

COMMENTARY

# Models of the human metabolic network: aiming to reconcile metabolomics and genomics

Philip W Kuchel\*

## Abstract

The metabolic syndrome, inborn errors of metabolism, and drug-induced changes to metabolic states all bring about a seemingly bewildering array of alterations in metabolite concentrations; these often occur in tissues and cells that are distant from those containing the primary biochemical lesion. How is it possible to collect sufficient biochemical information from a patient to enable us to work backwards and pinpoint the primary lesion, and possibly treat it in this whole human metabolic network? Potential analyses have benefited from modern methods such as ultra-high-pressure liquid chromatography, mass spectrometry, nuclear magnetic resonance spectroscopy, and more. A yet greater challenge is the prediction of outcomes of possible modern therapies using drugs and genetic engineering. This exposes the notion of viewing metabolism from a completely different perspective, with focus on the enzymes, regulators, and structural elements that are encoded by genes that specify the amino acid sequences, and hence encode the various interactions, be they regulatory or catalytic. The mainstream view of metabolism is being challenged, so we discuss here the reconciling of traditionally quantitative chemocentric metabolism with the seemingly 'parameter-free' genomic description, and vice versa.

## Clash of giants: relative complexity of metabolic pathways and genomes

There are approximately ten times as many expressed genes (proteins) as there are different metabolites in most cells. Biochemical analysis of cells has been the art of the possible; you know about what you can detect. In the past, assays have largely focused on small organic (bio)

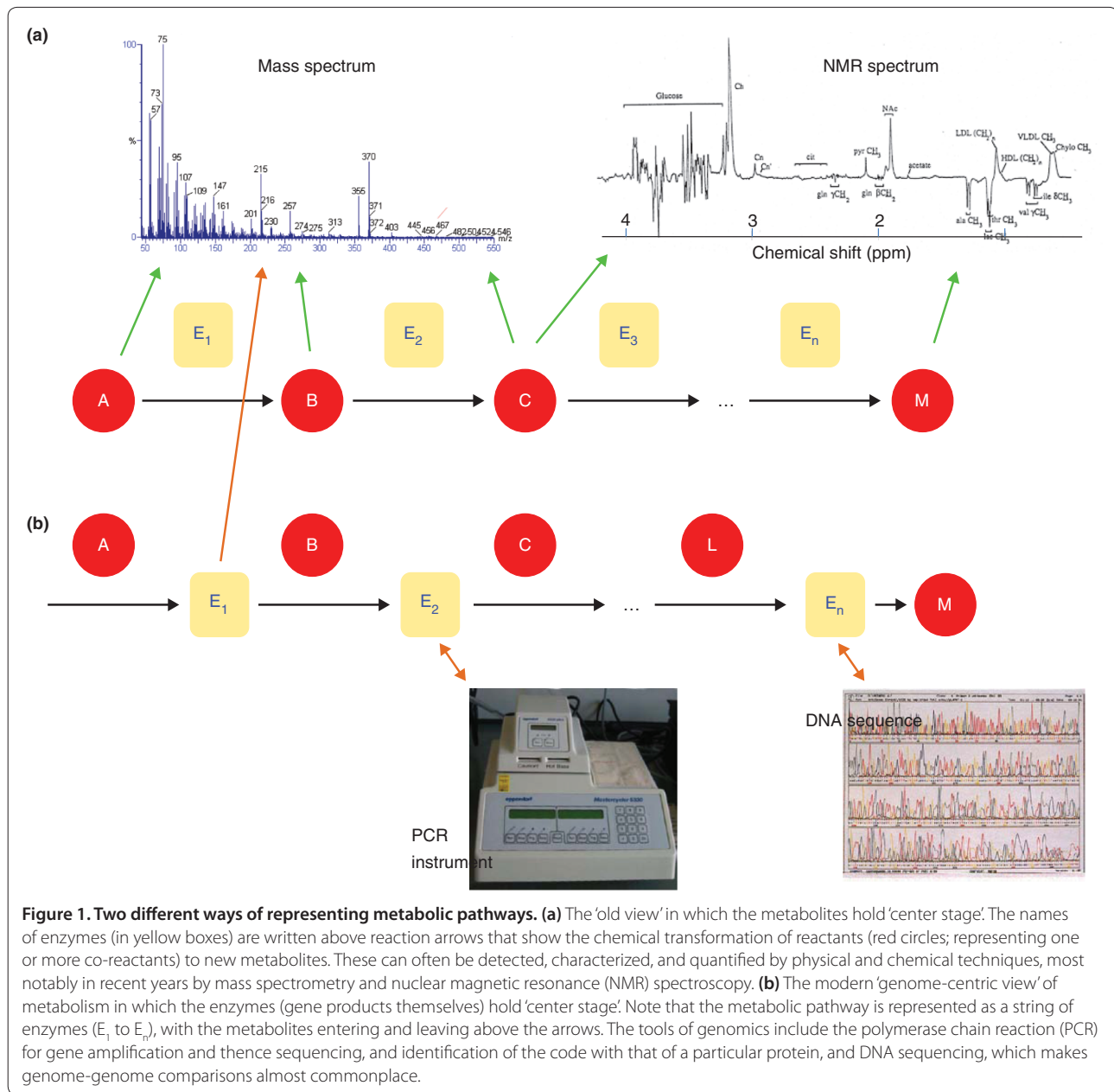
molecules analyzed by colorimetry or spectrophotometry. The genome projects have revealed a completely different data set from that of classical metabolic biochemistry, and a totally different perspective on metabolism. Two different perspectives, as neatly presented by Gerrard *et al.* [1], are presented in Figure 1; note how the genome draws attention to the proteins, many of which are enzymes, but many of which are not. So, measuring the concentrations of metabolites as we do in clinical biochemistry only indirectly reports on which of the enzymes, control proteins, or structural proteins are at fault in a case of chemical poisoning, drug side-effects, or in an inborn error of metabolism.

