transactions

Biochemical Society Transactions www.biochemsoctrans.org

Systems Analysis of Metabolism

A Biochemical Society Focused Meeting held at University of York, U.K., 22–24 March 2010, as part of the Systems Biochemistry Linked Focused Meetings. Organized and Edited by David Fell (Oxford Brookes, U.K.), Hans Westerhoff (Manchester, U.K., and Amsterdam, The Netherlands) and Michael White (Liverpool, U.K.).

Systems biochemistry in practice: experimenting with modelling and understanding, with regulation and control

Hans V. Westerhoff*†[1], Malkhey Verma*, Maria Nardelli*, Malgorzata Adamczyk*, Karen van Eunen†‡, Evangelos Simeonidis* and Barbara M. Bakker‡

*Manchester Centre for Integrative Systems Biology, Manchester Interdisciplinary Biocentre, University of Manchester, 131 Princess Street, Manchester M1 7DN, U.K., †Department of Molecular Cell Physiology, Netherlands Institute for Systems Biology, VU University Amsterdam, de Boelelaan 1085, NL-1081 HV Amsterdam, The Netherlands, and ‡Department of Paediatrics, Centre for Liver, Digestive and Metabolic Diseases, University Medical Centre Groningen, University of Groningen, Hanzeplein 1, 9713 GZ, Groningen, The Netherlands

Abstract

Biology and medicine have become 'big science', even though we may not always like this: genomics and the subsequent analysis of what the genomes encode has shown that interesting living organisms require many more than 300 gene products to interact. We once thought that somewhere in this jungle of interacting macromolecules was hidden the molecule that constitutes the secret of Life, and therewith of health and disease. Now we know that, somehow, the secret of Life is the jungle of interactions. Consequently, we need to find the Rosetta Stones, i.e. interpretations of this jungle of systems biology. We need to find, perhaps convoluted, paths of understanding and intervention. Systems biochemistry is a good place to start, as it has the foothold that what goes in must come out. In the present paper, we review two strategies, which look at control and regulation. We discuss the difference between control and regulation and prove a relationship between them.

Why systems biology?

Although it has become possible to characterize almost all its molecular components, no comprehensive function of any living organism has yet been understood in terms of a fully predictive biology. Thereby the implicit promise behind much of the life sciences, i.e. that we are working towards completely rational therapies of all our diseases and more generally towards a rational approach to Life, seems to be left unmet.

This broken promise becomes bitter as molecular biology, cell biology and genomics thrive in their successes. Progress from almost complete ignorance a century ago to the accumulated data and understanding of today has been enormous. However, the distance between where we are now and where we should be seems to have increased even more. With every new paper about p53, we understand less of its function [1]. We do not see the forest for the trees, and the information about the individual trees has been growing exponentially. Reasons are the enormous complexity of the subject matter and, perhaps, a wrong approach to that complexity. Systems biology more than the biology of living systems should help get us out of this predicament. It should perhaps utilize the natural organization of living systems to develop analysis methods that thereby deviate from those of chemistry and physics and help clarify the mysteries of Life.

The theoretical complexity of biology is vast

When setting up or deciding about the funding of research programmes, it becomes increasingly important to ensure that they follow a sensible strategy. When aiming to understand the common basis of all physical forces in terms of an as yet unidentified gauge boson [2], it makes little sense to study a pendulum in one's backyard. One probably needs to make elementary particles collide at huge energies, requiring an enormous accelerator at vacuum. Thus an enormous project emerges, involving lots of laboratories all over the world.

Key words: complexity, control analysis, regulation analysis, systems biology, yeast.

[1]To whom correspondence should be addressed (email Hans.Westerhoff@manchester.ac.uk).

©The Authors Journal compilation ©2010 Biochemical Society

Most of these laboratories do what has to be done, in a single context, using common standards.

Biology and molecular biology lack the tradition of 'big science'; they have been more like cottage industries. Indeed, there have been and there remain good reasons for working on certain components of living organisms, for instance on the structure of an important protein [3], on the genes' downstream binding sequences for a transcription factor or on the mechanisms by which chromatin modification may affect lyonization [4]: the structures and molecular mechanisms are both fascinating and important scientifically. An at least equally important aim of biology, however, is the understanding of Life, of functions that contribute to Life or of malfunctions that lead to disease. In a perhaps too literal reading of the word biology, this is what biologists are supposed to do, i.e. figure out the words and laws that gauge Life. Ultimately, this aim may require an even bigger science than that required for the Higgs boson. For the human is likely to be much more complex than the relationship between the four fundamental forces, and much more important for humankind as well.

Genomics has contributed a number of mind-shattering understandings of the living state. One is that the genome sizes of living organisms that we find interesting (e.g. humans, yeast or *Escherichia coli*) amount to thousands of genes. Already in this sense, biology is complex. Rather than the, say, three elementary particles that may feature at some relevant level in physics, or the 118 that do so in chemistry, one is dealing with at least 30 000 such elementary particles for the human.

Be that as it may, one could engage in an extensive structural biology initiative and overexpress, purify, crystallize and then determine the X-ray structure of each of the 30 000 proteins [5]. This is much and highly difficult work, because most proteins of a genome differ in structure and molecular mechanism. This feature corresponds to the complexity that is often associated with biology, i.e. much, much detail and little generality, and the project corresponds to a seemingly endless project of data collecting.

However, if chemistry would have been satisfied with determining the structures and material properties of all of the elements, this would not have led to as much understanding of the world as chemistry has contributed today. Most substances that affect the world are not elements, but compounds. Much of the relevance of material chemistry resides in the elements reacting with each other to form new entities; an infinite number of combinations.

