Teaching Metabolic Control Analysis and kinetic modelling

Towards a portable teaching module

by Jacky Snoep, Pedro Mendes and Hans Westerhoff

Not just another course!

In the traditional curriculum for biology students a gap separates the more quantitative and molecular courses in thermodynamics and biochemistry from the courses on higher organizational levels, such as microbial and plant physiology. For biologists, training in thermodynamics, kinetics, systems analysis and mathematics is limited. Furthermore most chemistry courses treat thermodynamics in a classic way, e.g. by basing it on steam machines; they do not relate it to the non-equilibrium features so crucial for biological systems and it thus remains very abstract to the students. Biochemistry courses are focused on isolated enzymes and tend to limit kinetics to the irreversible Michaelis–Menten type. Physiology courses may deal with whole pathways, organelles and cells, but do so in a qualitative way. The mathematics taught to biologists is heavily biased towards statistical analysis.

The lack of quantitative courses linking cell physiology to molecular biology could remain relatively unnoticed for as long as biochemistry and physiology were directed exclusively towards their own scientific niches. However, complete genomes have been sequenced, crystal structures appear at high rates and proteins have become over-expressable and purifiable with unprecedented ease. Rather than addressing the same type of molecular question again and again for a decreasing number of truly new proteins, some biochemists are turning to the challenge of discovering how proteins function *in vivo* and how they produce the complex phenomena that constitute the living cell. Indeed, this turning may well be massive with functional analyses of sequenced genomes becoming big science[1].

The complex behaviour of the living cell arises from the non-linear interactions of many molecules. The result of any combination of non-linear interactions cannot be predicted by simple extrapolation or interpolation. Only quantitative mathematical analyses, mostly even beyond what is possible analytically, are a prerequisite for understanding the living cell. If 21st-century biochemistry is to bring us to the understanding of life in terms of physics and chemistry, then it will need 21st-century biochemists that are at ease with mathematical modelling.

A course that shows the relation between the different disciplines and links them together is therefore necessary. At the Free University of Amsterdam and the University of Amsterdam, two of us have been teaching courses with this aim for several years. At the University of Amsterdam, Mathematical Biochemistry is a 1 month third-year course and is biased towards a statistical thermodynamic approach. At the Free University the courses Microbial Physiology, for second-year students, and Molecular Cell Physiology, for third-year students, are more focused on a quantitative approach to cell physiology. The aim of the course described in this article should be to bridge the gap between these different disciplines and to give the students the means to come to a quantitative understanding of biological pathways/systems. In international workshops this course has been given several times at the graduate-student level.

Kinetic modelling

Traditionally, kinetics have been taught in biochemistry courses in terms of enzyme steady-state kinetics. This corresponds to a detailed study of the local properties of the individual enzymes. However, one can go further and create kinetic models of whole pathways. Such models are composed of coupled ordinary differential (for time courses) or algebraic (for steady states) equations. These equations are non-linear and most often without analytical solution. This means that they can only be studied through numerical algo-

rithms, such as the Newton method for solving non-linear equations and numerical integrators. Biochemical systems are also rich in time scales and thus require sophisticated methods for the numerical solution of the differential equations that describe them.

Kinetic modelling of metabolic pathways may be carried out by computer software that simulates the behaviour of a real pathway. Simulation then resembles a true experiment: one sets the initial concentration of the metabolites and the software then produces the time evolution and/or the steady state of these concentrations. There is, however, an additional amount of information to be supplied in the case of simulation: the differential equations describing the kinetics of the pathway and values for all the parameters involved in these equations.

There are programs that minimize the mathematical effort of setting up these equations. Advantages of simulation over experiment are: (i) it is much cheaper; (ii) it is faster; (iii) students can investigate extreme conditions that are very hard to achieve in the lab; (iv) simulation programs supply the concentrations of all intermediates and all the fluxes, which are very hard to determine experimentally, especially by undergraduates; (v) unnecessary harm to animals is reduced; and (vi) occasionally, simulations can be done for experiments that are impossible with existing technologies. On the other hand, simulations have the disadvantage that the students do not have direct contact with the real biochemical system.