Figure 2 reminds us that there are at least 5,000 different enzymes, with as many metabolites in pathways that interconvert molecules in well-ordered sequences of reactions in an 'average' human cell. Figure 3 emphasizes that any one metabolite (denoted  $\gamma$  in this case) can modulate reactions from within its own pathway, across pathways, and even alters expression of genes and translation of messenger RNA into protein. An enzyme can also serve to modulate the activity of another enzyme, and affect its level of expression. Cations, including  $H^+$ , and extraneous compounds such as xenobiotics ( $H$  in Figure 3), also exert effects on enzymes and metabolites that potentially affect fluxes through multiple pathways.

## Traditional clinical biochemistry versus metabolomics

A modern and emerging form of advanced diagnostic strategy in chemical pathology is metabolomics, also called metabonomics [2]. There is a semantic and operational difference between these 'omics'. The former is the study of an extensive collection of metabolites present in a cell or tissue under a particular set of conditions (the metabolome) generating a biochemical profile. The latter involves the same profiling but in response to an influence (drug, toxin, or genetic defect) and then prediction of metabolic pathway(s) for the process(es). The approaches adopt an overview strategy that is superficially described as 'fingerprinting'. The investigator does not need to have a preconceived notion of what the metabolic problem might be with a patient because the methodology is

\*Correspondence: [philip.kuchel@sydney.edu.au](mailto:philip.kuchel@sydney.edu.au)  
School of Molecular Bioscience, University of Sydney, NSW 2006, Australia; Centre for Mathematical Biology, University of Sydney, NSW 2006, Australia



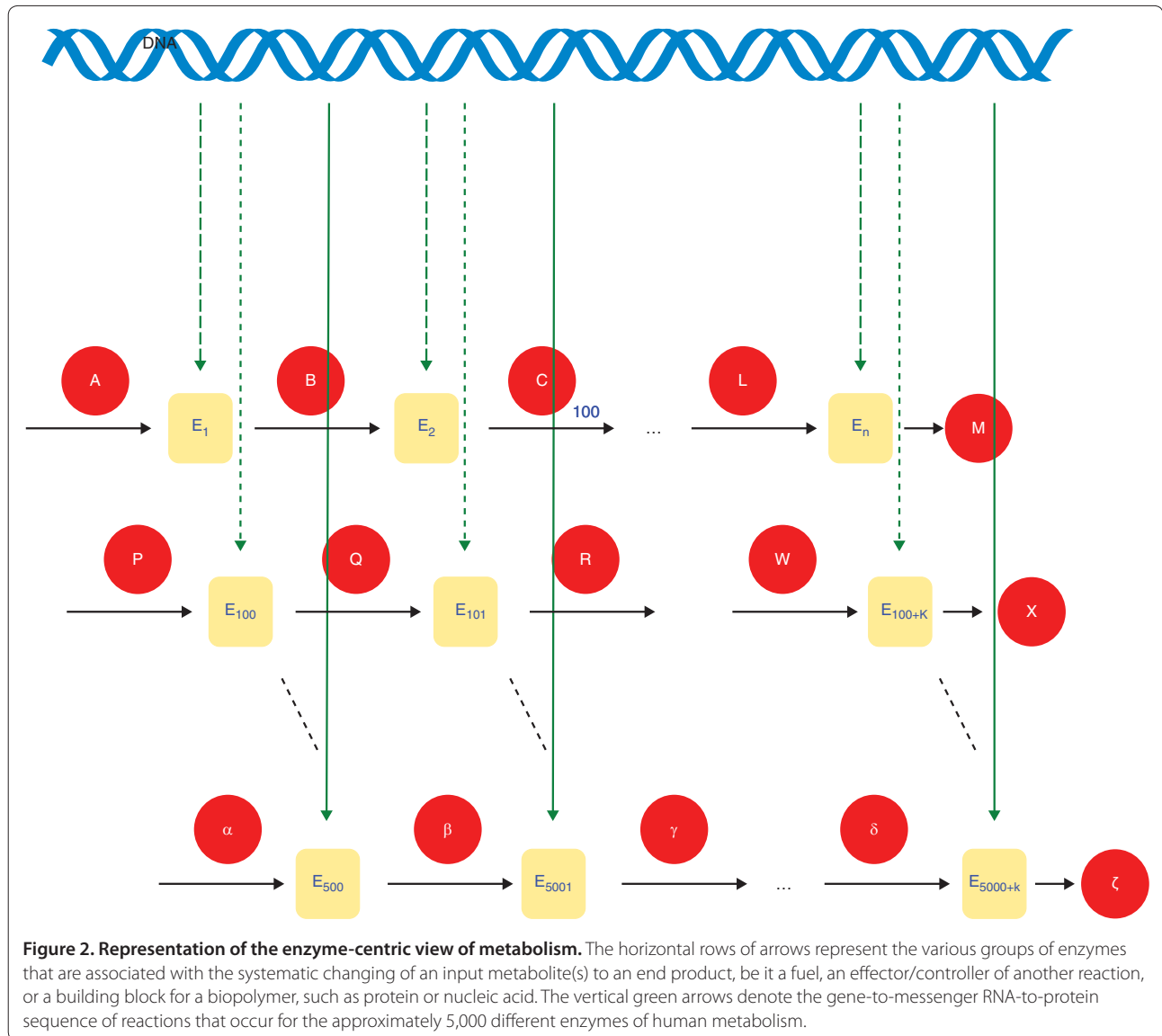
**Figure 1. Two different ways of representing metabolic pathways. (a)** The ‘old view’ in which the metabolites hold ‘center stage’. The names of enzymes (in yellow boxes) are written above reaction arrows that show the chemical transformation of reactants (red circles; representing one or more co-reactants) to new metabolites. These can often be detected, characterized, and quantified by physical and chemical techniques, most notably in recent years by mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy. **(b)** The modern ‘genome-centric view’ of metabolism in which the enzymes (gene products themselves) hold ‘center stage’. Note that the metabolic pathway is represented as a string of enzymes (E<sub>1</sub> to E<sub>n</sub>), with the metabolites entering and leaving above the arrows. The tools of genomics include the polymerase chain reaction (PCR) for gene amplification and thence sequencing, and identification of the code with that of a particular protein, and DNA sequencing, which makes genome-genome comparisons almost commonplace.

