The Computer Users Group Workshop on Modelling Metabolic Systems
at the 124th Meeting of the Society for General Microbiology
at University of Kent, Canterbury, UK, 6 January 1993

Demonstration Programs Illustrating the Modelling of Metabolic Systems

Compiled by Douglas B. Kell with Hans Westerhoff[1], David Fell[2]
Simon Thomas[2] and Pedro Mendes

Biological Sciences, University of Wales, Aberystwyth, Dyfed SY23 3DA UK
Electronic mail: DBK@ABER. AC.UK and
PRM@ABER. AC.UK
[1]Nederlands Kanker Instituut, H5, Plesmanlaan 121, NL1066 CX Amsterdam
The Netherlands
Electronic mail: HW@NKI.NL
[2]School of Biological and Molecular Sciences, Oxford Brookes University,
Gipsy Lane, Oxford OX3 0BP, UK
Electronic mail: DFELL@BROOKES.AC.UK and
S8980696@BROOKES.AC.UK

Douglas Kell (University of Wales, Aberystwyth, UK) organised the above workshop at the SGM Kent meeting. The main aims were to indicate to a relatively lay audience the advantages of producing mathematical models of the metabolic systems one was studying, and to illustrate (with practical demonstrations) some of the many computer programs which are now available for doing this in a relatvely painless fashion. The workshop intentionally followed the Genetics & Molecular Biology and the Physiology & Biochemistry Groups' joint Symposium on ***Genetic regulation of Metabolic Flux.*** *It consisted of an overview lecture by Dr Hans Westerhoff (Nederlands Kanker Instituut, Amsterdam, The Netherlands) entitled* ***Modelling Metabolism; What, Why and How?*** *followed by a number of demonstrations of programs outlined below, by Hans Westerhoff, David Fell and Simon Thomas (Oxford Brookes University, UK) and Pedro Mendes (University of Wales, Aberystwyth, UK). What follows was contributed by them and subjected only to minor editing. The local organizer was Alan Bunch, Biological Laboratories, University of Kent, UK who cheerfully and graciously ensured that we were provided with all necessary computing requirements, and plied with wine by courtesy of the SGM Computer Users Group.*

Three generations of programs that are used to simulate biological systems were distinguished by Hans Westerhoff. The oldest generation is characterized by the fact that both the program language and its communication with the user are in mathematical terms. That is, the program statements are formulated as mathematical operations between real or integer numbers and the user has to describe the biological system in terms of mathematical equations. These would typically be the differential equations describing how the time dependence of the metabolite concentrations depends on those concentrations. Examples of such languages are FORTRAN and C. Languages such as MLAB, MatLab and MAPLE belong in the same category, although they have the important advantage that many mathematical operations (such as an integration) do not have to be programmed in terms of primitive mathematical operations but can be invoked by a single command.

The second generation of simulation languages is still structured in mathematical terms, but its user interface has been adjusted to the needs of the biologist. Consequently, the user may describe the biological problem of interest in terms of actual processes which change the concentrations of metabolites. The program will then automatically integrate, determine steady states, or control properties. Examples of this generation (in approximate order of increasing devotion to biology) are: STELLA, SCoP, ESSYNS, METAMOD, Meta Model, MetaCon (which specializes in control properties), SCAMP and Gepasi. MetaCon and Gepasi are described later. These simulation programs are considered to have the advantage that they are readily understandable for the layman, reasonably well debugged and that in many cases the authors are still actively interested in their application to biological and metabolic problems.

The third generation of programs is not yet at the stage of being widely available and fully debugged. They have been programmed in a language that is itself couched in terms of the biological object. That is, instead of real numbers and integer numbers, enzymes, and metabolites are used and instead of multiplication, reaction. This is possible because of the development of object-oriented programming languages such as C++. An examples is the program METASIM. An advantage is that the structure of the program itself is now directly related to the structure of the biological system that is being programmed, such that the program may be actually complete, yet understandable by its biological user. This could reduce errors arising from communication failures between programmers and biologists. Disadvantages are that these programs cannot directly implement all the best available mathematical tools that have been developed over the centuries, precisely because the tools have been formulated in mathematical rather than biological terms. In addition, the program is still at the testing stage.

