Systems biology may work when we learn to understand the parts in terms of the whole

A. Cornish-Bowden[1] and M.L. Cárdenas
CNRS UPR-9036, Institut Fédératif "Biologie Structurale et Microbiologie", 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France

Abstract

The first point to note about whether systems biology will work is that the essential idea of systems biology is not new: there has been interest in it, as well as efforts to apply it, since the middle of the 20th century. The difference now is that it has become fashionable, with an explosion in the number of publications using the two words, albeit not always with the same meaning. The reductionist approach remains dominant, however, and systems biology is often seen as no more than integration of diverse data into models of systems. This way of thinking needs to be changed if systems biology is to lead to an understanding of life and to provide the benefits that are expected from it. The emphasis ought to be on the needs of the system as a whole for understanding the components, not the converse. General properties of metabolic systems, such as feed-back inhibition, can be properly understood by taking account of supply and demand, i.e. the requirements of the system as a whole, but this is often overlooked. Metabolism tends to be viewed as static, although enzymes (and proteins in general) are continuously synthesized and degraded. The fact that they are themselves therefore metabolites introduces great complexity to metabolism, including an implication of infinite regress; understanding how living organisms escape from this will be an essential step towards understanding life.

Introduction

According to Bains [1], 'the genome has turned out to be a relatively poor source of explanation for the differences between cells or between people'. This statement may appear too negative, but there can be little doubt that the present capacity to generate mountains of new data has far outstripped our capacity to make sense of it all. Before many genomes had been elucidated it was widely assumed that the function of an unidentified gene could be determined by examining the effects of deleting it. However, many genes prove to be 'silent': an organism remains viable when such a gene is deleted, sometimes with normal growth and normal metabolic fluxes. Clearly the solution here is not simply to gather more observations, but to search for a better understanding of the observations already available; an important step is to recognize that metabolite concentrations are far more sensitive than fluxes to perturbations. Studying effects of gene deletion on concentrations, preferably combined effects on multiple concentrations, therefore provides a much more sensitive probe into gene function than studying effects on fluxes [2,3]. All of this implies a need for a biology of systems that studies an organism as an entity and not just as a collection of components.

Systemic ideas in biology

After a long period during which systemic ideas excited very little interest among biochemists, 'systems biology' has suddenly come into vogue, so suddenly that more than half of the publications found by searching PubMed for this term at the beginning of January 2005 date from the single year of 2004; even 2005 accounts for more than all the years before 1998 added together. Does this reflect a real resurgence of interest in systems theory as this would have been understood by Bertalanffy [4] or Rosen [5], for example, or even in the systemic ideas [6] that form the basis of metabolic control analysis? Probably not, even though a small part of the current research output is genuinely system-oriented.

Long ago Bertalanffy [4] complained that 'the only goal of science appears to be analytical, i.e. the splitting up of reality into ever smaller units'. It is unlikely that he would be enthusiastic about molecular biology as currently practised, with or without the label of systems biology, recently characterized [7] as the integration of 'knowledge from diverse biological components and data into models of the system as a whole.' This sort of definition is entirely reductionist, and makes 'systems biology' into little more than a euphemism for the type of approach that Bertalanffy criticized: instead of using a view of the whole system as a way to understand its components, it seeks to explain the whole in terms of a vast list of components.

Reductionism

The reductionist approach to biochemistry started about a century ago with Buchner's demonstration that the capacity of yeast cells to convert glucose into ethanol did not need an explanation in terms of a vital force [8]. This discovery sounded the death knell of vitalism in biology, and the development of biochemistry that followed was driven by reductionism, with cells separated into their components,

Key words: infinite regress, metabolism, systems biology.
[1]To whom correspondence should be addressed (email acornish@ibsm.cnrs-mrs.fr).

which were then separated into smaller components, and then studied in isolation. The reductionist stage was certainly necessary, and 20th century biochemistry could not have achieved the successes that it did if components had never been studied one at a time. The time has come, however, to move beyond this, beyond even studying the interactions of the components with one another, because all of them form parts of a whole, and their presence in the whole can only be understood by considering the needs of the whole.