non-selective for particular metabolites, and yet specifically detects a broad range of them. In contrast, what has traditionally been done in clinical biochemistry is to work with a diagnostic hypothesis because only a limited set of tests exists to apply to a patient’s blood, or biopsy tissue, to help make a diagnosis. So focus is placed on a biochemical system; if the test points in a particular direction of enquiry, then another test is ordered, and so forth. Not so with the metabol(n)omics ‘shotgun’ approach!

Now that genes can be inserted into cells to correct metabolic defects in animals (for example, [3]), and presumably ultimately in humans, it will be important to be able to predict and monitor the metabolic consequences

of these genetic manipulations, thus bringing together the two paradigms: namely delineating metabolism by perturbing it with small molecules such as toxins and drugs, and perturbing it by manipulating gene expression, thus affecting enzyme activities.

To elaborate on the previous point, ‘Will the insertion of a “good” gene into a baby who has inherited a defective gene lead to them having a normal life?’ On contemplating this point, it becomes obvious that: (1) the gene must be able to be targeted to those tissues where it usually functions; (2) it must be delivered in sufficient quantities to transform a large enough fraction of the cells in the tissues to a normal state with normal

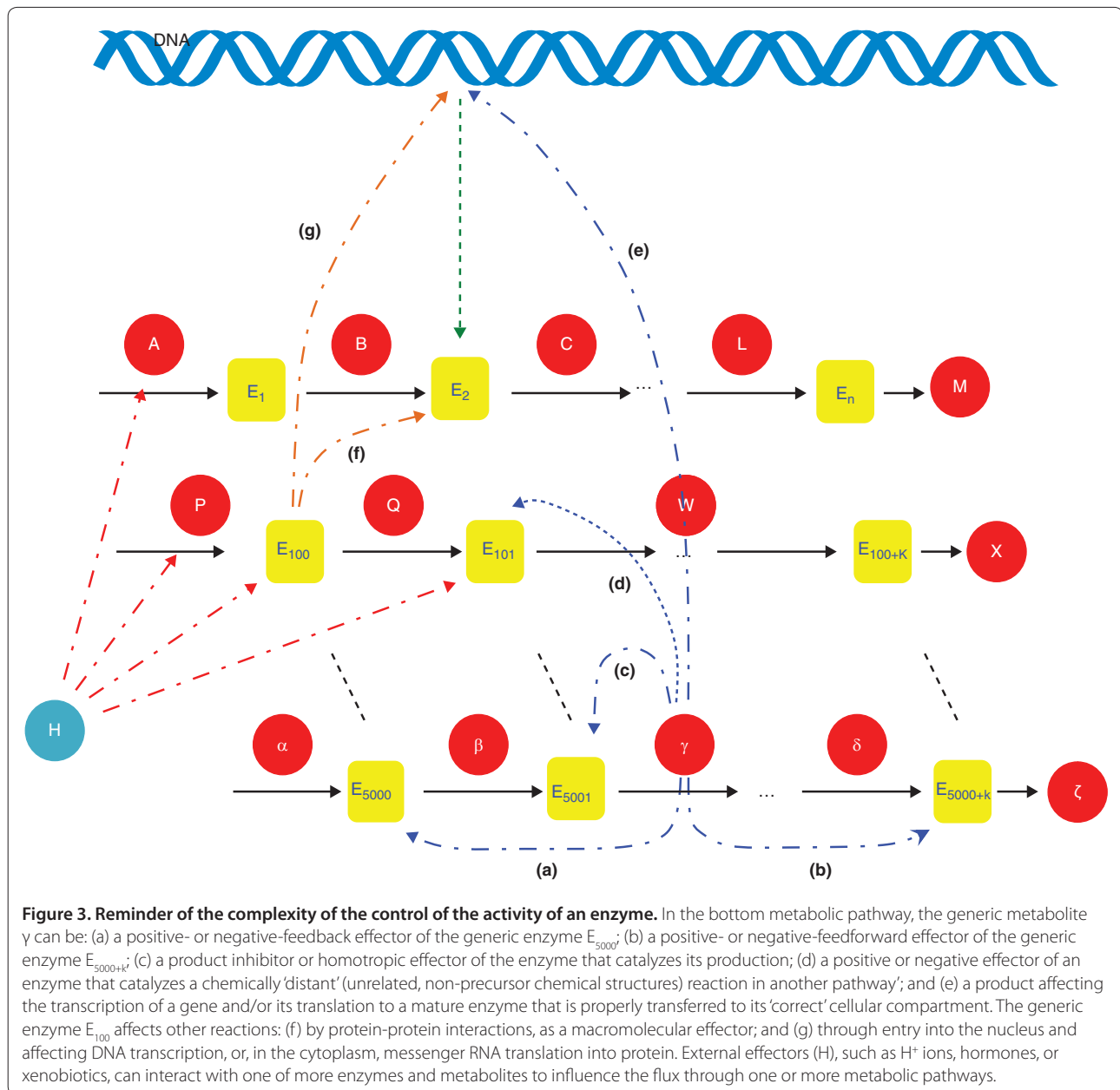


responses to nervous and endocrine ‘cues’; and (3) ‘What if only a small fraction of the cells were transformed? What is the minimum fraction that would lead to “rescuing” the metabolic state of the whole organ(s) and hence the individual?’

### Quantitative prediction of metabolic responses

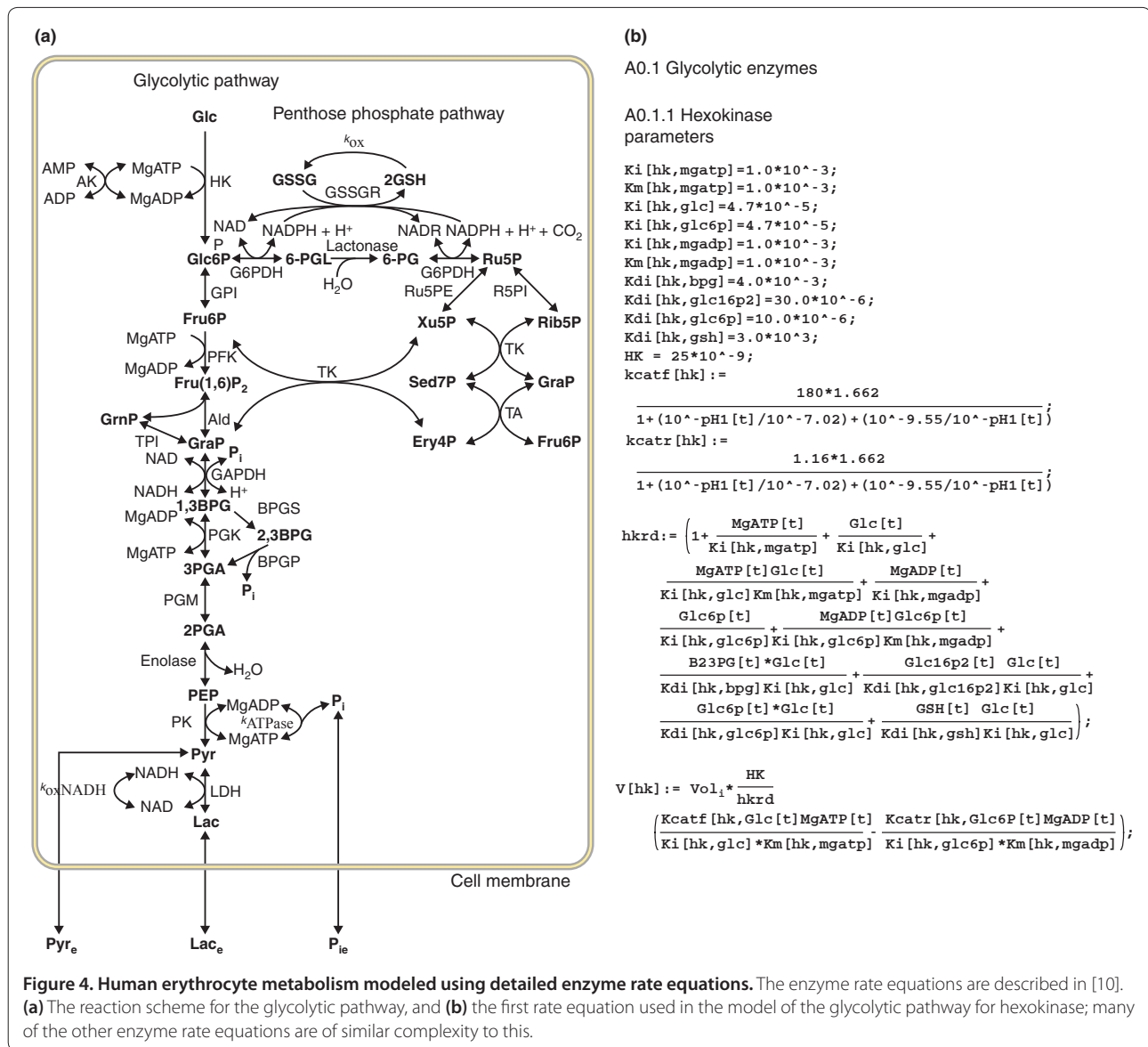
How do we begin to predict the metabolic responses to experimental genetic manipulations in something as chemically complex as a baby (or even a mouse), when we struggle to describe metabolism in quantitative terms for even the simplest of cells, notably erythrocytes (for example, [4-10])? To give an impression of the task at hand, consider glycolysis and the pentose phosphate pathway of the human erythrocyte (Figure 4a): there are approximately 25 enzymes involved (but there are as

