Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 252 (2008) 391–401
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Isotopomer subspaces as indicators of metabolic-pathway structure

Philip W. Kuchel[a,b,*], David J. Philp[c]

[a]School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia
[b]Centre for Mathematical Biology, University of Sydney, NSW 2006, Australia
[c]National Centre for Epidemiology and Population Health, Australian National University, Canberra, ACT 0200, Australia

Received 18 April 2007; received in revised form 11 May 2007; accepted 15 May 2007
Available online 5 July 2007

This paper is dedicated to the memory of Professor Reinhart Heinrich of Berlin

Abstract

The relative abundances and rates of formation of particular isotopic isomers (isotopomers) of metabolic intermediates from ^{13}C-labelled substrates in living cells provide information on the routes taken by the initial ^{13}C-atoms. When a primary substrate such as [U,^{13}C] D-glucose is added to human erythrocytes, the pattern of labels in terminal metabolites is determined by a set of carbon-group exchange reactions in both glycolysis and the pentose phosphate pathway (PPP). Of a given terminal metabolite, not all possible isotopomers will be produced from each possible primary substrate isotopomer.

There are only 8 different ^{13}C-isotopomers of lactate but *not all* of these are produced when one of the 64 possible ^{13}C-isotopomers of glucose is used as the input substrate; thus a subset of all 63 glucose isotopomers $\times$ 8 lactate isotopomers + 1 unlabelled glucose $\times$ 1 unlabelled lactate = 505 *pattern associations*, would be produced if a complete experimental analysis were performed with all the glucose variants. The pattern of labelling in this *isotopomer subspace* reflects the nature of the re-ordering reactions that 'direct' the metabolism. Predicting the combinatorial rearrangements for particular sets of reactions and comparing these with real data should enable conclusions to be drawn about which enzymes are involved in the real metabolic system. An example of the glycolysis–PPP system is discussed in the context of a debate that occurred around the F- and L-type PPPs and which one actually operates in the human RBC. As part of this discussion we introduce the term '*combinatorial deficit*' of all possible isotopomers and we show that this deficit is less for the F- than the L-type pathway.

Keywords: Isotopomers; ^{13}C NMR; Combinatorics; *Mathematica*; Pentose phosphate pathway; Erythrocyte metabolism

1. Introduction

1.1. General

We introduce and discuss two new concepts relating to the use of isotopic isomers (isotopomers) to delineate pathways of carbon flux in cellular metabolism. The concepts are '*subspaces of isotopomers*' and '*combinatorial deficits*' of sets of isotopomers, and they are presented in the context of experiments and computer-based simulations of particular pathways of metabolism. These concepts are useful when attempting to settle arguments concerning the likely route(s) taken by carbon atoms that are introduced in a substrate such as glucose and pass through various possible enzymes to a product such as lactate.

In an analogy with astronomy, the carbon atoms (stars) in each different metabolite are a set of potential labelling sites (equivalent to a constellation) within a set of metabolites (the galaxy). The analysis presented here determines which of these atoms (stars) are allowed, by the combinatorial rules, to be visited in a trip through the metabolic pathway (galaxy) from a given starting position on an input molecule (constellation).

Abbreviations: MCA; metabolic control analysis; NMR; nuclear magnetic resonance; PPP; pentose phosphate pathway; RBC; red blood cell

*Corresponding author. School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia.
Tel.: +61 2 9351 3709; fax: +61 2 9351 4726.
E-mail address: p.kuchel@mmb.usyd.edu.au (P.W. Kuchel).

0022-5193/$ - see front matter
doi:10.1016/j.jtbi.2007.05.039

1.2. NMR

Isotopomer analysis relies on the ability of ^{13}C NMR spectroscopy to identify and quantify the isotopomers of a given metabolite in a complex mixture based on the changes that occur in the spectra when there are one or more ^{13}C-labelled atoms in the molecule: the changes depend both on the location of the labelled atom, and its proximity to other labelled atoms (London, 1988). The method was first applied in studies of bacterial metabolism (Walker et al., 1975, 1976, 1982; Walker and London, 1987) and was then used on ^{13}C-acetate and ^{13}C-pyruvate dissimilation in Krebs cycle intermediates in the perfused rat heart (Chance et al., 1983); this work was extended by Malloy et al. (1988). Amongst other systems that have been studied with NMR isotopomer analysis are isolated rat renal tubules (Cowin et al., 1994) and human red blood cells (RBCs; Berthon et al., 1993).

1.3. Computer models of metabolism

Modelling carbon flux via metabolic pathways by using numerical integration to solve the nonlinear differential equations that describe the kinetics of enzyme and membrane transport reactions is a mature science (e.g., Kauffman et al., 2002; Mulquiney et al., 1999; Mulquiney and Kuchel, 1999a, b, 2003; Maher et al., 2003). Software packages like *Mathematica* (Wolfram, 2003; Barnett et al., 2004) have greatly facilitated the process of model formulation because of their flexibility in manipulating and representing symbolic expressions. However, isotopomer analysis remains an area of metabolic simulation that has not yielded readily to computer modelling, because this requires accurate entry of a differential equation for each possible combination of isotopomers of substrates and then solving a large array of such complicated differential equations. For example, without accounting for isotopomers, our RBC system contains ~100 differential equations, each with at most a few dozen terms. Explicitly including isotopomers requires a few thousand equations, with up to a few hundred terms each. The first comprehensive simulation of isotopomer production in the PPP of RBCs was by Berthon et al. (1993). And more recently Selivanov et al. (2004, 2005) developed an algorithm that yields close fits of a model of isotopomer flux to experimental data from cancer cells in culture.

1.4. Pentose phosphate pathways

In the 1980s there was considerable debate and controversy regarding the possible existence of a pentose phosphate pathway (PPP) that was different from the standard 'text book model' (Horecker, 2002; Williams et al., 1987). The latter was known to operate in adipocytes (fat cells) and was called the F-type PPP to distinguish it from the newly proposed one that putatively operates in the liver, called the L-type PPP. Definitive NMR-based experiments that refuted the operation of the L-type pathway in human RBCs were reported in 1989 (McIntyre et al., 1989); this conclusion was made in spite of the inference by Williams et al. (1987) that the L-type system operates in them.

What caught the attention of many biochemists was the fact that the new L-type proposal requires for its operation: (1) two 'new' enzymes that are not part of the F-type pathway, viz., a phosphotransferase that exchanges a phosphate group from octulose 1,8-bisphosphate to sedoheptulose 7-phosphate (and vice versa), and arabinose-5-phosphate epimerase; (2) the presence of four 'new' metabolites, octulose 1,8-bisphosphate, octulose 8-phosphate, sedoheptulase 1,7-bisphosphate, and arabinose 5-phosphate; and (3) the non-participation of the ubiquitous cytoplasmic enzyme, transaldolase; this enzyme is notoriously heat labile, a property that was used to advantage in pathway-reconstitution experiments (McIntyre et al. 1989). While the idea that the L-type PPP operates to any measurable extent in RBCs has been refuted, the violation of the deeper philosophical point of 'maximal parsimony' in formulating a hypothesis, as it relates to metabolism in the liver, especially since transaldolase is present there in abundance, seemed to have occurred.