Examples of simulations using languages from each of these generations were shown. The following are typical statements in an MLAB

0226-304X/93/020047+03 $03.00/0

BINARY Vol. 5 © 1993 BioLine

programme designed to simulate ATP production during (part of) glycolysis

```
function v1(ATP,G)=k1*ATP*G;
function v3(ATP,I,H)=k3*(ADP)*I*(Pi)*NAD;
function ATP'T(t)=-2*v1+4*v3;
m=integrate (ATP'T, t0:tend:tstep);
draw m in window;
view window;
```

MLAB is a commercial programs, and is available from Civilized Software, MD 20814, USA (Tel: 301 652 4714).

In the SCoP program the interface is already in terms of biologically relevant properties. There is one block in which one has to define the constants, one in which one has to state the independent variable (usually time), one in which the metabolite concentrations must be defined, one in which any conservation conditions must be mentioned, and then one in which the equations are mentioned. This is all done in a file, which is then used in a compilation procedure to make a fast simulation program. SCoP is also a commercial program, and is available from Mailen Kootsey, NBSR, Duke University Medical Center, Durham, NC 27710, USA.

In the object oriented METASIM program (Stoffers *et al.*, 1992 and available from Hans Westerhoff) the statements look like:

```
substance       DNA, mRNAhis3, His3
Irreversibleenzyme  His3DNAtranscription,
                    His3mRNAtranslation,
simulator       time(0, 15, 10~-1)
while (time) {X.output}
X.output
```

where the biological objects are recognized in the program.

Gepasi

Gepasi (as in GEneral PAthway SImulator), written and demonstrated by Pedro Mendes, University of Wales, Aberystwyth, UK is a software package for the simulation of the dynamics, steady-state and control analysis of metabolic pathways. The simulations produce values for the concentrations of metabolites and the magnitudes of fluxes starting from an initial state. The parameters of these models are the kinetic constants of the steps and the concentrations of metabolites that are either fixed (buffered) or constantly flowing into or out of the system. The user interaction is handled by a program that runs under MS *Windows* (presently on IBM PC-compatible computers). This program makes extensive use of menus, dialogue boxes, push buttons, and other controls. This form of input, accompanied by keyboard short cuts, minimizes the time taken by a first-time user to get accustomed to the mechanics of the program. By taking advantage of MS *Windows* own help engine, this front-end program has an extensive help system that covers not only immediate instructions on how to use the package but also has explanations of common concepts in metabolic control, for example, those of internal or external metabolites. There is also a section with full references to articles, reviews and books covering subjects related to metabolic control and modelling. It is considered that this feature should ensure that the package will be a useful tool for education.

One feature of Gepasi that is particularly useful for an extensive study of a model is the ability to scan various parameters. The simulator will produce several simulations in sequence (each with different values for those parameters) and put the output on different rows of the same file, effectively producing a map of the behaviour of the model (its variables) in a region of parameter space. Virtually all parameters of the model can be scanned (but unless the model is small, scanning all would become prohibitive). These 'scans' can be done with linear or logarithmic intervals: lower and upper limits and the number of intervals must specified for each parameter. They can also be done by assigning random values (with linearly or logarithmically uniform distributions) for the parameters within lower and upper limits. In this case a total number of simulations is selected.

Gepasi runs on personal computers with MS DOS version 3.2 or above and MS *Windows* version 3.1 in enhanced mode. It is available from the author, who can be contacted via electronic mail to PRM@ABER.AC.UK. The package is supplied by post only if suitably formatted diskettes are sent to the author (with a minimum of 1.2 Mbytes free). In JANET, the latest version of the package is available for downloading from the UKUUG archive at UK.AC.IC.DOC.SRC in the directory /packages/ibmpc/simtel20/biology. In the US it can be obtained by anonymous FTP from wsmr-simtel20.army.mil in the directory <MSDOS.BIOLOGY>.