many, again, doing other things, not included here, such as peptidases, phospholipases, catalase, carbonic anhydrase, and so on), and hexokinase, the first enzyme in the pathway, has the level of details shown in Figure 4b to account for its reaction rate as a function of the concentration of substrates, products and effectors, including  $H^+$ . In order to account for the exquisite pH dependence of the steady-state concentration of 2,3-bisphosphoglycerate, the pH dependence of all the key reactions (enzymes) needed to be incorporated into the expressions for the various equilibrium and kinetic constants. Only then was it possible to analyze the mathematical model to identify the fact that  $H^+$  ions exerted their effect on the concentration of 2,3-bisphosphoglycerate mostly via three different enzymes, two of which are far removed in the pathway. Such is the behavior of a system that in



effect is run by a committee! This type of analysis was only made possible by performing a type of meta-analysis on the model using the guiding principles of metabolic control analysis [11] and especially the important idea of co-response coefficients [12,13]. In other words, having done an experimental study of a metabolic system, a mathematical model consisting of rate equations is formulated; and the simulations are used to test hypotheses that relate to control of the reaction network. This abstraction is then used to inform further experiments on the real system, and so forth, in a series of iterative loops between numerical simulation and real experiment, thus refining understanding of the real system.

Metabolic processes in unicellular organisms such as bacteria and yeast have been studied using this approach, but they turn out to be even more complex than the human erythrocyte. This is because they have the full complement of metabolic machinery that is required to maintain an autonomous existence and to reproduce themselves; the human (mammalian) erythrocyte is an end-stage differentiated cell and thus, while relatively simpler, it is still complex. The human erythrocyte has been subjected to the most detailed biochemical analysis and computer modeling of all known cell types, and has been a fruitful guide to the future of metabolic simulations and quantitative analysis of metabolic



responses [7-9]. This analysis probably already includes most of the concepts that will be necessary to scale up to a model of the whole human metabolic network.

### Computer models of metabolism

It is intriguing that the first serious attempts to model metabolism in cells considered yeast, hepatocytes, and myocytes, and the models began with a high level of complexity. Consideration was given to the detailed mechanisms of the individual enzymes in many metabolic pathways, such as those shown in stylized form in Figure 1a, with control of enzymes by small molecules as is represented in Figure 3. Such work was exemplified by that of Britton Chance, Edwin Chance and Joseph Higgins, and later by that of David and Lillian Garfinkel and colleagues

[14]. As it was obvious 40 years ago, and is even more apparent today, it is difficult to obtain the coherent/consistent sets of data required to guide the development of quantitative models of metabolism in a particular tissue [7-9]. Future developments will need some, and more, of the blanket approaches to identify and quantify metabolites that have been used in metabol(n)omics, such as chromatographic methods linked to mass spectrometry and nuclear magnetic resonance spectroscopy [15,16]; also called 'hyphenated modalities'.