What has been lacking is some form of test that could be applied to experimental results on metabolite flux, or numerical models of the systems, to resolve this controversy. Meléndez-Hevia and Isidoro (1985) analysed the rearrangement of the carbon backbones of sugar phosphates in the two proposals for the PPP and for the Calvin cycle of photosynthesis. They introduced a *hypothesis of simplicity* that they applied to the PPPs and concluded that the least number of steps with the least number of carbon atoms per metabolite exist in the two PPP proposals; in other words they are 'optimal' for the specified sets of enzymes. The question remains, though, whether the choice of the set of enzymes, especially for the L-type PPP, is correct. More recent analysis posits that 'flexibility' in substrate usage by an enzyme is an adaptive advantage and the basis of evolution of metabolic pathways (Montero et al., 1997); as such this expansion of the substrate specificity of enzymes could lead to the appearance of more different metabolites in some pathways and evolution of alternative flux streams.

1.5. Isotopomer simulations

General strategies for the simulation of isotopomer distributions and intracellular fluxes have been developed, and in principle they could be used to explore the kinetic behaviour of variants of the PPPs (Zupke and Stephanopoulos, 1994; Wiechert, 2001; Wiechert et al. 2001). Experimental data could be interpreted by using numerical simulations of the various candidate reactions schemes, bearing in mind that such schemes need to take account of tens to hundreds of possible isotopomers per metabolite. The analysis requires state-of-the-art computer methods

and hardware for the simulations but the simulations are still slow (minutes to hours), making reiteration and model refinement a slow process.

An alternative approach is to ask a simpler set of questions based on combinatorics and not kinetics, viz., what are the contents of the set of all possible isotopomers of a product, for a given isotopomer input? Our program written for this purpose exploits the versatile syntax of *Mathematica* to express a large number of logical statements about isotopomer reactions, with a few purpose-built functions (see Appendix A for more details).

Fig. 1 shows an open-chain representation of D-glucose with its numbering system as used here. Our graphical representation of all possible isotopomers of glucose is shown in Fig. 2; and an example is as follows. Glucose labelled in the C-1 position is denoted by $Glc_{0,0,0,0,0,1}$ where the '1' specifies a ^{13}C-label, and all the isotopomers are graphically arrayed via ListDensityPlot. All other metabolites are represented in a similar way. To explain this further, consider the following example: The glucose isotopomer whose representation is located at row-3 and column-4 of the graphical array in Fig. 2 has the pattern of light and dark bands that indicate the labelling {0, 1, 0, 0, 1, 1}, or more conventionally it is $^{12}C_6$-$^{13}C_5$- $^{12}C_4$- $^{12}C_3$-$^{13}C_2$- $^{13}C_1$ thus specifying ^{13}C-label is in positions 1, 2 and 5 (recalling that we read the pattern from *right to left*). On the other hand, lactate at 'coordinate' (3, 4) in the array in Fig. 3A has all 8 possible isotopomers; while in Fig. 3C the lactate is labelled as {1,1,0}, {1,0,0}, {0,1,0} and {0,0,0} showing only a subset of all 8 possible isotopomers.

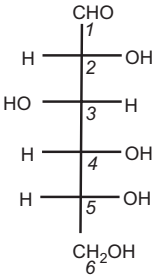

Fig. 1. Glucose molecule in an open chain representation showing the standard numbering system that was used in the present analyses.

Now we in a position to address the question whether all possible isotopomers of every metabolite are produced by a given set of reactions, using only knowledge of the way in which the enzymes 'shuffle' the order of carbon atoms in their reactants. This analysis has been done for metabolism in the human RBC with several variants of the two types of PPP.

1.6. RBC metabolism

The metabolic networks in the human RBC are well-characterized (e.g., Fig. 1 in Kuchel, 2004) and consist of: glycolysis with the special feature of the Rapoport-Luebering shunt that synthesizes 2,3-bisphosphoglycerate; the PPP; glutathione synthesis; the glyoxalase pathway; the purine nucleoside salvage pathway; multiple seemingly disconnected enzymic reactions; and transport processes, many of which are functionally linked to ATP generation by glycolysis. Most enzymes transform substrates to products without modifying the order of the carbon atoms in their carbon backbone. In some notable cases, a part of

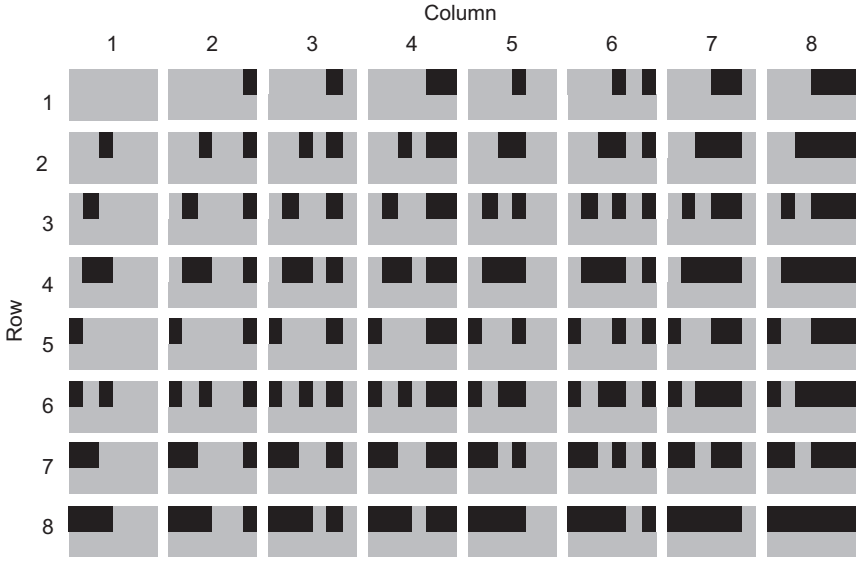

Fig. 2. Glucose isotopomers represented as a series of six small boxes inside a larger one; black denotes a ^{13}C atom in the respective position in the carbon chain. The lower half of each small box represents unlabelled glucose; the box in position (row 1 and column 1) (i.e., (1,1)) represents all unlabelled glucose. Because *Mathematica* by default orders lists of numbers with the least significant digit on the right, the carbon atoms are labelled from right to left, i.e., with C-1 on the right. The large 8 × 8 array (8 rows and 8 columns) has cells that each contain a representation of a unique glucose isotopomer in the upper half and a row of blank bars denoting unlabelled glucose in the lower half. This combination of the unlabelled isotopomer of glucose and fully *unlabelled* glucose represents the input to the pathway. Other combinations with more than two isotopomers are possible.