Similarly, the macromolecules of biology function exclusively through their interactions with other macromolecules. The cell-division cycle is one of the most fundamental and characteristic processes of Life. Mathematical models exist in which a number of proteins interact, effecting covalent modification and degradation of each other [6]. This entails cell cycling. None of the proteins in isolation would cycle, but the ensemble of interacting proteins does. Accordingly, a phenomenon such as the transition of the cell through certain characteristic points of the cell cycle appears to be controlled, not by a single protein, but by multiple proteins [7]. Oxidative phosphorylation is the main source of the ATP that drives virtually every process sufficiently away from equilibrium to provide flux that is relevant for Life. The ATPase responsible for the synthesis of the ATP is driven by protons tumbling down their concentration gradient/Gibbs free energy difference across the inner mitochondrial membrane [8] if it is thermodynamically proficient [9]. The process would run in the wrong direction if only this enzyme were involved. The mitochondrial electron-transfer chain is essential to crank up that electrochemical gradient. None of the participating proteins in isolation would catalyse the synthesis of ATP, only their network does [10]. Similarly, β-cell function, in the sense of the secretion of insulin in response to increased levels of glucose, would not happen if only the glucose-sensing system were in place or only the insulin production machinery. These are a mere three examples of the fact that functions of living organisms depend critically on the interactions between components.

In an additional sense, all biological functions ultimately depend on interactions. All functions depend on the maintenance of the living state of which they are part, which in turn depends on a minimum number of interacting gene products. This principle of Life is reinforced by the smallest genome size of living organisms being not smaller than some 400 genes [11,12]. The significance of this is that the smallest form of Life does not consist of 100 organisms of four genes each, but that all these 400 genes have to be together in a single organism for any of the individual genes to survive (*pace* [13]).

The number of potential binary interactions of 30 000 components is almost a billion, and if interactions can involve any number of components, each only once, the number of theoretically possible interactions ($30\,000! \approx 10^{121202}$, i.e. a 1 with more than 120 000 zeros) exceeds the number of atoms in the Universe (which is on the order of 10^{80}) [14]. For the smallest genome we know, the numbers are 400!, amounting to 10^{866}. Because these numbers exceed the number of genes astronomically, and because 30 000 would be manageable for a large research programme, but 10^{121202} would not be, the problem in the complexity of biology resides less in the large number of components than in the importance of their interactions for function.

The actual complexity of biology might be manageable

Are living organisms so complex as to depend on more than 10^{866} interactions? If so, then the experimental assessment of the strengths of these interactions would be impossible within the lifetime of the human species, and well beyond possible resourcing. Likewise, the mathematical modelling of the corresponding networks would require computers that cannot be envisaged to be built or even powered. In this section, we show that the complexity is far smaller than suggested by this astronomical number.

Living organisms engage in diverse functions that are often much more sophisticated than the function of an ideal gas.

©The Authors Journal compilation ©2010 Biochemical Society

It might seem that this sophistication might require the enormous complexity calculated above. However, the part of non-equilibrium thermodynamics that deals with self-organization has shown otherwise: through the non-linear interactions of a very limited number of components, highly complex patterns can emerge [15], which are reminiscent of the patterning in developmental biology [16]. This means that, although many biological functions are complex, they might depend on few components and interactions [15].

This now constitutes a paradox. If generating strong complexity requires so few components and interactions, why do living organisms require so many (>400) components? And, do they really require as many as the 400! possible interactions between these components? This paradox is profound and affects much of the methodology of the life sciences. On the one hand, some life scientists view biology as a can of worms, every new case requiring a *de novo* acquisition of experimental information where conclusions can only be drawn empirically and where there is no role for the application of generic laws and principles. On the other hand, (other) life scientists wish the science of biology to become a predictive science, where the application of more generic principles and laws to actual cases suggests hypotheses that are then tested experimentally and rigidly. When projected to the extreme, the former group sees biology as a can of >400! worms, whereas the latter sees it as in essence just four or five general principles being instantiated in large numbers of special cases. We argue that neither view is correct, that biology is neither simple nor maximally complex, and that an additional line of thinking, i.e. hysteresis through evolution, is unique for the life sciences, which thereby requires an entirely new methodology [17,18].

By accounting for the minimal requirements for Life, for the number of different chemical compounds required, for the fact that biochemical synthesis occurs in sequences of simple chemical reactions and for the required micro-environment for catalysis [19], one finds that Life requires more than 100 reaction steps, perhaps 300 [18]. In linear biochemical pathways, each protein interacts with two other proteins through an intermediary metabolite, leading to a number of (indirect) protein–protein interactions that is equal to the number of enzymes. Because of the requirement of free energy transduction and because of convergence of reactions, the actual number of such interactions would be a bit higher. The interactions that underpin the chemistry of living organisms are protein–metabolite interactions. If there are only linear pathways, one has $2n$ enzyme–metabolite interactions when there are n enzymes. For the 30 000-gene human, this would mean some 60 000 real interactions rather than the 10^{121202} theoretical ones (see above); for the minimum organism, 800 rather than the 400!.

This estimate of 60 000 may be a bit conservative, because there are also reactions that have more than one substrate, e.g. those involving free-energy or redox coenzymes, and there are ample indications that at least some enzymes interact with other enzymes or with other active proteins in the cell (see, e.g., [20]). However, even though one might then require 10^5 interactions for the human, this number is astronomically smaller than the theoretical number of 10^{121202}.

There are other reasons that Life is less complex than it could have been: looking at equal-size spheres, we find that any one can only interact with six others at the same time. In a multi-tissue organism, not all genes are expressed in all tissues. In micro-organisms, not all genes are expressed at the same time. Yet, this differential expression in space and time gives evolutionary advantage. The total number of genes expressed at any one time in any human tissue might well be as low as 4000. This decreases the maximum number of binary interactions to some 10 million; very many, but perhaps not too many for a very large research programme of 'big biology'. More important than the precise magnitude of these numbers, however, is the suggestion that living organisms are not as complex as the theoretical number of 400! suggested.