Those interested in optimizing batch cultures of microorganisms for the industrial production of substances such as antibiotics, or even simple ethanol, have adopted a more phenomenological approach to their models [17,18]; in other words, an attempt is made to represent

or describe a phenomenon without trying to infer a detailed underlying mechanism for each enzymic reaction. While some of these models of metabolism are very complicated, they do not (generally) involve the fine details of pre-steady-state or even steady-state rate equations for the respective enzymes. The set of simultaneous linear and non-linear differential equations that constitute deterministic models can be investigated using a form of sensitivity analysis (developed in the 1960s by chemical engineers [19], and now a part of metabolic control analysis [11]) to help identify flux-controlling steps (enzymes) that then become the target for genetic manipulations of the organism [5].

The main proponent of large-scale modeling of metabolism is Professor Bernhard Palsson and his team at the University of California, San Diego, California, USA. Their work to date has largely been phenomenological and can be classified as 'biochemical engineering'; it is of a kind that also attracted attention to the late Professor James Bailey, who nevertheless recognized the need to consider genomics in formulating the next generation of metabolic models [20]. The emphasis is on process output and the amount of detail used, as in pragmatic engineering, is just sufficient for describing the bio-processing task in hand. The models are fundamentally different from those that biochemists have constructed of human erythrocyte metabolism [7-10]. However, in the process of setting up their massive databases, Palsson and colleagues have established a means of storing information relating to vast arrays of individual enzymes. This 'library' system could, in principle, contain, and be used to curate, all the data compiled in any other highly enzyme-mechanism-based model; indeed, they have already subsumed some of the more mechanistic equations from other models, such as in [6].

Thus, the large-scale and very ambitious projects in metabolic modeling have identified the need to curate data from disparate sources and make it available to one model. Palsson's team recently listed 45 bacteria, 2 archaea, and 11 eukaryotes, including *Homo sapiens*, among those with detailed models of metabolism in their database [21]. To obtain some idea of the complexity involved, consider *Bacillus subtilis*: there are 4,114 genes that express 1,103 enzymes/proteins involved in 1,437 reactions with 1,138 metabolites [21,22]. Keeping track of the metabolites and the reaction kinetics with experimental data to justify particular choices of parameter values demands elegant file-handling programs and powerful computers.

The process of setting up the differential rate equations that are solved to predict time courses of metabolism under various conditions rests on a central idea that is well described in the book by Heinrich and Schuster [11], namely the stoichiometry matrix, and it has been

implemented in other well-known programs (for example, [23], and also in [10]). This is a mathematical construct that has a list of reaction names (enzyme names) in the metabolic system across the top of the columns of the matrix. The matrix is often gigantic, having as many columns as there are enzymes, and the metabolite names (reactants), which can number in the thousands, down the rows. Automatic writing of the differential equations that describe the rates of the biochemical reactions is done by the computer program (for example, [21]; this has also been done, on a smaller scale, in *Mathematica* [10]); the process involves accessing a separate list (the velocity vector) of rate equations that contains the kinetic descriptions of each reaction, either at the level of steady-state kinetics - for example, the Michaelis-Menten equation - or represented as simple first and second order rate equations where the enzyme concentration is implicit in the value of a rate constant. Thus, there are as many differential rate equations as there are metabolites. In other words, the model can engulf all previous estimates of metabolite concentrations and enzyme kinetic data relevant to the metabolic pathway under consideration.

The massive library of metabolic information, organized around the velocity and substrate vectors and the stoichiometry matrix, can readily be expanded to incorporate control networks, such as hormone effects (for example, [17]). However, a major question that emerges from combining all these data is how do conflicts between disparate data sets, from different investigations/investigators with different techniques, get resolved? The problem has not been systematically resolved and has been left to individuals to do the filtering of the data (for example, [24]).

### A coarser grained view

The major effort in quantitative holistic human modeling is the Human Physiome Project [25]. The Human Physiome Project runs under the aegis of the International Union of Physiological Societies, and the Institute of Electronic and Electrical Engineers' Engineering in Medicine and Biology Society, and it was made the main focus of the International Union of Physiological Societies for the decade commencing in 1993, and it continues today [26]; but the temporal and structural scales have not been those of metabolism - they are more those of tissue/anatomical structure. The Human Physiome Project is divided into 12 major systems, with the heart and cardiovascular system appearing to attract most attention (for example, [27,28]). The blood in this system (hematopoietic tissue plus circulating erythrocytes; also called the erythron) constitutes approximately 6 kg of the average adult mass (8.6%), with the approximately 2 kg of erythrocytes visiting all tissues, being a

major antioxidant via plasma membrane oxidoreductases and intracellular glutathione; and blood is also the main vehicle for the distribution (and degradation) of hormones. A model of the blood should be a key aspect of the quantitative human physiome; it will tie all the 12 systems together, with hormone signaling, nutrient and O<sub>2</sub> delivery, and metabolite and CO<sub>2</sub> disposal, as relevant to all tissues. On the other hand, there appear to be few signs that models of human erythrocyte metabolism are about to be included in the Human Physiome Project; so inclusion of the much more complex metabolic models of Pálsson *et al.* (for example, [21,22]) into the Human Physiome Project appears remote at this juncture.