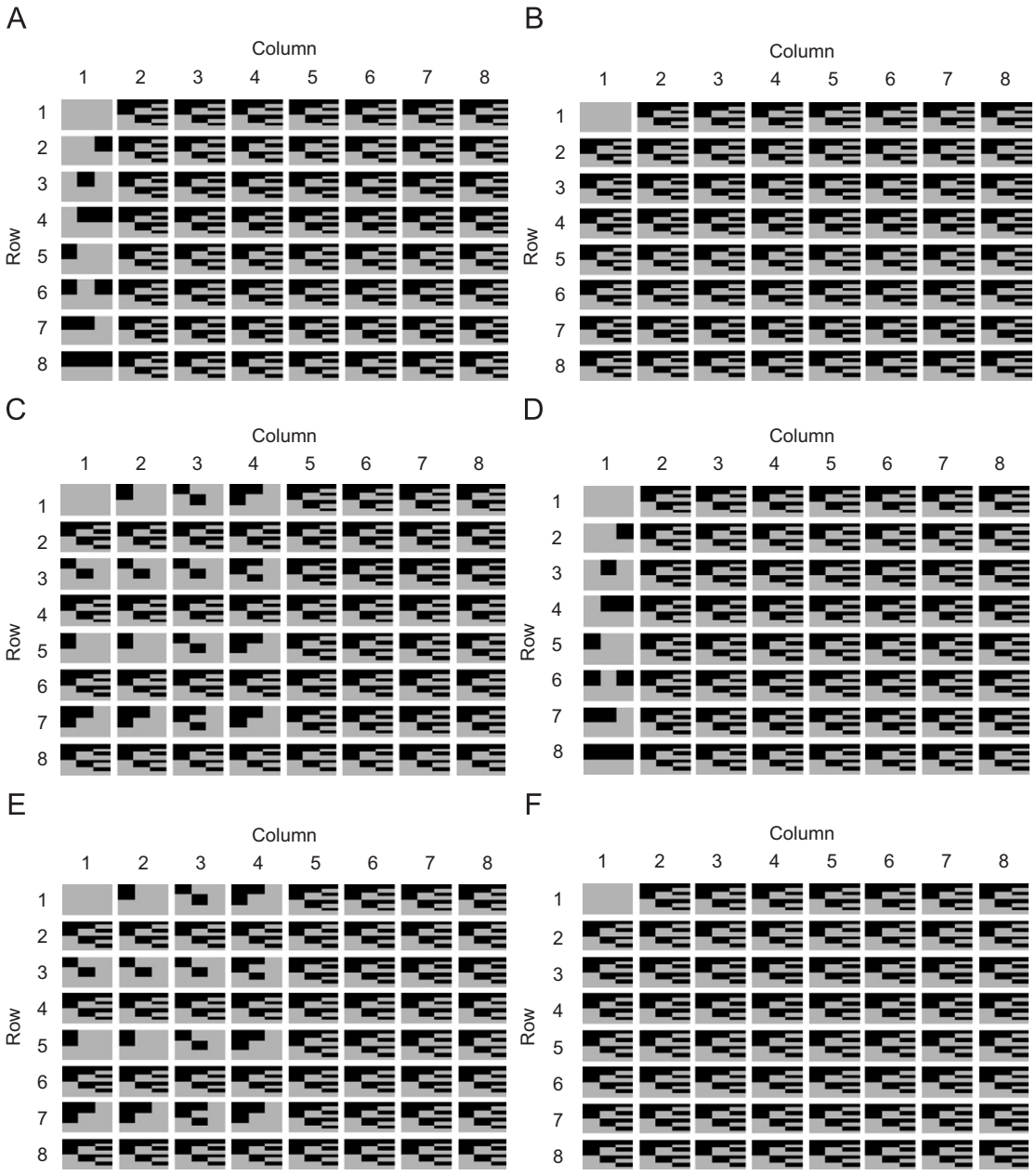

Fig. 3. The set of all possible isotopomers of lactate derived from a given input of labelled glucose: (A) from the 'classical' F-type PPP; (B) from the F-type PPP but including fructose-1,6-bisphosphatase; (C) the L-type PPP which includes no action by transaldolase; (D) the L-type PPP but including transaldolase; (E) L-type PPP but including fructose-1,6-bisphosphatase; and (F) L-type PPP but including both fructose-1,6-bisphosphatase and transaldolase. Each small rectangle in the three 8×8 arrays denotes a set of isotopomers of lactate. The three *columns* in each small rectangle correspond to one of the 3-carbon atoms, labelling 1, 2, and 3 from right to left. Each small rectangle has a maximum of 8 *rows* (each carbon can potentially be labelled or unlabelled, giving $2^3 = 8$ possible isotopomers). E.g., consider array (C), row 1 and column 4; there are only three rows, the first has label in positions 2 and 3, the second row has label in position 3, and the third row has no label at all. This particular set of isotopomers of lactate is derived from the mixture of unlabelled and labelled glucose that has the same position in *its* array; and according to row 1, column 4 of the glucose input array, the glucose is labelled in carbons 1 and 2. In other words, each small rectangle corresponds to a set of isotopomers that are derived from glucose with the same row and column number in its array, Fig. 2.

the carbon chain is moved from one substrate to a specific place in another. For example, in the oxidative part of the PPP, carbon-1 is removed as CO_2 from the 6-carbon precursor 6-phosphogluconate, then in the non-oxidative part transketolase removes a 2-carbon moiety (carbons 1 and 2) from a ketose phosphate and condenses it onto carbon-1 of an aldose phosphate; and transaldolase does the same but with a 3-carbon moiety. Thus there are well-defined 'rules' that describe carbon-group exchange reactions in any metabolic pathway, and symbolic manipulation of isotopomers is carried out in the computer program via these rules.

2. Results of combinatorial analysis

Fig. 3A shows the 8 × 8 array of rectangles that represent the isotopomers of lactate that are formed from the glucose isotopomers represented in Fig. 2, obtained with our *Mathematica* program (Appendix B). The metabolic system was that of human RBC metabolism as it is currently understood to operate (e.g., Mulquiney and Kuchel, 2003). The PPP is the F-type. Note that the glucose isotopomers in the first column of Fig. 2, viz., those labelled in the lower three C-atoms do not give all possible corresponding lactate isotopomers. In fact, out of a possible maximum of $8 \times 63+1 = 505$ labelled 'cells' the total number is less by $6 \times 7 = 42$; this is defined as the *combinatorial deficit* for this particular metabolic scheme. On the other hand, Fig. 3B was the output of the analysis obtained by including in the metabolic scheme fructose-1,6-bisphosphatase which is known to have very low activity in the human RBC. However, it is a key gluconeogenic enzyme in the liver. The combinatorial deficit in this situation is 0.