Why then did Life not grow as complex as it could have? Evolution presumably began at a suboptimal state and then moved to a variety of better states, adapted to ecological niches. Because of strong selection pressure, there has been limited horizontal evolution, i.e. at each intermediate stage of optimality, there have been insufficient mutations to sample all possibilities of improvement. If evolution is seen as the movement down from the top of a mountain, then it cannot be predicted which side of the mountain, hence in which of the many valleys surrounding the mountain, evolution ends up. Evolution cannot be seen as a Markov process, i.e. the state that will be achieved ultimately will not be determined completely by the ecological niche, by all the possibilities of chemistry and physics and by selection for optimality. It will also be determined by accidental 'choices' during the process of evolution. An example is the fact that all amino acids in proteins are of the L-stereoisomeric form, whereas mirror organisms with all D-amino acids would have been equally viable in terms of physics and chemistry. The difference with physics, or at least with the classical picture thereof, is that in biology one cannot expect to solve a Schrödinger type of equation for minimum energy solutions and find biological reality after evolution as the theoretical minimum energy state.

Living systems are not at equilibrium or even near equilibrium [21]. Even where relatively simple, linear, dynamic relationships exist [22], these relationships differ from those that could be derived from near-equilibrium non-equilibrium thermodynamics [23]. Often the system is 'pumped up' to a steady state far away from equilibrium by continuous expenditure of free energy harvested from food. The properties of such a state do not correspond to the non-mechanistic properties near equilibrium [24]. They reflect mechanisms and regulation [25] that may have been selected by evolution. Yet, within a basin of attraction (a valley in the mountainous landscape), evolutionary optimization may have occurred to some extent. Provided that certain limitations imposed by evolutionary history (such as L-amino acids) are taken into account, one may be able to predict the behaviour of the system. In addition, in different species, some of the evolutionary pressures are identical and hence

©The Authors Journal compilation ©2010 Biochemical Society

some generic principles may apply across species. And then, because of the conservative nature of evolution discussed above, there is substantial homology, which also has the effect of significant predictability (e.g. the prediction that a newly observed bacterial species will have DNA with essentially only four bases, essential in the sense that the four bases are used in DNA synthesis, and allowing for moderate levels of subsequent modifications).

Our suggestion is that experimental insight into how biological systems happen to work reduces the potentially astronomical complexity to perhaps manageable actual complexity. The true complexity of Life is not so high as to make its modelling unthinkable. The understanding of Life will not come from a grand minimization plus simplicity principle applied to all degrees of freedom of living systems [18]. It will require analysis in terms of the actual components of, and the actual interactions in, actual living organisms. Therewith experimental biology is a crucial component of systems biology [26].

Why systems biochemistry?

Biology is a well-organized combination of chemistry and a bit of physics. Chemistry is everywhere in Life. Gene expression involves transcription and translation, i.e. the chemical association of nucleotides and amino acids. Signal transduction leads to activation or inactivation by covalent modification. One can look at all of this as the gigantic single network of chemical reactions that it is.

One can also look at this in a way that distinguishes between two dimension types, i.e. the 'horizontal' ones of chemical conversions and the 'vertical' ones of information transfer [27,28]. The synthesis and degradation of any mRNA is a horizontal dimension in which a molecule is synthesized and degraded. The synthesis, covalent modification and degradation of a protein is another such chemical dimension. And the rather extensive connected metabolic network of living organisms is yet another chemical dimension. These dimensions are connected essentially through 'allosteric' interactions, i.e. influences that do not permanently transfer mass. A specific mRNA directs translation of amino acids into the protein it encodes, without itself being consumed in the process. That protein may catalyse a metabolic reaction, in which it is not itself consumed. Of course, the level of protein metabolism is connected by mass transfer with the level of intermediate metabolism, because protein synthesis requires amino acids as substrates, but we neglect this aspect: the biological essence of enzymes is to be catalysts.

If, for now, we restrict the term 'systems biochemistry' to the part of systems biology that proceeds at any horizontal, chemical, dimension, we note that it has a property that is absent from the vertical, information-transfer, dimension. This is the property that is often called 'mass conservation', although it is, in fact, quite a bit stronger than that. It is the property that at any reaction in a chemical dimension, the amount of each chemical element (and sometimes even a constellation of elements called a moiety [29]) is conserved.

Does it make sense to distinguish this systems biochemistry from mainstream systems biology? It would if it led to something special. And this is what it does: one result of this distinction between systems biochemistry and systems biology deals with control properties of the network. It is fairly well known that the sum of the control exerted by all of the reaction steps in the total network on any steady-state flux in that same network, is always 100 % [30,31]. This gives one law that governs the systems biology of an entire living cell. It is perhaps less known that within each chemical dimension the same is true, e.g. the sum of the control of all reactions in the chemical dimension of a certain mRNA molecule on the steady-state flux of synthesis of that mRNA must also equal 100 % [28,32,33]. Perhaps even less familiar is the law that the total control of all reactions in one chemical dimension on the steady-state flux in any other dimension must be zero. An implication is that control by transcription, if important, is not more important than control by mRNA degradation, a phenomenon that everyone knew all along . . .

Systems biochemistry has also inspired a second approach to systems biology, called flux analysis. This uses the conservation of chemical elements to identify so-called 'stoichiometric' constraints to the possible behaviour of metabolic networks [34]. The approach provides improved definitions of what are metabolic pathways, in terms of extreme pathways [35] and elementary modes [36,37]. In its analytical mode, flux analysis enables one to determine fluxes through networks by measuring the external disappearance or appearance of substances. In an even more powerful form, it uses mass-isotope-labelled internal substances to determine internal fluxes [38]. A related method, termed FBA (flux-balance analysis), serves a more predictive function. It deals with the redundancy of possible solutions of non-isotope doped flux analysis by making assumptions of optimality, such as maximum yield [39]. The method has appreciable predictive power, although it is also limited, unless enhanced further [40–42].