As an example, many cases of cooperative feedback inhibition of metabolic pathways are known, such as the inhibition of aspartokinase in bacteria by lysine. This type of observation is often explained by supposing that the biosynthetic flux is regulated by this feedback inhibition, and would be subject to uncontrolled variations if there were no feedback loop. However, this explanation is wrong, because fluxes can be controlled perfectly well without feedback inhibition, whether cooperative or not [9]. The need comes not from flux control but concentration control: without feedback inhibition in this pathway the rate at which lysine would be synthesized would still match the rate at which it is used in protein synthesis, but there would be huge and potentially damaging variations in the concentrations of lysine and the intermediates in the pathway from aspartate. This sensitivity of metabolite concentrations to perturbations has major implications for the regulatory design of metabolism in living organisms. To understand this it is necessary to represent biosynthetic pathways in a way that allows analysis in terms of supply and demand [10]; that is to say, in a more complete way than is usual in textbooks of biochemistry. These typically show, for example, the biosynthesis of lysine as a series of reactions that begin with aspartate and end with lysine. However, lysine is not in any meaningful sense the end-product: it is made not as an end in itself but as a starting material for other processes, principally, in this case, protein synthesis. As protein synthesis accounts for most of the metabolic demand for lysine, it determines the rate at which it needs to be synthesized from aspartate. Omitting the conversion of lysine into protein from the pathway means omitting the one step that explains the feedback inhibition of aspartokinase by lysine. This inhibition cannot be explained solely in terms of the components concerned, aspartokinase and lysine, but requires consideration of the whole system, including protein synthesis.

The case of phosphofructokinase is striking: despite having an activity that is modulated by several effectors, it has very little control over the glycolytic flux: all the attempts made by different research groups to increase glycolytic flux in diverse organisms by genetically increasing its activity have failed (e.g. see [11]). This appears surprising if the enzyme is considered in isolation, but it makes sense as a property of the whole system: as for the biosynthesis of lysine just discussed, the glycolytic flux is not determined by phosphofructokinase, but by the metabolic use that is made of the products of glycolysis. The regulatory properties of phosphofructokinase allow changes in flux to occur with minimal changes in metabolite concentration.

A final example of the need to analyse systems as systems, and not as mere collections of parts, is provided by studies of the effects of combining deletions of each of nine galactose-utilization genes with two dietary states, with and without galactose [12]. Comparison of the effects of 20 perturbations on cellular gene expression showed that the levels of 997 mRNAs varied significantly with one or more of these perturbations. The results were thus far from the naive expectation that altering a gene should affect expression of just that gene alone. Integrating mRNA and protein expression responses with the global set of protein–protein and protein–DNA interactions allowed a network involving hundreds of genes to be deduced. This generalized pleiotropy, predicted on the basis of systemic arguments many years ago by Kacser [13], underlines that genes act in concert with one another and with the environment. The more complex the level at which one seeks to explain a living system, the greater the need to examine the network of interactions that lie behind the genome [14].