### Metabonomics and its challenges

A recent application of metabonomics has been in experimental pancreatitis in animals in which major changes in blood chemistry are seen in response to arginine overloading. The interpretation of the metabolic profiles is based on known biochemical pathways, and yet the interpretation is still only qualitative. Nevertheless, the work appears to lend itself to quantitative metabolic modeling, which could make predictions more robust before it is applied to humans [29]. In spite of the huge amount of biochemical information available in such studies, much more information is required to make an enzyme-mechanistic model of the system of the kind developed for the human erythrocyte [7-10].

### Complicating issues

Thus far we have considered straightforward comparisons between standard enzyme kinetics and the prediction of metabolic responses. However, it is well known that some reactions inside cells do not follow the kinetics predicted from studies *in vitro*. One of the hopes for magnetic resonance spectroscopy is to study the kinetics of reactions as they occur *in situ* in cells or tissues. A complication that arises *in situ* is metabolite/substrate channeling, and yet the only model to date that has been based on real experimental data is that of arginine channeling in the urea cycle of isolated rat hepatocytes [30]. How much more complicated would be the kinetic characterization of metabolite channeling in the human liver *in vivo*?

One way to begin to look more closely at the flux of carbon atoms in metabolites through intersecting metabolic modules is to use <sup>13</sup>C nuclear magnetic resonance isotopomer analysis (for example, [31]). The ensuing increase in computational complexity brought about by the requirement to keep track of all combinations of <sup>13</sup>C labels in isotopomers has seen this area of computer modeling move very slowly. Nevertheless, the recent example of *B. subtilis* metabolism is an important advance [22]. And there is another subtlety: not all sites

in an end product of a metabolite may ever be labeled because of the particular subset of combinatorial shuffling of carbon atoms at different positions in a metabolite in a cell type. This realization both complicates possible experimental interpretations and could also serve as a type of diagnostic test, identifying which of a set of possible reactions are in operation in a tissue or cell type in a given time interval [32].

### Conclusions

It appears that the methods of metabol(n)omics that generate massive data sets on metabolite concentrations might tempt speculation that a detailed quantitative predictive model of the whole human metabolic network is imminent. On the other side of the 'conceptual divide', modelers of complicated metabolism, who have solved the problem of data curation, and fast and accurate numerical integration of differential rate equations, imply that the 'all that is needed are some data'; their methods are ready, waiting, and up to the task. Unfortunately, even modeling the metabolism of the simplest mammalian cell, the erythrocyte, has and still does require painstaking experimental analysis by a range of techniques; the latest addition in this area (on glutathione synthesis) was 6 years in the making [24]!

In conclusion, it would be demoralizing to base our predictions of a date when the whole human metabolic network would be complete on present technology. What is needed is the counterpart of the sort of breakthrough in technology that saw the Human Genome Project reach fruition 'from left field' via shotgun DNA sequencing, which is utterly reliant on massive computer power. It appears that, in the present case, we have the computing power and methods, but what we lack are the techniques of metabolite analysis, and various means of rapidly recording protein-protein and ligand-protein interactions. Furthermore, the genome-centric view of metabolism is identifying new modes of metabolic regulation, such as the indirect effects of interfering RNAs, and these will need to be incorporated in models of metabolism and its control. Therefore, there is much to be done before computer models of metabolism form part of the suite of methods used in clinical management.

### Competing interests

The author declares that he has no competing interests.

### Author's information

PWK is McCaughey Professor of Biochemistry at the University of Sydney. The main biological focus of his work is the human erythrocyte; his technological focus is NMR spectroscopy; and data from biochemical and physical systems are analyzed and modeled using numerical and statistical approaches, with heavy reliance on *Mathematica*.

### Acknowledgements

Thanks to Drs Tim Larkin and Anthony Maher, and Professor Lindy Rae for critical comments on the manuscript. The work was funded by a Discovery Project Grant from the Australian Research Council.