Regulation analysis

The distinction between the horizontal (chemical, metabolic) and the vertical (gene-expression) dimension also led to another extension of methodology, now to address an important issue in systems biology. Analysis of cell function had bifurcated into two large areas. One was that of biochemistry and focused on intermediate and energy metabolism. The other was called molecular genetics and looked much at transcription and genes. The two areas of molecular biology discussed cell function largely in the absence of mention of the other, partly because methodology to integrate the two was lacking. In intermediate metabolism, regulation was discussed in terms of the activation of rate-limiting enzymes. In molecular genetics, consensus sequences were discussed that would or would not bind transcription factors that activated transcription, which would then lead to the synthesis of yet other transcription factors which would again activate, or perhaps inactivate, the transcription of some other transcription-factor-encoding gene.

©The Authors Journal compilation ©2010 Biochemical Society

Sauro [43] had proposed a way to describe regulation quantitatively. He wrote an equation for the rate of a biochemical reaction in terms of its substrates, products and other allosteric or competitive effectors. The regulation of the rate of that enzyme-catalysed process could be disentangled in terms of the contributions of these various effectors. This analysis did not discriminate between the horizontal and vertical dimensions. Ter Kuile and Westerhoff [44] then added that distinction and thereby came to integrate metabolic biochemistry with gene-expression molecular genetics. They found that total regulation could at least in some way be seen as the sum of biochemical and molecular genetic contributions: in a new law, now for regulation of the steady-state rate of a process, the sum of the direct regulation of an intracellular reaction rate by the vertical dimension, i.e. gene expression, and by the horizontal dimension, i.e. metabolic factors, equalled 1 [44]. The method has since been extended to time-dependent hierarchical regulation analysis [45–47].

Regulation versus control

In passing, ter Kuile and Westerhoff [44] brought home a message that was left a bit implicit in the early paper by Sauro [43]: although often the words 'control of cell function' and 'regulation of cell function' are used indiscriminately, they tend to refer to two concepts rather than one. Before the advent of Metabolic Control Analysis, discussions of control or regulation of metabolism had been about phenomena and features that were important for the subject, such as futile cycling, the presence of many metabolic factors affecting an enzyme such as phosphofructokinase, feedback loops, position in a pathway and irreversibility [48], but did not by themselves create a comprehensive understanding. The definitions and laws [49–54] of Metabolic Control Analysis then put this to an end, especially after they had been implemented experimentally (see, e.g., [55–59]).

An ambiguity remained, however, which we illustrate by discussing the importance of an enzyme for the flux through the pathway of which it is part. One may ask whether that enzyme is the rate-limiting step, or more subtly to what extent that enzyme controls or limits the flux. The operational definition of this concept of control is the effect of a small modulation of the activity of the enzyme (e.g. its concentration, but see [60]) at constant activities of all other enzymes, on the flux [61]. If for a 10 % activation of the enzyme, the effect on the steady state flux is 7 %, then the flux control coefficient of that enzymes with respect to that flux is 0.7 [61]. Here the modulation of the activity of the enzyme is not carried out by the system itself; it is the experimenter, either *in vitro* (see, e.g., [55,62]) or *in silico* (see, e.g., [63,64]) who modulates. The magnitude of the flux control coefficient bears no reference to the question of whether or not the enzyme activity is now or ever increased or decreased by or in the organism, other than by this external intervention aimed at measuring the control.

The issue whether the organism itself does or does not alter the activity of the enzyme, for instance when trying to cope with an external challenge such as shortage of substrate for energy metabolism, is a different, but likewise interesting, issue. In order to distinguish from the above concept of control, we have proposed to restrict the word 'regulation' to dealing with this issue [44,65]. The word 'control' had been pretty well defined to describe the above concept of 'limitation' [61].

Rather than to define regulation as the change effected by the organism, we proposed to relate the change to what is being regulated. Perhaps the clearest case is that of the regulation of the flux through a chemical reaction. An organism can change such a flux in three types of ways. One is by changing the concentrations of the metabolites that affect the rate at which the enzyme that catalyses the reaction functions. A second is by changing the concentration of that enzyme. And the third is by changing the catalytic properties of the enzyme by irreversible (not necessarily irrevertible [23]) modification, such as adenylylation [66].

The first of the three is of 'horizontal' systems biochemistry and corresponds directly to metabolic regulation. The second corresponds to the most direct form of regulation through gene expression. The third may be called regulation through signal transduction, but we have not yet dealt with this component explicitly; until now it has been grouped together with gene expression regulation, often under the term 'hierarchical regulation' [44,45].

The gene expression regulation is quantified by taking the relative change in enzyme concentration or (in the case of hierarchical regulation coefficient comprising both gene expression and signal transduction regulation) enzyme V_{max} and divide this by the relative change in regulated function; in simple cases, the flux through the same enzyme. At least, this is what is done for small changes. For larger changes, the ratio of the differences of the logarithms is taken.

For most enzyme-catalysed reactions, the rate equation can be written as a multiplication of the enzyme concentration and a function of metabolite concentrations [67]. The metabolic regulation coefficient is defined as the relative change in the latter function divided by the relative change in flux. For larger changes, this is calculated as the change in logarithm of that function divided by the change in logarithm of the flux. If regulation is quantified in this way, the hierarchical regulation and the metabolic regulation coefficients sum to 1 also for larger changes in flux. When gene expression regulation is split into components such as transcription regulation and mRNA stability regulation, similar laws apply [68].