MetaCon

MetaCon (Thomas & Fell, 1993; demonstrated by David Fell, Oxford Brookes University) is a computer program for the evaluation of the 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 automatically be 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 (Fell & Sauro, 1985; Sauro *et al.*, 1987). The vast majority of MetaCon's input and output is from and to simple text (ASCII) files. The only exceptions are a few simple messages, and sometimes prompts to the user which require keyboard input. Processing commands in the input file are in an understandable, English-like language.

The stages of processing carried out by MetaCon are to:

1. Analyse the pathway, and create the corresponding elasticity matrix (a term used in the matrix method), which can be written to an output file. This step is the actual Metabolic Control Analysis. The variables in the elasticity matrix are:
   i. elasticities,
   ii. pathway fluxes, in branched pathways,

iii. metabolite concentrations, in pathways with conserved groups of metabolites. The variables in the elasticity matrix represent the set which must be measured to calculate the control coefficients.

2 Solve the matrix equation to produce algebraic expressions (they are multivariate polynomials) for the control coefficients. The flux control coefficients of all steps in the pathway on one specified flux (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, C/v, and scaled sensitivities

$$(\partial C/\partial v).(v/C).$$

The sensitivity analysis is carried out algebraically, to produce expressions of the same form as those for the control coefficients.

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 - 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 was intended to be used mainly by experimentalists working in Metabolic Control Analysis, as it

1. automates all stages of the calculations of the control coefficients, and
2. can help in experimental design by informing the user which variables are to be measured and by carrying out a sensitivity analysis on the values of the control coefficients. Because processing is algebraic, useful information about the control structure of a pathway can be obtained before the values of all the variables have been determined. However these features should also prove to be useful to workers who are investigating the control of metabolic pathways theoretically, rather than experimentally.

MetaCon runs under DOS on IBM-compatible PCs. A version to run on a DEC 2100 workstation (UNIX) is also being implemented. It is distributed under the GNU General Public Licence Version 1. The DOS version is obtainable by sending a formatted 3.5 inch disk to the authors Simon Thomas and David A. Fell, at the School of Biological and Molecular Sciences, Oxford Brookes University, Gipsy Lane, Oxford OX3 0BP, UK. The executable program is distributed with a comprehensive manual (MS Word for *Windows* 2 is needed to read it; a less readable ASCII version is also distributed), and sample input and output files.

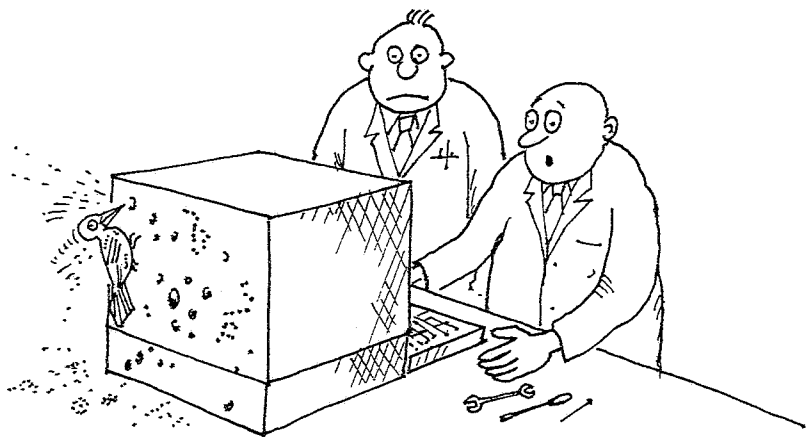
References

FELL, D.A. & SAURO, H.M. (1985). *European Journal of Biochemistry* **148**, 555-561.

SAURO, H.M., SMALL, J.R. & FELL, D.A. (1987). *European Journal of Biochemistry* **165**, 215-221.

STOFFERS, H.J., SONNHAMMER, E.L.L., BLOMMESTIJN, G.J.F., RAAT, N.J.H. & WESTERHOFF, H.V. (1992). METASIM: object-oriented modelling of cell regulation. *Computer Applications in the Bioscience* **8**, 443-449. Available from Hans Westerhoff (Fax: 31 20 5122029; email: HW@NKI.NL).

THOMAS, S. & FELL, D.A. (1993). *Biochemiical Journal*. In press.



"No, I can still hear it!"