Published: 28 July 2010

## References

- Gerrard JA, Sparrow AD, Wells JA: **Metabolic databases - what next?** *Trends Biochem Sci* 2001, **26**:137-140.
- Lindon JC, Nicholson JK: **Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics.** *Ann Rev Anal Chem* 2008, **1**:45-69.
- Cunningham SC, Kok CY, Dane AP, Carpenter KH, Kuchel PW, Alexander IE: **Production and rescue of a severe phenotype of ornithine transcarbamylase deficiency in the Spf-ash mouse model using adeno-associated viral vectors and RNAi technology.** *J Gene Med* 2009, **11**:843.
- Kirk K, Kuchel PW: **Red-cell volume changes monitored using P-31 NMR - a method and model.** *Stud Biophys* 1986, **116**:139-140.
- Raftos JE, Chapman BE, Kuchel PW, Lovric VA, Stewart IM: **Intraerythrocyte and extraerythrocyte pH at 37°C and during long-term storage at 4°C - P-31 NMR measurements and an electrochemical model of the system.** *Haematologia* 1986, **19**:251-268.
- Thorburn DR, Kuchel PW: **Regulation of the human-erythrocyte hexose-monophosphate shunt under conditions of oxidative stress - a study using NMR-spectroscopy, a kinetic isotope effect, a reconstituted system and computer-simulation.** *Eur J Biochem* 1985, **150**:371-386.
- Mulquiney PJ, Bubb WA, Kuchel PW: **Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations in vivo kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using C-13 and P-31 NMR.** *Biochem J* 1999, **342**:567-580.
- Mulquiney PJ, Kuchel PW: **Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: equations and parameter refinement.** *Biochem J* 1999, **342**:581-596.
- Mulquiney PJ, Kuchel PW: **Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis.** *Biochem J* 1999, **342**:597-604.
- Mulquiney PJ, Kuchel PW: *Modelling Metabolism with Mathematica.* Boca Raton, FL: CRC Press; 2003.
- Heinrich R, Schuster S: *The Regulation of Cellular Systems.* New York: Chapman and Hall; 1996.
- Cornish-Bowden A, Hofmeyr JHS: **Determination of control coefficients in intact metabolic systems.** *Biochem J* 1994, **298**:367-375.
- Hofmeyr JHS, Cornish-Bowden A: **Co-response analysis - a new strategy for experimental metabolic control analysis.** In *What Is Controlling Life?: 50 Years after Erwin Schrödinger's What Is Life?* Edited by Gnaiger E, Gellerich FN, Wyss M. Innsbruck: Innsbruck University Press; 1994:109.
- Garfinkel D, Garfinkel L, Pring M, Green SB, Chance B: **Computer applications to biochemical kinetics.** *Annu Rev Biochem* 1970, **39**:473-498.
- Wishart DS: **Quantitative metabolomics using NMR.** *Trends Anal Chem* 2008, **27**:228-237.
- Kirschenlohr HL, Griffin JL, Clarke SC, Rhydwen R, Grace AA, Schofield PM, Brindle KM, Metcalfe JC: **Proton NMR analysis of plasma is a weak predictor of coronary artery disease.** *Nat Med* 2006, **12**:705-710.
- Papin JA, Palsson BO: **The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis.** *Biophys J* 2004, **87**:37-46.
- Vaidyanathan S, Harrigan G, Goodacre R: *Metabolome Analyses: Strategies for Systems Biology.* New York: Springer; 2005.
- Rosenbrock HH, Storey C: *Computational Techniques for Chemical Engineers.* Oxford: Pergamon Press; 1966.
- Bailey JE: **Complex biology with no parameters.** *Nat Biotechnol* 2001, **19**:503-504.
- Feist AM, Herrgård MJ, Thiele I, Reed JL, Palsson BØ: **Reconstruction of biochemical networks in microorganisms.** *Nat Rev Microbiol* 2009, **7**:129-143.
- Dauner M, Bailey JE, Sauer U: **Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*.** *Biotechnol Bioeng* 2001, **76**:144-156.
- Mendes P: **Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3.** *Trends Biochem Sci* 1997, **22**:361-363.
- Raftos JE, Whillier S, Kuchel PW: **Glutathione synthesis and turnover in the human erythrocyte: alignment of a model based on detailed enzyme kinetics with experimental data.** *J Biol Chem* 2010, in press.
- Hunter PJ: **The IUPS Physiome project.** *J Physiol Sci* 2009, **59**:46.
- Hunter PJ: **Modeling human physiology: The IUPS/EMBS physiome project.** *Proc IEEE* 2006, **94**:678-691.
- Smith NP, Crampin EJ, Niederer SA, Bassingthwaite JB, Beard DA: **Computational biology of cardiac myocytes: proposed standards for the physiome.** *J Exp Biol* 2007, **210**:1576-1583.
- Smith NP, Hunter PJ, Paterson DJ: **The cardiac physiome: at the heart of coupling models to measurement.** *Exp Physiol* 2009, **94**:469-471.
- Bohus E, Coen M, Keun HC, Ebbels TM, Beckonert O, Lindon JC, Holmes E, Noszál B, Nicholson JK: **Temporal metabonomic modeling of L-arginine-induced exocrine pancreatitis.** *J Proteome Res* 2008, **7**:4435-4445.
- Maher AD, Kuchel PW, Ortega F, de Atauri P, Centelles J, Cascante M: **Mathematical modelling of the urea cycle - a numerical investigation into substrate channelling.** *Eur J Biochem* 2003, **270**:3953-3961.
- Berthon HA, Bubb WA, Kuchel PW: **<sup>13</sup>C NMR isotopomer and computer-simulation studies of the nonoxidative pentose-phosphate pathway of human erythrocytes.** *Biochem J* 1993, **296**:379-387.
- Kuchel PW, Philp DJ: **Isotopomer subspaces as indicators of metabolic-pathway structure.** *J Theor Biol* 2008, **252**:391-401.

doi:10.1186/gm167

Cite this article as: Kuchel PW: Models of the human metabolic network: aiming to reconcile metabolomics and genomics. *Genome Medicine* 2010, **2**:46.