The connections between regulation and control

In the regular case, where the rate of a reaction is proportional to the concentration of the enzyme catalysing it, the flux control coefficient is the change in the logarithm of the

©The Authors Journal compilation ©2010 Biochemical Society

flux divided by the change in the logarithm of the concentration of the enzyme:

$$C_i^{J_i} \equiv \left(\frac{\partial \ln [J_i]}{\partial \ln(e_i)}\right)_{e_{j\neq i},\ \text{steady state}}$$

The gene-expression regulation coefficient is the change in the logarithm of the enzyme concentration divided by the change in the logarithm of the flux:

$$\rho_g \equiv \left(\frac{\Delta \ln [e_i]}{\Delta \ln(J_i)}\right)_{\text{steady state}}$$

Capital Δ refers to both to changes of substantial and to changes of small magnitude. In the latter case, the above equation can be written as $\rho_g \equiv \frac{d \ln [e_i]}{d \ln(J_i)}$, which suggests that the two coefficients might be each other's inverse, or that $\rho_g \cdot C_i^{J_i} = 1$. But are they each other's inverse? Is gene-expression regulation the inverse of control?

The symbol ∂ refers to the fact that a partial derivative is considered, i.e. enzyme e_i is modulated at constant concentrations (activities) of all other enzymes (as indicated in the subscript and at constant magnitudes of all other parameters of the system): as mentioned below, control by an enzyme is the phenomenon that if an experimenter modulates the activity of that enzyme, then the flux changes considerably. Of course, in such an experiment, one should not at the same time modulate the concentrations of the other enzymes, as one would then compound the control exercised by those other enzymes with the control of the enzyme under study. The concentrations (activities) of the other enzymes should be kept constant; also, if the living system itself would try to change them; we consider the 'metabolic control coefficients' [69].

When we study regulation, however, it is not the experimenter who modulates the enzyme concentration. It is the living system itself that does this, for instance in response to an external challenge. But, typically, the system changes the concentrations of many enzymes at the same time. Therefore, unless the system only changes the expression level of one enzyme, the gene-expression regulation coefficient is not the inverse of the flux control coefficient.

We now consider a thought-up example, where a cell is challenged by something that does not affect the flux through a two-enzyme pathway with intermediate X. The cell responds by changing the activities of the two genes encoding the two enzymes by $d\text{lng}_1$ and $d\text{lng}_2$ respectively. Writing the control coefficients measuring the extent to which gene 1 controls the steady-state concentration of enzyme 1 as $C_{g_i}^{e_i}$, and indicating the traditional flux control coefficients by capital C, one finds:

$$\rho_g^{v_1} = \frac{C_{g_i}^{e_i} \cdot d \ln g_1}{C_1^J \cdot C_{g1}^{e_1} \cdot d \ln g_1 + C_2^J \cdot C_{g2}^{e_2} \cdot d \ln g_2}$$

This confirms that the regulation coefficient is not just equal to the inverse of the corresponding flux control coefficient, unless the enzyme is the sole rate-limiting step of the pathway, or the enzyme is the sole enzyme of the pathway that is regulated by gene expression.

This finding that the gene-expression regulation coefficients need not be equal to the inverse of the corresponding flux control coefficients is perhaps a bit counterintuitive. One might expect some relationship between control and regulation to be in place. For instance, the induction of steps without control would be unsuccessful, hence not contribute to regulation of the flux. Multiplying regulation coefficients with the corresponding flux control coefficients should have some meaning . . .

And indeed, it has. For a linear pathway:

$$\left(\sum_{i=1}^{n} C_i^{J_i} \cdot \rho_g\right)_{\text{linear pathway}} \equiv 1$$

This can be seen by changing indices 1 and 2 in the above equation and then evaluating the sum of the two products of regulation and control coefficients, or by considering the following proof:

$$\sum_{i=1}^{n} C_i^J \cdot \rho_g^{v_i} = \left(\frac{\partial \ln[J_i]}{\partial \ln(e_i)}\right)_{e_{j\neq i},\ \text{steady state}} \cdot \left(\frac{d \ln[e_i]}{d \ln(J_i)}\right)_{\text{steady state}}$$
$$= \left(\frac{d \ln(J_i)}{d \ln(J_i)}\right)_{\text{steady state}} = 1$$

Strategies of regulation

Above, we have shown that the regulation coefficient is not just equal to the inverse of the corresponding flux control coefficient, unless the enzyme is the sole rate-limiting step of the pathway, or the enzyme is the sole enzyme of the pathway that is regulated by gene expression. The two exceptions correspond to classical views on metabolic control and regulation. One of these applies where a pathway has a single rate-limiting enzyme. In that case, all other flux control coefficients equal zero, and if the cell were to change the expression level of any of the other enzymes, there would be no effect on the flux. The pathway flux can only change if the rate-limiting enzyme changes in expression level, hence the gene expression-regulation coefficient of the rate-limiting enzyme must equal 1.

In the second case, flux control may be distributed over the pathway enzymes, yet the cell may regulate only by overexpressing one of the enzymes. Then the gene-expression regulation coefficient of that enzyme equals the inverse of the flux control coefficient. When the flux should be increased, then it might be best for the cell to overexpress the enzyme with the highest flux control coefficient and only that enzyme, because then the required increase in enzyme concentration might be the smallest. These strategies of regulation could be called 'rate-limiting step regulation' and 'single step regulation' respectively.

Kacser and Acerenza [70] and Fell and Thomas [71] have identified a particularly robust way in which an engineer or an organism respectively could regulate a pathway flux. The way is robust because it does not involve changes in metabolite concentrations and therefore leaves cell behaviour as optimal as it was before the regulation took place. The

©The Authors Journal compilation ©2010 Biochemical Society

regulation is then such that the expression levels of all enzymes are increased by the same factor. The consequence in the above equation is that the two gene-expression regulation coefficients should both be equal to the inverse of the sum of the flux control coefficients, which is 1 [30]. Consequently, all metabolic regulation coefficients should equal zero.