Fig. 3C shows the output of lactate isotopomers when the F-type PPP is replaced by the L-type scheme in human RBCs. Recall that this pathway does not involve the 3-carbon transfer activity of transaldolase. It is clear that many fewer isotopomers are produced; and on counting the deficiency of black cells it is seen that the combinatorial deficit is 76. In contrast, Fig. 3D shows that if transaldolase is *included* with the L-type pathway, the output of all possible lactate isotopomers is the same as for the F-type pathway (Fig. 3A); its combinatorial deficit is 42.

Fig. 3E shows the result of not including transaldolase but adding fructose-1,6-bisphosphatase to the system that includes the L-type PPP. The pattern is exactly the same as for the L-type PPP alone. In other words the addition of fructose-1,6-bisphosphatase to this scheme adds nothing to the combinatorial scrambling capacity. It underscores the major importance of transaldolase in the carbon-scrambling process in the PPP.

Fig. 3F shows the array of all possible lactate isotopomers when the L-type PPP operates in human RBCs with addition of both transaldolase *and* fructose-1,6-bisphosphatase. The output includes all possible lactate isotopomers for every glucose isotopomer that is a substrate; there is no combinatorial deficit, as in the system giving Fig. 3B (F-type PPP plus fructose-1,6-bisphosphatase).

3. Discussion of results

The F-type PPP has a carbon-scrambling action that acts differently on the two halves of the glucose backbone. A carbon at position 1, 2 or 3 of glucose may end up at any of the sites in the final lactate product. However, a carbon at position 4 can only end up at position 1, 5 at 2, and 6 at 3. This outcome is relaxed if the carbon atom fragment 1-2-3 can be swapped in position with carbon atoms 4-5-6 in fructose 1,6-bishosphate. This can occur via the hydrolase and the subsequent condensation reaction catalysed by aldolase. In contrast, the pattern of isotopomers in Fig. 3C shows that not all possible labelling sites in the set of all lactate isotopomers can be targeted by ^{13}C from glucose that enters the pathway as various isotopomers that have label in positions 4, 5 and 6. The subsets of isotopomers have peculiar patterns that are diagnostic of the underlying allowed group-exchange reactions in the metabolic pathway. Adding transaldolase to the L-type PPP introduces the 3-carbon transfer reaction and hence makes the combinatorial scrambling the same as for the F-type PPP (Fig. 3A). On the other hand, merely adding the capability to swap carbon atoms 1-2-3 and 4-5-6 in generating fructose 6-phosphate from fructose 1,6-bisphosphate does not alter the pattern (Fig. 3E). This underscores the key role of transaldolase in bringing about a high degree of label-scrambling.

It is perhaps not surprising that including both transaldolase and fructose-1,6-bisphosphatase with the L-type PPP brings about all possible lactate isotopomers (Fig. 3F); all possible labelling sites are targeted from all possible glucose isotopomers.

4. General discussion

The patterns of isotopomer distributions of lactate evident in Figs. 3C and 3E constitute a form of order, or reduced entropy, in label distribution. There does not seem any reason why such a bias should exist in a natural, highly evolved system. Therefore a working hypothesis is that the proposed mechanism of this pathway, the L-type PPP, is incorrect not only in the human RBC but in other tissues as well.

Notwithstanding this negative standpoint, the predictions of the model provide a basis whereby the veracity of the scheme could be tested further by the use of appropriately labelled isotopomers of glucose incubated with RBCs and measured by using ^{13}C NMR spectroscopy (e.g., Berthon et al., 1993).

Another advance in the analysis would be a full-scale kinetic simulation of the metabolic schemes. By doing this, not only the existence of the various isotopomers would be predicted but also the time courses of their evolution would emerge. At this stage there are major impediments for routine simulation, in terms of computer speed and memory capacity. However, the existence of small isotopomer subspaces in those schemes involving the L-type PPP suggest that this scheme is incomplete; and the calculation of the *combinatorial deficit* appears to be useful for applications to other systems.

A further level of comparison between likely candidates for predicted pathways is between the isotopomer patterns themselves, such as the fact that Figs. 3A and 3D are the same, as are Figs. 3C and 3E, and Figs. 3B and 3F. The common underlying features in the combinatorial scram-

bling are clear: in the first case it is the inclusion of transaldolase in the L-type PPP, and it is already present in the F-type PPP; in the second case the outcome shows that the end-to-end swapping of the 3-carbon units in the 6-carbon chain of fructose 1,6-bisphosphate, brought about by fructose-1,6-bisphosphatase and aldolase, achieves nothing extra in the combinatorial shuffling of label than is already achieved by transketolase and the 'new' enzymes of the L-type PPP. In the last case, both schemes achieve a zero combinatorial deficit; the mechanisms of the group shuffling reactions of the additional enzymes are the key to achieving this.

In this paper, that focuses on theoretical issues it seems relevant to explore whether the combinatorial shuffling invoked by the metabolic schemes could be treated as a formal algebraic system. Any enzyme or pathway of enzymes will bring about a reordering of the ^{13}C atoms in the input and intermediate metabolites; as such it is an operator. The output of products with various label distributions constitutes a (sub)space of isotopomers. Appendix A contains further discussion on this matter, but in brief, although it seemed promising that the system might be a formal mathematical *group*, this turns out not to be the case because of the inability to define a plausible inverse for each element of the input set of isotopomers. If a more abstract definition had been possible, then related theorems of group theory might have led to deeper insights than are immediately possible with direct inspection of the outcomes of computation.

It is worth re-iterating that our combinatorial analysis is not a substitute for more detailed modelling (e.g., Berthon et al., 1993; Malloy et al., 1988; Selivanov et al., 2004, 2005). Reference to the older literature reveals several notable examples of the combination of ^{13}C NMR and computer analysis (Chance et al., 1983). Specifically, though, the latter study dealt with a rat heart perfusion system and a computer model that described the Krebs cycle and related reactions; but it contained only ~200 differential flux equations. Thus the general approach used for data analysis was quite different from that used even in the later work noted above; and when the later methods (Mulquiney and Kuchel, 2003) are extended to isotopomers in RBC metabolism the number of differential equations rises to thousands. Hence the present combinatorial analysis was developed as a complement to the 'brute-force' methods; it gives an overview that is difficult to obtain with simulations that take minutes or hours to perform.

Our proposal to map the entire isotopomer space of a metabolic system has not yet been realized. There are potential practical problems associated with attempting to subject a metabolic system to the input of strictly pure well-defined mixtures of ^{13}C-enriched glucose. ^{13}C has a natural abundance of 1.1% thus 6.6% of glucose molecules would be expected to have one C atom as ^{13}C; while the probability of two or more labelled atoms in one molecule is of course very much less. A simple solution to this problem, at least in principle, is to synthesize glucose with 100% ^{12}C and to use this as the substrate for the cells during pre-incubations. This would circumvent the problem of having natural contaminating isotopomers; but in experimental practice it is likely that some impression of the importance of putative flux-pathways would emerge from determining *relative abundances* of the isotopomers. If only traces of product isotopomers were detected then a judgement would need to be made as to whether the amount constituted experimental 'noise' or not. In the final analysis, a full numerical simulation would still be needed to form a definitive conclusion about the effects of contaminating isotopomers that could be present at an abundance of around 1 in 100.