Compared with more conventional ways of learning, such as reading textbooks or watching video programmes, simulation has the advantage that it is interactive. With simulation the student is allowed to discover concepts by her/himself. Students with very inquisitive minds can progress further than expected because they can continue to 'experiment' with the model. In other more passive media these concepts are explained and the student takes a passive role.

We see simulation as a very useful complement to experiments, perhaps being able to reduce their number while still covering all the subjects. This well may be a very positive thing in these days when departments are being forced to decrease their spending and increase their scientific production.

Metabolic Control Analysis (MCA)

MCA is a system-level approach to the study of metabolism. It has become a standard way of analysing metabolism quantitatively. The central concept in MCA is that of the control coefficient: a quantitative measure of the extent to which the activity of a single enzyme determines the pathway flux or internal metabolite concentrations. The flux control coefficients are the rigorous way to quantify flux limitation. Control coefficients can be determined directly by measuring the effect of single-enzyme perturbations on the system, i.e. without detailed knowledge of the kinetics of those individual enzymes. One of the most important contributions of MCA is that control (rate limitation) is now seen as a property of the whole system, rather than as a property of the individual enzymes.

The elasticity coefficients quantify how much the rate of an enzyme isolated from the pathway is affected by substrates, products and other effectors. Thus the elasticity coefficients reflect the kinetics of the enzyme (they are scaled partial derivatives of the kinetic rate equations with respect to metabolite concentrations). MCA, through the summation and connectivity theorems, also allows one to relate the (global) systemic properties of the pathway to the (local) properties of the individual enzymes. Thus, from a detailed knowledge of the kinetics of all the pathway's enzymes, one can calculate the control distribution, linking the system behaviour to the individual steps. This is exactly what the computer programs mentioned below do.

MCA has been extended to deal with basically all types of metabolic structure, including hierarchical systems (such as when one considers transcription and translation or signal-transduction cascades), and when there are enzyme–enzyme interactions. How much detail of the theory should be included in an undergraduate-level course is debatable. We think that it is better to concentrate on the main concepts and perhaps leave details out. Of central importance are concepts such as: (i) rate limitation as a quantitative dynamic property; (ii) control being distributed; and (iii) the control exerted by one step being dependent on the state of the system (e.g. over-expressing an enzyme usually leads to a reduction of its control over the pathway).

Software

There are several software packages that can be used for modelling kinetics and control analysis of biochemical systems. Which software to choose should be determined by the objectives and perhaps the budget of the course. There are basically three types of package that can be used in this context: generic mathematical modelling software; metabolic simulators driven by specific languages; and metabolic simulators with graphical user-interfaces (GUIs).

General mathematical modelling software requires the student to write the mathematical model explicitly and to choose the appropriate numerical method for its solution. These programs are useful in courses that deal themselves with computer simulation, as the student will have to master the mathematics in order to run a simulation successfully. They are not recommended for courses that use simulation simply as a tool to teach biochemistry. Students would need to study calculus, matrix algebra, numerical analysis and the programming language that drives such software. Commercial packages such as Mathematica (Wolfram Research, Inc., Champaign, IL, U.S.A.), Matlab (The MathWorks, Natick, MA, U.S.A.) and ModelMaker (Charwell Scientific Publishing, Oxford, U.K.) belong to this category. There are also free packages of this type on the Internet, a popular example being GNU Octave (http://bevo.che.wisc.edu/octave/).