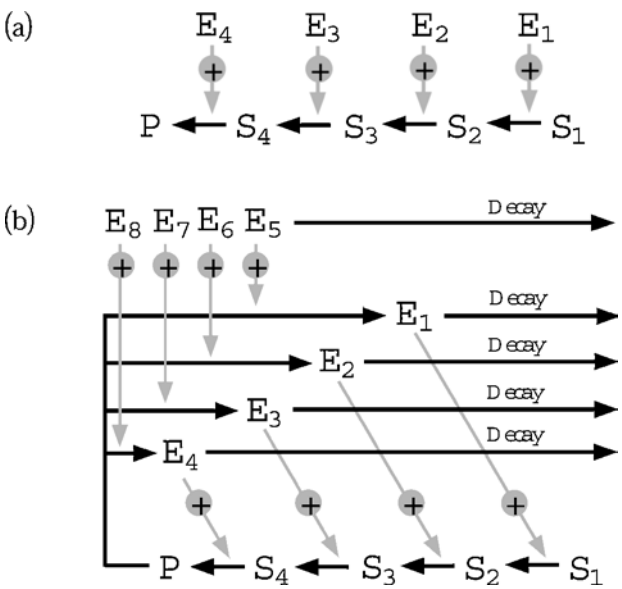
Definition of life

It is not clear whether systems biology as it is currently understood will lead us any closer to an understanding of life than we were when Schrödinger [15] raised the question 60 years ago, but without such understanding systems biology will remain little more than tinkering. It may sometimes 'work', certainly, just as evolution by natural selection applies mindless tinkering to create complex organisms that 'work'. The hope of systems biology should surely be more than this, however, that with a proper understanding of why living systems are organized in the way that they are we shall be able to design fundamentally new solutions to problems of technological interest.

Understanding the organization of living organisms will only be possible with a genuinely systemic approach, and long after Schrödinger's lectures in Dublin in 1944, we are still far from having a satisfactory definition of life. All living organisms both make themselves and repair any damage to themselves that results from ordinary wear and tear, but no machine exists or can easily be conceived that even approaches this level of complexity: the best we can do with present technology is to produce sophisticated machines that include a very limited amount of self-testing to warn their operators of faulty components, but the actual repair or replacement of such components relies on agencies external to the machine itself. In an effort to understand self-repair in biological organisms, Rosen [5,16,17] devoted the whole of his career to the development of what he called metabolism-repair systems, or (M,R)-systems.

The classical view of metabolism, illustrated in Figure 1(a), is static: a series of chemical reactions is catalysed by a series of fixed enzymes, and the product of the reactions is not used for anything in particular. This picture lacks the essential idea that the catalysts responsible for metabolism are not permanent entities, but are subject to inevitable wear and tear resulting in degradation and loss of activity. They must thus

Figure 1 | A schematic representation of metabolism (a) and a more realistic (but still highly simplified) view of metabolism (b)
(**a**) Metabolism is commonly regarded as a set of chemical reactions that convert starting materials, here represented by the single reactant S_1, into products, here represented by P. (**b**) The enzymes that catalyse metabolic reactions are themselves products of the same metabolism, and ought to be regarded as metabolites. Moreover, they are never indefinitely stable, and need to be continually resynthesized. This resynthesis itself requires catalysts, themselves normally enzymes, and thus likewise not indefinitely stable, and equally in need of resynthesis. All of this implies an infinite regress, and understanding how to escape from it is an essential step towards understanding life.



be considered as products of metabolism themselves, and there must be other enzymes to catalyse their synthesis. This additional complexity is hinted at in Figure 1(b), but this barely touches on the essential problem, that every level of components that one adds in order to solve the most obvious problems just brings a new set of problems of the same kind, because the protein-synthesizing enzymes are themselves subject to decay, as are the enzymes that catalyse protein degradation. In the 1970s, investigators of protein degradation in mammals were concerned that the specificity of protein degradation implied by the very different half-lives of different enzymes would require specific proteinases, leading to infinite regress. In the case of protein synthesis, it may appear misleading in Figure 1(b) to write four separate catalysts for the synthesis of four enzymes, given that a ribosome may catalyse the synthesis of many different proteins. However, as the complete catalyst includes a specific mRNA as well as the ribosome it is different for each protein synthesized.

Almost inevitably, therefore, one arrives at an infinite regress, without getting any closer to understanding how organisms, which are neither infinitely large nor infinitely complex, succeed in escaping it. Notice that these difficulties are absolutely fundamental in considering metabolism as the basis of life, before we even start to take account of such complications as compartmentation, transport across membranes, enzyme–enzyme association, and so on. It should also be evident that understanding will not come from accumulating more and more detailed knowledge, but from developing a theory of the living organism as a whole. This theory will need to explain how a living organism of finite size and complexity can maintain its identity, and hence its internal organization, for periods that are very long compared with the time constants of the chemical reactions within it.