Finally, it is not always possible to unambiguously identify isotopomers from multiplets in crowded regions of a ^{13}C NMR spectrum. This would be another practical impediment to the proposal for a comprehensive combinatorial analysis of a metabolic system. However, modern multi-dimensional NMR methods, including the now routine ^{13}C-^{13}C COSY (e.g., Berthon et al. 1993), suggest that cautious optimism is warranted. An experimentalist could use our combinatorial analysis to decide which combination of isotopomers to use, to distinguish between two or more proposals, in the least number of experiments.

5. Conclusions

In conclusion, a *large isotopomer subspace*, and a corresponding *small combinatorial deficit* is a possible indicator of the level of plausibility of a proposed metabolic pathway. This concept has merit over previous attempts to comment in a mathematical/game theory way on the matter (e.g., Meléndez-Hevia and Isidoro, 1985; Montero et al., 1997) because specific hypotheses are suggested by analysing the simulated isotopomer patterns and comparing them with real experimental data. This can be done not only for a single product, as has been done here with lactate, but for a series of accessible metabolites. The analysis is ready for theoretical applications to pathways other than those described here, and the practical utility of our proposal awaits experimental testing.

Acknowledgements

The work was supported by a grant from the Australian Research Council to PWK and by a grant that supported DJP from the Australian Research Council to Dr. Katie Glass. Valuable discussions were held in Barcelona with Dr. Vitaly Selivanov and Professor Marta Cascante. Dr. Bill Bubb is thanked for ongoing collaboration in the area of biological NMR spectroscopy.

This paper is dedicated to the memory of Professor Reinhart Heinrich of Berlin.

Appendix A. Comments on a formal algebraic definition of an isotopomer-exchange system

Since the metabolic system depicted in Fig. 4 involves sets of inputs and outputs and an, albeit complicated, overall combinatorial operation, it seems relevant to explore if it constitutes a formal algebraic *group*. If it did, then there would be theorems that could be used to extend our understanding of its behaviour under possibly obscure conditions.

A simple group G is a set (finite or infinite) of elements S, together with a *binary operation* (the group operation) that satisfies the four fundamental properties of *closure*, *associativity*, possessing an *identity* element, and having for each element an *inverse*.

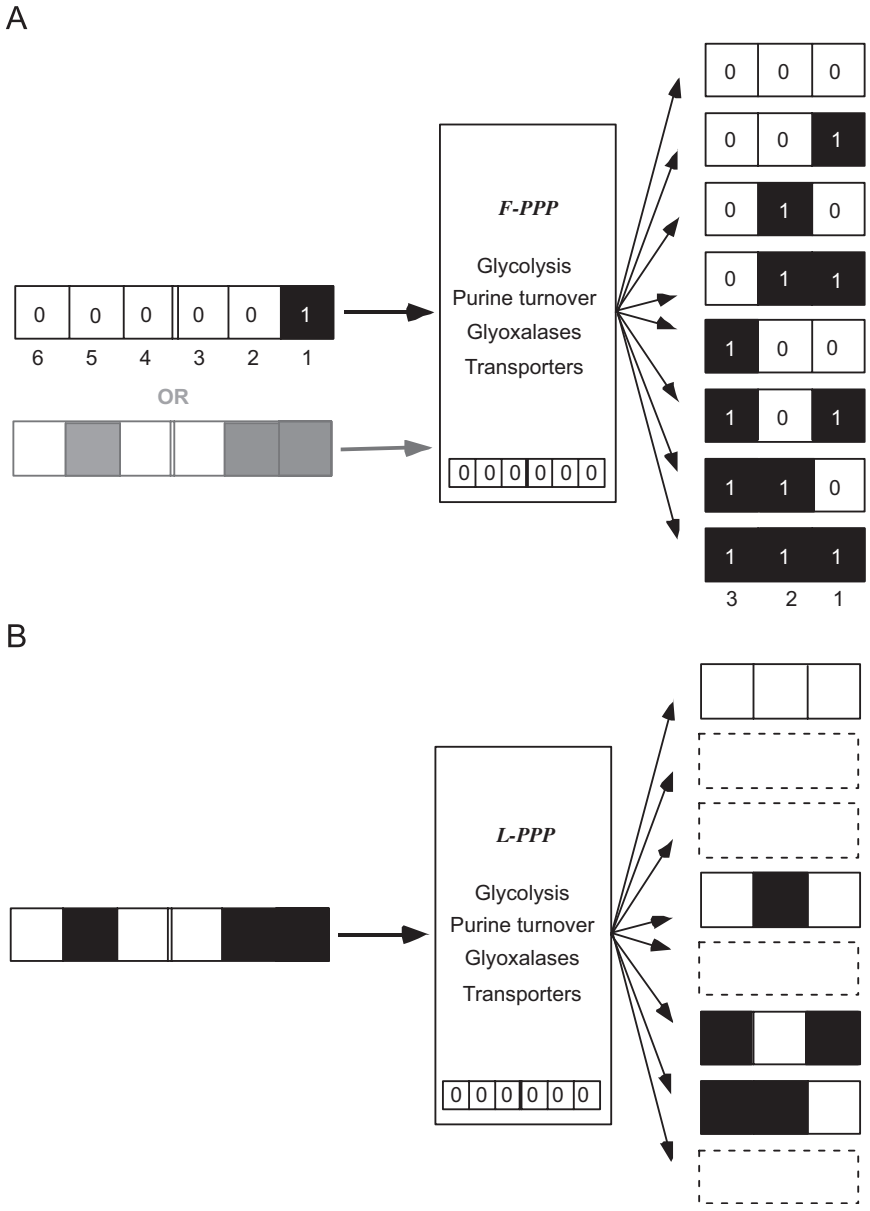

Fig. 4. Schematic representation of metabolism in the human RBC as a combinatorial scrambler of ^{13}C-labels in the input substrate glucose giving rise to isotopomers of the output, lactate. The input glucose is represented by a row of square boxes. (A) In the particular case shown in bold outline the glucose isotopomer is labelled in the C-1 position. The double vertical lines separating the boxes representing C-3 and C-4 were drawn to emphasize the reflection symmetry about the centre of the molecule that can operate in the reactions of the glycolytic/gluconeogenic pathway. In the *Mathematica* program, the presence of ^{13}C-label is indicated by the number '1' in a list while ^{12}C is denoted by '0', hence the use of these numerals in the boxes in the figure. The central large rectangle represents the metabolic system, and the six boxes in its base represent unlabelled glucose that will be mixed with the input isotopomer of glucose. The second grey representation of glucose is of an isotopomer at position (row 3 and column 4) in the array in Fig. 2, i.e., labelled in positions 1, 2 and 5, and also denoted by {0, 1, 0, 0, 1, 1} in the *Mathematica* program. It also gives rise to all possible isotopomers of lactate in a reaction scheme that includes the F-type PPP. Hence it is shown with the same output as the glucose represented above it in the diagram. (B) Glucose isotopomer labelled in positions 1, 2 and 5 as above, and it is seen from Fig. 3C to give only the four indicated isotopomers of lactate from the L-type PPP.