Some software packages are dedicated metabolic simulators and therefore require less mathematical knowledge. They already include the appropriate numerical methods but the student still has to write down kinetic equations for each step. These packages are driven by a specific programming language. Even though these languages are closer to biochemistry by the use of the appropriate nomenclature (e.g. 'concentration' instead of 'variable'), the students still have to learn a rigid syntax and have to write 'metabolic' programs. The free program SCAMP[2] and the commercial SCoP (Simulation Resources, Inc., Redlands, CA, U.S.A.) belong to this category.

A third class of software package is based on a GUI front end, by which the student defines and runs the simulation. These packages have predefined rate equations for the most common kinetic functions and they may even hide the existence of the differential equations and matrices. This means that the student can define and run a metabolic model with a minimal knowledge of the mathematics *or* computer programming. This is indeed recommended for courses where the emphasis is on the biochemistry. This class of software is composed mainly of free packages: Gepasi[3] (see http://gepasi.dbs.aber.ac.uk/softw/Gepasi.html), MIST[4] and DBSolve[5].

Modelling and teaching

The structure of our courses is such that lectures are given in the morning and in the afternoon the students apply the theory in computer simulations. The emphasis is on the integration of single enzymes into a system. The course is built up starting with reversible enzyme kinetics, essential for a system containing more than one enzyme. A link between thermodynamics and kinetics is made by deriving the Haldane equation and by showing the relation between the kinetic constants, K_{eq}, ΔG^{o} and ΔG, and the driving forces for the reaction. Some of these principles are shown using a very simple model with a pathway of three identical enzymes. Subsequently, the characteristics of the isolated enzymes are translated to the system. At this point MCA is introduced as a tool for studying and understanding systems.

Starting from the simple model, we introduce a moiety-conserved cycle and then progress to a branched model and finally to a branched model with moiety conservation. The latter is a misleadingly simple model that leads to counterintuitive results. Among other things it serves to illustrate the need for quantitative analysis. After this set of 'simple' models, some more applied and specific types of model are introduced. These models relate to dynamic systems, growing systems, biological sub-systems in a non-biological environment, optimization principles or realistic models (for instance, glycolysis).

In each of the models a new principle is introduced or attention is given to a specific aspect related to the morning lectures. The latter cover topics such as steady state, time evolution, parameter versus variable, units, structured versus unstructured models, volumes in models, subsystems in systems, e.g. microorganisms grown in a chemostat, and optimization. Emphasis lies on steady-state systems but some dynamic behaviour is also studied (steady oscillations).

The combination of modelling and theory has a number of advantages. (i) The learning process is an active one and this results in a better understanding and more definitive memorization. Furthermore, the students will know themselves whether or not they understood the lectures and, by actively performing the simulation experiments, will find out how the system works. (ii) The students can work at their own speed and get to an understanding at the level they can cope with. During the modelling sessions the students use a handout with exercises and questions. These questions can be answered at various levels, and in subsequent sessions each of the questions is treated in detail. (iii) In contrast with experiments the simulations are quick and always work. They are ideal for showing principles.

Logistics of the course

The teacher/student ratio is crucial in courses such as these, which, computer notwithstanding, rely on direct contact. The morning sessions can be given as

a formal lecture, although a tutorial-group format is preferable. The modelling sessions, however, rely heavily on interaction between the students and the teachers. Although the students can in principle work independently, it is important to have close contact to overcome trivial problems such as typographic errors, as can occur in certain modelling programs, or to help the students to come to a deeper understanding of the problem, e.g. through answering the question in a more analytical, quantitative way. The courses are generally considered to be rather difficult and can best be given in the third year of undergraduate studies or as a PhD course. This makes it easier to limit the number of students, but foregoes the aim of reaching all biologists. The ideal number of students lies around 25. The maximal number of students per computer is three; the optimum two and an experienced instructor can assist some 12 students. The simulations are usually not very computer-intensive, i.e. simple models are used to illustrate a principle (although some of the more realistic and detailed models can become heavier on the computer). The software used to do the simulation is often what determines the hardware/software platform requirements. SCAMP is a DOS-based program, but can run under Windows 3.1/95/NT, whereas Gepasi runs only in 32-bit environments such as Windows 95 and Windows NT.