The implications of these three regulation strategies for the gene-expression regulation coefficients are clear, hence subject to experimental validation or falsification. There exist still only few experimental regulation analyses. Surprisingly, in none of these do any of the three regulation strategies seem to be used by the living organism examined [44,69,72–76]. Regulation may serve many more purposes than adjusting flux, and this may account for the more complex regulatory patterns found [72].

Concluding remarks

Living systems have large numbers of components and astronomical numbers of potential molecular interactions. Without making use of what we know of the inherent organization of living organisms, it will be impossible to understand how all molecular interactions lead to function. But by putting in the dimensionality of the pathways in living cells, fairly simple extrapolations of intuition can be established that are useful to understand and dissect regulation and control in living organisms.

Funding

We thank the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council [grant numbers BBD0190791, BBC0082191, BBF0035281, BBF0035521, BBF0035521, BBF0035361, BBG5302251 and *SySMO P49*], European Union Framework Programme 7 [(BioSim, NucSys, EC-MOAN (*Escherichia coli* Modeling and Analysis)] and other funders (http://www.systembiology.net/support/) for support.

References

1 Lazebnik, Y. (2002) Can a biologist fix a radio? Or, what I learned while studying apoptosis. Cancer Cell **2**, 179–182
2 Higgs, P.W. (1964) Broken symmetries and the masses of gauge bosons. Phys. Rev. Lett. **13**, 508
3 MacPherson, K.H.R., Xu, Y.B., Cheah, E., Carr, P.D., van Heeswijk, W.C., Westerhoff, H.V., Luque, E., Vasudevan, S.G. and Ollis, D.L. (1998) Crystallization and preliminary X-ray analysis of *Escherichia coli* GlnK. Acta Crystallogr. Sect. D Biol. Crystallogr. **54**, 996–998
4 Monkhorst, K., de Hoon, B., Jonkers, I., Achame, E.M., Monkhorst, W., Hoogerbrugge, J., Rentmeester, E., Westerhoff, H.V., Grosveld, F., Grootegoed, J.A. and Gribnau, J. (2009) The probability to initiate X chromosome inactivation is determined by the X to autosomal ratio and X chromosome specific allelic properties. PLoS ONE **4**, e5616
5 Matthews, B.W. (2007) Protein structure initiative: getting into gear. Nat. Struct. Mol. Biol. **14**, 459–460
6 Novak, B., Csikasz-Nagy, A., Gyorffy, B., Chen, K. and Tyson, J.J. (1998) Mathematical model of the fission yeast cell cycle with checkpoint controls at the G_1/S, G_2/M and metaphase/anaphase transitions. Biophys. Chem. **72**, 185–200
7 Conradie, R., Bruggeman, F.J., Ciliberto, A., Csikasz-Nagy, A., Novak, B., Westerhoff, H.V. and Snoep, J.L. (2010) Restriction point control of the mammalian cell cycle via the cyclin E/Cdk2:p27 complex. FEBS J. **277**, 357–367
8 Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature **191**, 144–148
9 Westerhoff, H.V., Simonetti, A.L.M. and Van Dam, K. (1981) The hypothesis of localized chemiosmosis is unsatisfactory. Biochem. J. **200**, 193–202
10 Van Dam, K., Westerhoff, H.V., Krab, K., Van Der Meer, R. and Arents, J.C. (1980) Relationship between chemiosmotic flows and thermodynamic forces in oxidative phosphorylation. Biochim. Biophys. Acta **591**, 240–250
11 Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M. et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science **270**, 397–403
12 Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchison, C.A., Smith, H.O. and Venter, J.C. (2006) Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. U.S.A. **103**, 425–430
13 Dawkins, R. (1976) The Selfish Gene, Oxford University Press, Oxford
14 Noble, D. (2006) The Music of Life: Biology Beyond Genes, Oxford University Press, Oxford
15 Turing, A. (1952) The chemical basis of morphogenesis. Philos. Trans. R. Soc. London Ser. B **237**, 37–72
16 Gierer, A. and Meinhardt, H. (1972) Theory of biological pattern formation. Kybernetik **12**, 30–39
17 Boogerd, F.C., Bruggeman, F.J., Hofmeyr, J.H.S. and Westerhoff, H.V. (2007) Systems Biology: Philosophical Foundations, Elsevier Science, Amsterdam
18 Westerhoff, H.V., Winder, C., Messiha, H., Simeonidis, E., Adamczyk, M., Verma, M., Bruggeman, F.J. and Dunn, W. (2009) Systems biology: the elements and principles of life. FEBS Lett. **583**, 3882–3890
19 Westerhoff, H.V. and Welch, G.R. (1992) Enzyme organization and the direction of metabolic flow: physico-chemical considerations. Curr. Top. Cell. Regul. **33**, 361–390
20 Olah, J., Klivenyi, P., Gardian, G., Vecsei, L., Orosz, F., Kovacs, G.G., Westerhoff, H.V. and Ovadi, J. (2008) Increased glucose metabolism and ATP level in brain tissue of Huntington's disease transgenic mice. FEBS J. **275**, 4740–4755
21 Westerhoff, H.V., Hellingwerf, K.J. and Van Dam, K. (1983) Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. Proc. Natl. Acad. Sci. U.S.A. **80**, 305–309
22 Van Der Meer, R., Westerhoff, H.V. and Van Dam, K. (1980) Linear relation between rate and thermodynamic force in enzyme-catalyzed reactions. Biochim. Biophys. Acta **591**, 488–493
23 Westerhoff, H.V. and Van Dam, K. (1987) Thermodynamics and Control of Biological Free-Energy Transduction, Elsevier, Amsterdam
24 Onsager, L. (1931) Reciprocal relations in irreversible processes. I. Phys. Rev. **37**, 405–426
25 Keizer, J. (1987) Statistical Thermodynamics of Nonequilibrium Processes, Springer-Verlag, Berlin
26 Alberghina, L. and Westerhoff, H.V. (2005) Systems Biology: Definitions and Perspectives, Springer, Berlin
27 Jensen, P.R., Van Der Gugten, A.A., Bier, M., Van Heeswijk, W.C., Rohwer, J., Molenaar, D., Van Workum, M., Richard, P., Teusink, B., Bakker, B.N. et al. (1995) Hierarchies in control. J. Biol. Syst. **3**, 139–144
28 Westerhoff, H.V., Koster, J.G., Van Workum, M. and Rudd, K.E. (1990) On the control of gene expression. in Control of Metabolic Processes (Cornish-Bowden, A. and Cárdenas, M.L., eds), pp. 399–412, Plenum Publishing, New York
29 Hofmeyr, J.H., Kacser, H. and van der Merwe, K.J. (1986) Metabolic Control Analysis of moiety-conserved cycles. Eur. J. Biochem. **155**, 631–641
30 Kacser, H. and Burns, J.A. (1973) The control of flux. Symp. Soc. Exp. Biol. **27**, 65–104
31 Rapoport, T.A., Heinrich, R., Jacobasc, G. and Rapoport, S. (1974) Linear steady-state treatment of enzymatic chains: mathematical-model of glycolysis of human erythrocytes. Eur. J. Biochem. **42**, 107–120
32 Kahn, D. and Westerhoff, H.V. (1991) Control theory of regulatory cascades. J. Theor. Biol. **153**, 255–285