The elements of S are the isotopomers (say those of glucose) and the overall operator is the action of the set of carbon-scrambling enzymes. Our 'combinatorial scrambler' represented by the central large rectangle (box) in Fig. 4 is a shown here as a *binary* operator. It accepts a glucose isotopomer as one input and operates on it with endogenous *unlabelled* glucose as the second input. Unlabelled glucose is shown in Fig. 1A as the smaller series of six boxes at the bottom of the operator box. The output (result) of the operation (that entails rearrangement and re-combination of the carbon atoms) is a set of metabolites that have combinations of label that are different from those of the inputs.

An attempt to define the system as a *simple group* is complicated by the fact that in glycolysis the 6-carbon glucose molecule is split into two 3-carbon lactate molecules, and usually there are many more than just one possible isotopomer as output. To highlight the scission of glucose in the way mentioned, double lines are drawn between the third and fourth boxes on the left in Figs. 4A and 4B. However to make progress, we can juxtapose, in all possible combinations, the two 3-carbon units from the list of outputs. This recombination is akin to what takes place in gluconeogenesis from pyruvate; it does not occur in the human RBC but it does in the liver. The recombination is a valid procedure in the present mathematical investigation: if it is done with all eight of the lactate isotopomers shown in Fig. 4A it is readily seen that all 64 possible labelling patterns of glucose are obtained. On the other hand, for the particular choice of input shown in Fig. 4B, {0,1,0,0,1,1} with the L-type PPP version of the operator, the output is only four isotopomers; these are recombined in the manner suggested above to give the set {{0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 1, 0}, {0, 0, 0, 1, 0, 1}, {0, 0, 0, 1, 1, 0}, {0, 1, 0, 0, 0, 0}, {0, 1, 0, 0, 1, 0}, {0, 1, 0, 1, 0, 1}, {0, 1, 0, 1, 1, 0}, {1, 0, 1, 0, 0, 0}, {1, 0, 1, 0, 1, 0}, {1, 0, 1, 1, 0, 1}, {1, 0, 1, 1, 1, 0}, {1, 1, 0, 0, 0, 0}, {1, 1, 0, 0, 1, 0}, {1, 1, 0, 1, 0, 1} and {1, 1, 0, 1, 1, 0}}. Thus only 16 of the 64 possible isotopomers are produced, leaving a *deficit* of 48 possibilities. In other words, the F-type PPP generates only a *subset* of all possible isotopomers from the particular input shown in Fig. 4B. It does this for certain other inputs as well, as discussed above. Overall, though, with all possible input isotopomers the binary operation only produces output elements (using our re-combination procedure) that are in the input set; so the system has the property of *closure*.

The *associativity* requirement is satisfied, as the order of adding inputs is irrelevant.

An *identity* can be contrived by juxtaposing the appropriate two 3-carbon products to give the same labelling pattern as the original input; but this procedure involves a particular selection of pairs of outputs and excludes many other obvious recombinations; so the relationship between the number of products and the inputs is *supra*-1:1 and is not what is seen with a simple group. This is the first impasse.

Another impasse is encountered when trying to construct the *inverse* of one of the inputs. This would be tantamount to invoking the operation in a way that led to the complete loss of label to generate the identity element {0,0,0,0,0,0}. This can be achieved with single pairs of molecules in the oxidative PPP (via 6-phosphogluconate dehydrogenase) but under no circumstances does it occur with a whole ensemble of input molecules. So we conclude that there is no formal inverse.

While it was useful to try and represent the set of isotopomers and the 'combinatorial scrambler' operator as a group, this fails because of the inability to plausibly define a way of representing the inversion of an input isotopomer. On the other hand a *monoid* is a mathematical system that is a group without the inverse property; so our modified (with respect to representing an identity) system could be called a *monoid*. But, apart from providing a simple mathematical classification it appears that designating the system a monoid it appears to provide no insights into its properties and mathematical behaviour that help in the present work.

Appendix B. Outline of *Mathematica* program to calculate isotopomer distributions

Inputs

The aim was to develop a notation that accounts for all exchanges of label from substrates to products in a given reaction. Each reaction, whether actually enzymic or not, is represented as a *name*, a *list of reactants* (substrates and products) and the *label pattern* of each reactant. For example adenosine deaminase (adod) is entered as follows:

EnzymeReaction{adod, ReactantList[Reactant[Ado, AdoLabels:{_, _, _, _, _}]], ReactantList[Reactant[Amm, {}], Reactant[Ino, AdoLabels]]],

where Ado denotes adenosine, the underscore-list denotes the label pattern (in the ribose moiety) that will be filled-in prior to calculations. Amm denotes ammonia and the associated null set denoted by the {}-term indicates that no ^{13}C-label ever appears in this product and Ino denotes inosine with its label pattern being that of the substrate adenosine, namely AdoLabels.

The whole scheme of reactions is placed in a list:

allEnzymeReactions = {
{EnzymeReaction[{adod, ReactantList[Reactant[Ado, AdoLabels:{_, _, _, _, _}]], ReactantList[Reactant[Amm, {}], Reactant[Ino, AdoLabels]]] }],

{EnzymeReaction[{adok,], etc

};

Next the names of all reactants and the number of associated exchangeable carbon atoms are listed. E.g.,

allCarbons = { { Ado, 5}, …,{Amm, 0}, …, {Ino, 5},…, etc}

The *Mathematica* Utility that enables a subscript notation to be used for isotopomers is called, and then the specific notational transformations are listed:

<<Utilities 'Notation'

…

Notation[Ado $_{i5\_,\ i4\_,\ i3\_,\ i2\_,\ i1\_}$ ⇔ Reactant[Ino, { i5_, i4_, i3_, i2_, i1_}]]

…, etc.

Processing

One iteration of the program requires a list of available substrates (including details of their isotopomer variants present) as input. For each enzyme, all substrates that the enzyme uses are collated into a list. Each possible combination of substrates is passed into the enzyme, and the products are added to the list of available substrates for the next stage of the computation. Central to this operation is the outer product function and the intrinsic characteristic of *Mathematica* to order elements of Lists according to a strict hierarchy.

Outputs

The output is given in the form shown in Figs. 2 and 3 for any specified metabolite using the *Mathematica* functions Raster and GraphicsArray.

A copy of the program is available from the authors.

