-state condition) from all other vectors which have a non-zero entry in the row. The result would be one vector with a non-zero value in the reference-flux row, and $\text{nullity}-1$ vectors each with a zero in the reference-flux row. The first vector would be ignored, and the remaining vectors used to create the branch-point rows of the Fell and Sauro (1985) matrix.

For the following nullspace basis of the example pathway, which is the one which would be created by MetaCon:

1.00	0.50
1.00	0.00
0.00	0.50
0.00	1.00
0.00	1.00
0.00	1.00

choosing the flux through R1 as the reference flux, subtracting (vector 1)/2 from vector 2 produces a new vector 2: $[0 \ -0.5 \ 0.5 \ 1 \ 1 \ 1]^T$. Normalizing the control coefficients and K produces the sixth row of Matrix 1. This method also assures that the right-hand sides of the equations are always 0.

This procedure is effected in MetaCon by deleting the reference-flux column from the stoichiometry matrix, to create a nullspace basis, K' , of the resulting matrix. K' has one vector less than K , and each vector has one less row, the missing row corresponding to the entry for the reference flux. For the example pathway, assuming that the flux through step R1 is the reference flux, K' is:

-0.50
0.50
1.00
1.00
1.00

where the rows represent entries for steps R2–R6. Completing the branch-point row(s) simply requires division of the entries in the vector(s) by the corresponding reaction flux. Thus the nullspace vector above corresponds to the following branch-point equation for the flux-control coefficients:

$$\frac{-0.5C_{R2}^{J_{R1}}}{J_{R2}} + \frac{0.5C_{R3}^{J_{R1}}}{J_{R3}} + \frac{C_{R4}^{J_{R1}}}{J_{R4}} + \frac{C_{R5}^{J_{R1}}}{J_{R5}} + \frac{C_{R6}^{J_{R1}}}{J_{R6}} = 0 \quad (\text{A10})$$

APPENDIX 3

Checks for pathway validity and integrity

In this Appendix are described some of the properties of the various matrices created by MetaCon during the course of a Control Analysis, which MetaCon uses to check for the validity of the network. If any of the required conditions are not met, a suitable warning message is produced, and the program terminates.

1. As the stoichiometry matrix does not contain rows for pool metabolites, any reaction which includes only pool metabolites will produce a column of zeroes in the stoichiometry matrix. Such a reaction has no meaning in the reaction scheme, but could arise, for example, by confusing a pool and a variable metabolite when entering the scheme.

2. The requirement for a steady state to be attainable is that the nullity must be at least 1. Thus for any properly linked reaction scheme the following relationship is true:

$$n - m_0 > 0 \quad (\text{A11})$$

REFERENCES

- Fell, D. A. and Sauro, H. M. (1985) *Eur. J. Biochem.* **148**, 555–561
 Heinrich, R. and Rapoport, T. A. (1974) *Eur. J. Biochem.* **42**, 89–95
 Kacser, H. and Burns, J. A. (1973) *Symp. Soc. Exp. Biol.* **27**, 65–104
 Knuth, D. E. (1981) *The Art of Computer Programming*, vol. 2: Seminumerical Algorithms, 2nd edn., Addison-Wesley Publishing Co., Reading, MA
 Mazat, J.-P., Reder, C. and Letellier, T. (1990) in *Control of Metabolic Processes* (Cornish-Bowden, A. and Cárdenas, M. L., eds.), pp. 129–138, Plenum Press, New York
 Reder, C. (1988) *J. Theor. Biol.* **135**, 175–201

Should this relationship be found not to hold for a reaction scheme, it is likely that the mistake has arisen in translating the original network into reactions for the input file. Possible causes are: (i) a pool metabolite has been inadvertently entered as a variable metabolite, or (ii) one or more reactions have been omitted, splitting the scheme into several separate parts.

3. For a step to be able to carry a flux in the steady state the corresponding row in the nullspace must have at least one non-zero entry. Thus, if any row in the nullspace has all zero entries, the corresponding step cannot carry a steady-state flux. A probable cause for failing this check is for the reaction in question to be in a 'dead-end' in the network, i.e. in a branch not leading to or from a pool metabolite. If this is part of the intended original scheme (i.e. it is not a result of making a mistake while entering the scheme), the scheme will have to be re-defined, but will also arise if a pool metabolite at the end of a branch is inadvertently entered as a variable one.