©The Authors Journal compilation ©2010 Biochemical Society

33 Westerhoff, H.V. (2008) Signalling control strength. J. Theor. Biol. **252**, 555–567
34 Fell, D.A. and Small, J.R. (1986) Fat synthesis in adipose tissue: an examination of stoichiometric constraints. Biochem. J. **238**, 781–786
35 Papin, J.A., Price, N.D., Edwards, J.S. and Palsson, B.O. (2002) The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. J. Theor. Biol. **215**, 67–82
36 Poolman, M.G., Fell, D.A. and Raines, C.A. (2003) Elementary modes analysis of photosynthate metabolism in the chloroplast stroma. Eur. J. Biochem. **270**, 430–439
37 Dandekar, T., Schuster, S., Snel, B., Huynen, M. and Bork, P. (1999) Pathway alignment: application to the comparative analysis of glycolytic enzymes. Biochem. J. **343**, 115–124
38 Wiechert, W. and de Graaf, A.A. (1996) *In vivo* stationary flux analysis by ^{13}C labeling experiments. Adv. Biochem Eng Biotechnol. **54**, 109–154
39 Reed, J.L. and Palsson, B.O. (2003) Thirteen years of building constraint-based *in silico* models of *Escherichia coli*. J. Bacteriol. **185**, 2692–2699
40 Westerhoff, H.V., Kolodkin, A., Conradie, R., Wilkinson, S.J., Bruggeman, F.J., Krab, K., van Schuppen, J.H., Hardin, H., Bakker, B.M., Mone, M.J. et al. (2009) Systems biology towards life *in silico*: mathematics of the control of living cells. J. Math. Biol. **58**, 7–34
41 Molenaar, D., van Berlo, R., de Ridder, D. and Teusink, B. (2009) Shifts in growth strategies reflect tradeoffs in cellular economics. Mol. Syst. Biol. **5**, 323
42 Simeonidis, E., Murabito, E., Smallbone, K. and Westerhoff, H.V. (2010) Why does yeast ferment? A flux balance analysis study. Biochem. Soc. Trans. **38**, 1225–1229
43 Sauro, H.M. (1990) Quantification of metabolic regulation by effectors. in Control of Metabolic Processes (Cornish-Bowden, A. and Cárdenas, M.L., eds), pp. 225–230, Plenum Press, New York
44 ter Kuile, B.H. and Westerhoff, H.V. (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. FEBS Lett. **500**, 169–171
45 Bruggeman, F.J., de Haan, J., Hardin, H., Bouwman, J., Rossell, S., van Eunen, K., Bakker, B.M. and Westerhoff, H.V. (2006) Time-dependent hierarchical regulation analysis: deciphering cellular adaptation. IEE Proc. Syst. Biol. **153**, 318–322
46 Bevilacqua, A., Wilkinson, S.J., Dimelow, R., Murabito, E., Rehman, S., Nardelli, M., van Eunen, K., Rossell, S., Bruggeman, F.J., Blüthgen, N. et al. (2008) Vertical systems biology: from DNA to flux and back. SEB Exp. Biol. Ser. **61**, 65–91
47 van Eunen, K., Bouwman, J., Lindenbergh, A., Westerhoff, H.V. and Bakker, B.M. (2009) Time-dependent regulation analysis dissects shifts between metabolic and gene-expression regulation during nitrogen starvation in baker's yeast. FEBS J. **276**, 5521–5536
48 Newsholme, E.A. (1982) Application of principles of metabolic control to some aspects of hormone action. Prog. Clin. Biol. Res. **102**, 269–277
49 Kacser, H. and Burns, J.A. (1973) The control of flux. Symp. Soc. Exp. Biol. **27**, 65–104
50 Heinrich, R. and Rapoport, T.A. (1974) A linear steady-state treatment of enzymatic chains: general properties, control and effector strength. Eur. J. Biochem. **42**, 89–95
51 Westerhoff, H.V. and Chen, Y.D. (1984) How do enzyme activities control metabolite concentrations? An additional theorem in the theory of metabolic control. Eur. J. Biochem. **142**, 425–430
52 Westerhoff, H.V. and Kell, D.B. (1987) Matrix-method for determining steps most rate-limiting to metabolic fluxes in biotechnological processes. Biotechnol. Bioeng. **30**, 101–107
53 Reder, C. (1988) Metabolic control theory: a structural approach. J. Theor. Biol. **135**, 175–201
54 Kahn, D. and Westerhoff, H.V. (1991) Control theory of regulatory cascades. J. Theor. Biol. **153**, 255–285
55 Flint, H.J., Tateson, R.W., Barthelmess, I.B., Porteous, D.J., Donachie, W.D. and Kacser, H. (1981) Control of the flux in the arginine pathway of *Neurospora crassa*: modulations of enzyme activity and concentration. Biochem. J. **200**, 231–246
56 Groen, A.K., Wanders, R.J., Westerhoff, H.V., van der Meer, R. and Tager, J.M. (1982) Quantification of the contribution of various steps to the control of mitochondrial respiration. J. Biol. Chem. **257**, 2754–2757