Appendix C. Reinhart Heinrich connections, scientific and personal

Philip Kuchel

Scientific

Until the trilogy of papers in the European Journal of Biochemistry in 1974 from Professor Sam Rapoport's group (Heinrich and Rapoport, 1974a, b; Rapoport et al., 1974), major efforts to simulate the operation of metabolic pathways had concentrated on the 'realistic' representation of schemes of enzyme-catalyzed reactions. These were based on the emerging and burgeoning literature on the kinetics of individual enzymes measured in vitro. The masters of this simulation craft were Higgins and Chance and their intellectual inheritors the Garfinkels (Garfinkel et al., 1970). The third in the series of 1974 papers dealt with erythrocyte glycolysis, but it cut through the complexities of the ~13 individual enzyme reaction mechanisms and focussed on a reduced model that captured the essence of the control features of the metabolic scheme. The emphasis had switched to the structure of the control network, and which features of the system affect steady-state concentrations in the pathway, and material flux through it.

When discussing the birth of metabolic control analysis (MCA) it is usual to quote the 1974 papers alongside the one by Kacser and Burns (1973) that also deals with flux control. But it is worth noting that Burns (1969) had already drawn attention to what was a normalized sensitivity coefficient for the control of substrate concentration; so there must be something else that identifies this later set of papers as the initiator of a paradigm shift.

A normalized sensitivity coefficient was renamed in the first 1974 paper as an 'element of the control matrix' (Heinrich and Rapoport, 1974a), thus giving the derived parameter an evocative name, along with the term 'control strength'. This was an important start in identifying a coherent area of Biochemistry with its own sub-language, but it is not the main reason; the main reason was the proof of two theorems that stated that the sum of the elements of the control matrix is zero, and the sum of the flux control coefficients is 1. Thus, mathematical rigour, new nomenclature, and successful application of the new MCA to the real human RBC system were the key to the projection of MCA to broader awareness in the biochemical community (e.g., Heinrich and Schuster, 1996).

My scientific efforts at the time (1973–1975) were directed at understanding the control of the urea cycle in mammalian liver. The computer model that was written (in Fortran with a numerical integrator to handle stiff differential equations) simulated the time dependence of metabolite concentrations in normal and disease states (Kuchel et al., 1977). It was developed along the lines of Garfinkel's models. The passage of the emergent manuscript to press was not smooth, and when I moved to Oxford I learnt why (but I was also well aware of the limitations too!): Sir Hans Krebs commented (I paraphrase) that so little was known about the operation of enzymes inside cells that methods to study this were sorely needed, and only then might the models be of some value. Around this time the first non-invasive ^{31}P NMR experiments were reported on ATP levels and pH in whole erythrocytes and myocytes (Moon and Richards, 1973; Hoult et al., 1974) and Oxford Biochemistry was abuzz with excitement with the prospect of studying metabolism in vivo. The first ^{1}H NMR studies of metabolism were reported after a means of minimizing the overwhelming signal from water from the spectra was implemented (Brown et al., 1977). Then a vast set of metabolites could be monitored in suspensions of intact erythrocytes and other cells and the much sought after kinetic data could begin to be amassed. In fact, much additional information was obtained by acquiring ^{1}H, ^{13}C and ^{31}P NMR spectra in an interleaved way and this approach provided data sets that formed the basis of our own model of human erythrocyte metabolism (Mulquiney et al., 1999; Mulquiney and Kuchel, 1999a, b, 2003). In the build up to our detailed model of human erythrocyte metabolism valuable data had been obtained on rates of enzyme-catalyzed reactions in the pentose phosphate pathway by using ^{13}C-labelled glucose and analysis of isotopic isomers (isotopomers) that were formed over time. Thus our model is based on a consistent set of data mostly obtained with NMR spectroscopy; but the numerical model is also

amenable to analysis using all the ideas and coefficients of MCA, and stability analysis that was is well described in the 1977 review from the Berlin group (Heinrich et al., 1977; Mulquiney and Kuchel, 2003).

Personal

That Reinhart Heinrich and I were distant relatives is an interesting personal aside: it is based on the fact that my grandmother on my father's side of our family was Ottilie Agnes *Heinrich*, born in 1887 at Kilkerran, South Australia.

I first met Reinhart at the Xth International Berlin Symposium: Structure and Function of Erythroid Cells, August 29–September 1, 1983; and after discussions with him we suspected that there may be a family link, so when I returned to Berlin in 1986 I took with me a copy of the recently compiled book on the Heinrich family in Australia by Rhoda Heinrich (Heinrich, 1979). A family of Heinrichs (parents plus four children) had emigrated as Lutheran religious refugees from Silesia in 1839 to the free colony of South Australia; there the successive families prospered during the 19th century mostly in farming, and during the 20th century in a host of other professions. Reinhart was captivated by a possible distant, spatial and temporal link and it was his father, a retired professor of mathematics in Dresden (Helmut Heinrich), who listed the genealogy and the likely common ''Urururgrossvater Ernst Phillip Heinrich'' born in 1786 in Herzogwald near Seiffersdorf in Silesia; the latter village name was re-used in the Barossa Valley of South Australia where many of the German pioneers settled.

Needless to say this, albeit distant, connection (which for any of us would exist for other acquaintances, but probably in a less-well documented way) was a facet of our friendship that added to matters which would have brought us in contact anyway, namely common areas of interest in metabolic science and a deep curiosity in what makes cells 'tick'.

References

Barnett, M.P., Capitani, J.F., von zur Gathen, J., Gerhard, J., 2004. Symbolic calculation in chemistry: selected examples. Int. J. Quantum Chem. 100, 80–104.

Berthon, H.A., Bubb, W.A., Kuchel, P.W., 1993. ^{13}C NMR isotopomer and computer simulation studies of the non-oxidative pentose phosphate pathway of human erythrocytes. Biochem. J. 296, 379–387.

Brown, F.F., Campbell, I.D., Kuchel, P.W., Rabenstein, D.L., 1977. Human erythrocyte metabolism studies by ^{1}H spin echo NMR. FEBS Lett 82, 12–16.

Burns, J.A., 1969. Steady states of general multi-enzyme networks and their associated properties. Computational approaches. FEBS Lett. 2, S30–S33.

Chance, E.M., Seeholzer, S.H., Kobayashi, K., Williamson, J.R., 1983. Mathematical analysis of isotope labelling in the citric acid cycle with applications to ^{13}C NMR studies in perfused rat hearts. J. Biol. Chem. 258, 13785–13794.

Cowin, G.J., Willgoss, D.A., Bartley, J., Endre, Z.H., 1994. Serine isotopomer analysis by C-13-NMR defines glycine–serine interconversion in situ in the renal proximal tubule. Biochim. Biophys. Acta 1310, 32–40.

Garfinkel, D., Garfinkel, L., Pring, M., Green, S.B., Chance, B., 1970. Computer applications to biochemical kinetics. Ann. Rev. Biochem. 39, 473–498.

Heinrich, R., 1979. Heinrich of Seiffersdorf 1839–1979 Gottlob and Anna Rosina Heinrich and their Descendants 140 Years in Australia. Lutheran Publishing House, Adelaide, South Australia.

Heinrich, R., Rapoport, T.A., 1974a. A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur. J. Biochem. 42, 89–95.