In mathematical terms, metabolism can be seen as a mapping f that transforms one metabolic configuration into another, and Rosen's solution to escape the infinite regress is to suppose that the mapping is invertible – a property that the overwhelming majority of mappings do not have. In a sense, f satisfies the bizarre equation $f(f) = f$, i.e. it is a function that acts on an instance of itself to produce another instance of itself. Apart from extremely trivial functions such as addition of zero and multiplication by unity, ordinary mathematical functions do not behave like this, and it has been doubted [18] whether non-trivial mathematical examples are even possible; we have dealt with this point elsewhere [19]. At a moment when new discoveries in virology [20] are reopening the question of what it means to be alive, the definition of life is more relevant than ever, and the progress since Schrödinger [15] has been depressingly scanty.

Conclusions

We do not believe that systems biology will deliver the benefits expected of it as long as it remains just a new term for large-scale classical biological experimentation. Real progress will require a serious effort to understand systems at a systemic level. With this in mind we have started on an effort to express Rosen's analysis [5,16,17] in more accessible language [19,21], but much remains to be done.

References

1 Bains, W. (2001) Nat. Biotechnol. **19**, 401–402
2 Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J. et al. (2001) Nat. Biotechnol. **19**, 45–50
3 Cornish-Bowden, A. and Cárdenas, M.L. (2001) Nature (London) **409**, 571–572
4 Bertalanffy, L. (1975) General Systems Theory in Human Communication (Ruben, B.D. and Kim, J.Y., eds.), pp. 6–20, Hayden Books, Rochelle Park, New Jersey
5 Rosen, R. (1991) Life Itself. Columbia University Press, New York
6 Kacser, H. (1957) The Strategy of the Genes (Waddington, C.H., ed.), pp. 191–249, George Allen and Unwin, London
7 Ideker, T. (2004) Nat. Biotechnol. **22**, 473–475
8 Buchner, E. (1897) Ber. Dt. Chem. Ges. **30**, 117–124; English translation by H.C. Friedmann, pp. 25–31, in New Beer in an Old Bottle: Eduard Buchner and the Growth of Biochemical Knowledge (Cornish-Bowden, A., ed.), Universitat de Valencia, Valencia
9 Cornish-Bowden, A., Hofmeyr, J.-H.S. and Cárdenas, M.L. (1995) Bioorg. Chem. **23**, 439–449
10 Hofmeyr, J.-H. S. and Cornish-Bowden, A. (2000) FEBS Lett. **476**, 47–51
11 Schaaff, I., Heinisch, J. and Zimmermann, F.K. (1989) Yeast **5**, 285–290
12 Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R. and Hood, L. (2001) Science **292**, 929–934

13 Kacser, H. (1963) Biological Organization at the Cellular and Supercellular Level (Harris, R.J.C., ed.), pp. 25–41, Academic Press, New York
14 Cornish-Bowden, A. and Cárdenas, M.L. (2001) Trends Biochem. Sci. **26**, 463–465
15 Schrödinger, E. (1944) What is Life? Cambridge University Press, Cambridge, Chapter 7
16 Rosen, R. (1958) Bull. Math. Biophys. **20**, 245–341
17 Rosen, R. (2000) Essays on Life Itself. Columbia University Press, New York
18 Landauer, C. and Belmann, K. (2002) AIP Conf. Proc. **627**, 59–70
19 Letelier, J.C., Soto-Andrade, J., Guíñez Abarzúa, F., Cornish-Bowden, A. and Cárdenas, M.L. (2004) Artificial Life IX (Pollack, J., Bedau, M., Husbands, P., Ikegami, T. and Watson, R.A., eds.), pp. 450–455, MIT Press, Cambridge, Massachusetts
20 Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M. and Claverie, J.M. (2004) Science **306**, 1344–1350
21 Cornish-Bowden, A., Cárdenas, M.L., Letelier, J.C., Soto-Andrade, J. and Guíñez Abarzúa, F. (2004) Biol. Cell **96**, 713–717

Received 18 January 2005