57 Westerhoff, H.V., Plomp, P., Groen, A.K., Wanders, R.J. A., Bode, J.A. and van Dam, K. (1987) On the origin of the limited control of mitochondrial respiration by the adenine nucleotide translocator. Arch. Biochem. Biophys. **257**, 154–169
58 Stuger, R., Woldringh, C.L., Van Der Weijden, C.C., Vischer, N.O.E., Bakker, B.M., van Spanning, R.J.M., Snoep, J.L. and Westerhoff, H.V. (2002) DNA supercoiling by gyrase is linked to nucleoid compaction. Mol. Biol. Rep. **29**, 79–82
59 Hornberg, J.J., Bruggeman, F.J., Binder, B., Geest, C.R., de Vaate, A., Lankelma, J., Heinrich, R. and Westerhoff, H.V. (2005) Principles behind the multifarious control of signal transduction: ERK phosphorylation and kinase/phosphatase control. FEBS J. **272**, 244–258
60 Kholodenko, B.N., Molenaar, D., Schuster, S., Heinrich, R. and Westerhoff, H.V. (1995) Defining control coefficients in nonideal metabolic pathways. Biophys. Chem. **56**, 215–226
61 Burns, J.A., Cornish-Bowden, A., Groen, A.K., Heinrich, R., Kacser, H., Porteous, J.W., Rapoport, S.M., Rapoport, T.A., Stucki, J.W., Tager, J.M. et al. (1985) Control analysis of metabolic systems. Trends Biochem. Sci. **10**, 16–16
62 Jensen, P.R., Michelsen, O. and Westerhoff, H.V. (1993) Control analysis of the dependence of *Escherichia coli* physiology on the H^+-ATPase. Proc. Natl. Acad. Sci. U.S.A. **90**, 8068–8072
63 Bakker, B.M., Michels, P.A. M., Opperdoes, F.R. and Westerhoff, H.V. (1999) What controls glycolysis in bloodstream form *Trypanosoma brucei*? J. Biol. Chem. **274**, 14551–14559
64 Chassagnole, C., Fell, D.A., Rais, B., Kudla, B. and Mazat, J.P. (2001) Control of the threonine-synthesis pathway in *Escherichia coli*: a theoretical and experimental approach. Biochem. J. **356**, 433–444
65 Rossell, S., Van Der Weijden, C.C., Lindenbergh, A., van Tuijl, A., Francke, C., Bakker, B.M. and Westerhoff, H.V. (2006) Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A. **103**, 2166–2171
66 Bruggeman, F.J., Boogerd, F.C. and Westerhoff, H.V. (2005) The multifarious short-term regulation of ammonium assimilation of *Escherichia coli*: dissection using an *in silico* replica. FEBS J. **272**, 1965–1985
67 Cornish-Bowden, A. (1995) Fundamentals of Enzyme Kinetics, Portland Press, London
68 Daran-Lapujade, P., Rossell, S., van Gulik, W.M., Luttik, M.A.H., de Groot, M.J.L., Slijper, M., Heck, A.J.R., Daran, J.M., de Winde, J.H., Westerhoff, H.V. et al. (2007) The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proc. Natl. Acad. Sci. U.S.A. **104**, 15753–15758
69 van der Gugten, A.A. and Westerhoff, H.V. (1997) Internal regulation of a modular system: the different faces of internal control. Biosystems **44**, 79–106
70 Kacser, H. and Acerenza, L. (1993) A universal method for achieving increases in metabolite production. Eur. J. Biochem. **216**, 361–367
71 Fell, D.A. and Thomas, S. (1995) Physiological control of metabolic flux: the requirement for multisite modulation. Biochem. J. **311**, 35–39
72 Rossell, S., Van Der Weijden, C.C., Kruckeberg, A.L., Bakker, B.M. and Westerhoff, H.V. (2005) Hierarchical and metabolic regulation of glucose influx in starved *Saccharomyces cerevisiae*. FEMS Yeast Res. **5**, 611–619
73 Cakir, T., Kirdar, B., Onsan, Z.I., Ulgen, K.O. and Nielsen, J. (2007) Effect of carbon source perturbations on transcriptional regulation of metabolic fluxes in *Saccharomyces cerevisiae*. BMC Syst. Biol. **1**, 18
74 Carrari, F., Baxter, C., Usadel, B., Urbanczyk-Wochniak, E., Zanor, M.I., Nunes-Nesi, A., Nikiforova, V., Centero, D., Ratzka, A., Pauly, M. et al. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. Plant Physiol. **142**, 1380–1396
75 Cakir, T., Patil, K.R., Onsan, Z., Ulgen, K.O., Kirdar, B. and Nielsen, J. (2006) Integration of metabolome data with metabolic networks reveals reporter reactions. Mol. Syst. Biol. **2**, 50
76 Even, S., Lindley, N.D. and Cocaign-Bousquet, M. (2003) Transcriptional, translational and metabolic regulation of glycolysis in *Lactococcus lactis* subsp. *cremoris* MG 1363 grown in continuous acidic cultures. Microbiology **149**, 1935–1944

Received 9 March 2010
doi:10.1042/BST0381189

©The Authors Journal compilation ©2010 Biochemical Society