Heinrich, R., Rapoport, T.A., 1974b. A linear steady-state treatment of enzymatic chains. Critique of the cross-over theorem and a general procedure to identify sites with an effector. Eur. J. Biochem. 42, 97–105.

Heinrich, R., Rapoport, T.A., Rapoport, S.M., 1977. Metabolic regulation and mathematical models. Prog. Biophys. Mol. Biol. 32, 1–82.

Heinrich, R., Schuster, S., 1996. The Regulation of Cellular Systems. Chapman & Hall, New York.

Horecker, B.L., 2002. The pentose phosphate pathway. J. Biol. Chem. 277, 47965–47971.

Hoult, D.I., Busby, S.J.W., Gadian, D.G., Radda, G.K., Richards, R.E., Seeley, P.J., 1974. Observation of tissue metabolites using ^{31}P nuclear magnetic resonance. Nature 252, 285–287.

Kacser, H., Burns, J.A., 1973. The control of flux. Symp. Soc. Exp. Biol. 32, 65–107.

Kauffman, K.J., Pajerowski, J.D., Jamshidi, N., Palsson, B.O., Edwards, J.S., 2002. Description and analysis of metabolic connectivity and dynamics in the human red blood cell. Biophys. J. 83, 646–662.

Kuchel, P.W., Roberts, D.V., Nichol, L.W., 1977. The simulation of the urea cycle: correlation of effects due to inborn errors in the catalytic properties of the enzymes with clinical-biochemical observations. Aust. J. Expt. Biol. Med. Sci. 55, 309–326.

Kuchel, P.W., 2004. Current status and challenges in connecting models of erythrocyte metabolism to experimental reality. Prog. Biophys. Mol. Biol. 85, 325–342.

London, R.E., 1988. ^{13}C labeling in studies of metabolic regulation. Prog. NMR Spectr. 20, 337–383.

McIntyre, L.M., Thorburn, D.R., Bubb, W.A., Kuchel, P.W., 1989. Comparison of computer simulations of the F-type and L-type non-oxidative hexose monophosphate shunts with ^{31}P-NMR experimental data from human erythrocytes. Eur. J. Biochem. 180, 399–420.

Maher, A.D., Kuchel, P.W., Ortega, F., Atauri, J.C., Cascante, M., 2003. Mathematical modelling of the urea cycle. Eur. J. Biochem. 270, 3953–3961.

Malloy, C.R., Sherry, A.D., Jeffrey, F.M.H., 1988. Evaluation of carbon flux and substrate selection through alternate pathways involving the citric acid cycle of the heart by ^{13}C NMR spectroscopy. J. Biol. Chem. 263, 6964–6971.

Meléndez-Hevia, E., Isidoro, A., 1985. The game of the pentose phosphate cycle. J. Theoret. Biol. 117, 251–263.

Montero, F., Nuño, J.C., Valdenebro, I., Pérez-Iratxeta, C., Meléndez-Hevia, E., 1997. Stoichiometric properties of the non oxidative phase of the pentose phosphate cycle. Nonlinear Anal. Meth. Appl. 30, 1865–1874.

Moon, R.B., Richards, J.H., 1973. Determination of intracellular pH by ^{31}P magnetic resonance. J. Biol. Chem. 248, 7276–7278.

Mulquiney, P.J., Bubb, W.A., Kuchel, P.W., 1999. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations 1: in vivo kinetic characterisation of 2,3-bisphosphoglycerate synthase/phosphatase using ^{13}C and ^{31}P NMR. Biochem. J. 342, 567–580.

Mulquiney, P.J., Kuchel, P.W., 1999a. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations 2: equations and parameter refinement. Biochem. J. 342, 581–596.

Mulquiney, P.J., Kuchel, P.W., 1999b. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations 3: computer simulation and metabolic control analysis. Biochem. J. 342, 597–604.

Mulquiney, P.J., Kuchel, P.W., 2003. Modelling Metabolism with Mathematica: Detailed Examples Including Erythrocyte Metabolism. CRC Press, Boca Raton, FL, USA, pp. 1–309.

Rapoport, T.A., Heinrich, R., Jacobasch, G., Rapoport, S., 1974. A linear steady-state treatment of enzymatic chains. Critique of the cross-over theorem and a general procedure to identify sites with an effector. Eur. J. Biochem. 42, 97–105.

Rapoport, T.A., Heinrich, R., Rapoport, S.M., 1974. The regulatory principles of glycolysis in erythrocytes in vivo and in vitro. A minimal comprehensive model describing steady states, quasi-steady states and time-dependent processes. Biochem. J. 154, 449–469.

Selivanov, V.A., Puigjaner, J., Sillero, A., Centelles, J.J., Ramos-Montoya, A., Lee, P.W.-N., Cascante, M., 2004. An optimized algorithm for flux estimation from isotopomer distribution in glucose metabolites. BioInformatics 20, 3387–3397.

Selivanov, V.A., Meshalkina, L.E., Solovjeva, O.N., Kuchel, P.W., Ramos-Montoya, A., Kochetov, G.A., Lee, P.W.-N., Cascante, M., 2005. Rapid simulation and analysis of isotopomer distributions using constraints based on enzyme mechanisms: an example from HT29 cancer cells. BioInformatics 21, 3551–3564.

Walker, T.E., London, R.E., Whaley, T.W., Barker, R., Matwiyoff, N.A., 1976. Carbon-13 nuclear magnetic resonance spectroscopy of [1-^{13}C]enriched monosaccharides. Signal assignments and orientational dependence of geminal and vicinal carbon–carbon and carbon–hydrogen spin–spin coupling constants. J. Am. Chem. Soc. 98, 5807–5813.

Walker, T.E., London, R.E., 1987. Biosynthetic preparation of L-[^{13}C]- and [^{15}N]glutamate by Brevibacterium flavum. Appl. Environ. Microbiol. 53, 92–98.

Walker, T.E., Han, C.H., Kollman, V.H., London, R.E., Matwiyoff, N.A., 1982. ^{13}C nuclear magnetic resonance studies of the biosynthesis by Microbacterium ammoniaphilum of L-glutamate selectively enriched with carbon-13. J. Biol. Chem. 257, 1189–1195.

Wiechert, W., 2001. ^{13}C metabolic flux analysis. Metab. Eng. 3, 195–206.

Wiechert, W., Möllney, M., Petersen, S., de Graaf, A.A., 2001. A universal framework for ^{13}C metabolic flux analysis. Metab. Eng. 3, 265–283.

Williams, J.F., Arora, K.K., Longenecker, J.P., 1987. The pentose pathway: a random harvest. Impediments which oppose acceptance of the classical (F-type) pentose cycle for liver, some neoplasms and photosynthetic tissue. The case for the L-type pentose pathway. Int. J. Biochem. 19, 749–817.

Wolfram, S., 2003. The Mathematica Book, third ed. Wolfram Media, Champaign, IL.

Zupke, C., Stephanopoulos, G., 1994. Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices. Biotechnol. Prog. 10, 489–498.