Course material consists of a summary of the lectures and copies of the transparencies. Furthermore, depending on the background and emphasis of the course, a choice of three books can serve as extended reading and examination material. These books are *Understanding the Control of Metabolism* by D.A. Fell[6], *The Regulation of Cellular Systems* by Heinrich and Schuster[7] and *Fundamentals of Enzyme Kinetics* by A. Cornish-Bowden[8]. The former two books deal with systems and how they can be understood on the basis of kinetics and regulation. Fell's book[6] concentrates on an experimental approach and is aimed at a broader range of readers, whereas Heinrich and Schuster's[7] contains more theoretical detail and some modelling and is for the more mathematically inclined students. Cornish-Bowden's book[8] is, as its title indicates, mainly about enzyme kinetics, but it makes the link to MCA and multi-enzyme systems. For reference to fundamental principles we have Hill[9,10] and Westerhoff and Van Dam[11] on standby.

Examination of the students is in a written form and focuses on the lecture material. One of the questions has a simulation experiment in which the students have to use the computer to get to an answer.

Students' impressions

In evaluations the course is well received by the students. They find the course difficult; in particular, biologists with a molecular-biological background experience problems with the mathematical approach. It may be important to have the course at the end of the curriculum. The students will then have been introduced to the different disciplines and are ready for the more integral approach at a slightly higher level of abstraction ●

References

1. Oliver, S.G. (1996) Nature (London) **379**, 597–600
2. Sauro, H.M. (1993) Comput. Appl. Biosci. **9**, 441–450
3. Mendes, P. (1997) Trends Biochem. Sci. **22**, 361–363
4. Ehlde, M. and Zacchi, G. (1995) Comput. Appl. Biosci. **11**, 201–207
5. Goryanin, I. (1996) in Biothermokinetics of the Living Cell (Westerhoff, H.V., Snoep, J.L., Sluse, F.E., Wijker, J.E. and Kholodenko, B.N., eds.), BioThermoKinetics Press, Amsterdam
6. Fell, D.A. (1997) Understanding the Control of Metabolism, Portland Press, London
7. Heinrich, R. and Schuster, S. (1996) The Regulation of Cellular Systems, Chapman and Hall, New York
8. Cornish-Bowden, A. (1995) Fundamentals of Enzyme Kinetics, Portland Press, London
9. Hill, T.L. (1977) Free Energy Transduction in Biology, Academic Press, New York
10. Hill, T.L. (1985) Cooperativity Theory in Biochemistry, Springer Verlag, New York
11. Westerhoff, H.V. and Van Dam, K. (1987) Thermodynamics and Control of Biological Free Energy Transduction, Elsevier, Amsterdam

Hans V. Westerhoff engages in Microbial Physiology at the Free University of Amsterdam, and in Mathematical Biochemistry at the University of Amsterdam. He hopes to help elucidate how direct and indirect interactions between macromolecules give rise to the properties that distinguish a living cell from dead molecules. This 'BioComplexity' interest is focused into the BioCentrum Amsterdam goal of 'Entering the Living Cell'. hw@bio.vu.nl

Pedro Mendes is a Research Associate in the Institute of Biological Sciences of the University of Wales, in Aberystwyth. His scientific interests are in computer simulation, kinetics of biochemical pathways and Metabolic Control Analysis. He is the author of the metabolic simulation program Gepasi. Pedro will soon join the National Center for Genome Resources in Santa Fe, New Mexico, USA. prm@aber.ac.uk

Jacky Snoep is Assistant Professor in the Department of Molecular Cell Physiology of the Free University in Amsterdam. His main interest lies in reaching a quantitative understanding of the living cell. This involves experimental as well as theoretical work in a physics as well as a biology context. Recently his scope has widened to structural dynamics of DNA supercoiling. js@bio.vu.nl