Abstract

Although the rapid advance of the various “omics” fields is striking, it consists mainly of accumulating data in ever-increasing detail at an ever-increasing rate. There is little basis for believing that this approach is leading to a deeper understanding of living organisms, and the hope of creating life de novo remains as remote as it has always been. Changing this will require a far greater integration of systemic ways of thinking into what is loosely called systems biology. At the most profound level this means incorporating a genuine theory of life, but even at a more superficial level the relatively simple ideas that have come from metabolic control analysis have yet to be fully incorporated into biotechnological practice.

Introduction

Proteomics, metabolomics, metabonomics, transcriptomics and so on, together with systems biology as it is currently understood, are essentially products of the 21st century, as one can see from the distribution of usages of the various terms listed in Table 1, which is based on data from PubMed (2007). The explosive increase in papers about genomics is, of course, a direct consequence of the vast amount of information that has come from sequencing genomes, and is a natural development of the subject. The other “omics” disciplines are fruits of the growing amount of genomic information, and studies of the proteome and transcriptome have already contributed greatly to understanding of the net of interactions that connect genes to phenotypes (Cornish-Bowden and Cárdenas, 2001a), and they have been useful tools for early diagnosis of medical problems. For example, powerful proteomic technologies now have great potential in cancer research for biomarker discovery, and for addressing the issue of cancer heterogeneity. Detecting cancer at an early stage, and predicting how a tumour will develop and how it will respond to therapy, are areas of research that are already benefitting from proteomics (Celis et al., 2004; 2006).

Systems biology, however, is more of a new name for an old approach than a genuinely new way of studying biology. The distinction is important because, as we shall argue in this paper, there is a real need for an integrated approach to biology in which the
components of a biological system are analysed in terms of their contributions to the organization of the whole system, but it is far from clear that that is what systems biology is in its current form. On the contrary, it appears to be just as reductionist as less fashionable areas of biochemistry and molecular biology, differing mainly in being based on an enormously increased body of detailed data available for study.

Expressing the same idea differently, the current obsession with the accumulation of detailed data is not leading towards a better understanding of organisms: a theory of biological organization will not appear spontaneously from beneath a mountain of data, but will need to be actively constructed. Systems biology in its present form has almost nothing in common with the general systems theory that Ludwig von Bertalanffy worked to develop, and would not escape his criticism that “the only goal of science appears to be analytical, i.e. the splitting up of reality into ever smaller units and the isolation of individual causal trains” (Bertalanffy, 1975). The question therefore arises of what sort of systemic ideas need to be added to the various “omics” fields to enable them to move away from the mere accumulation of data and towards a real contribution to biological understanding. In a sense one could hope to move biology away from being a purely descriptive science to become a predictive science.

Eighty years after Heisenberg’s uncertainty principle taught physicists that behaviour at the particle level cannot be predicted, and 25 years after studies of chaotic dynamics taught them that the long-term behaviour of many-particle ensembles cannot be predicted either, it may seem futile to try to make biology something that even modern physics no longer claims to be, but the implied criticism misses the point. The existence of two areas of physics that are now known to be less amenable to prediction than they were once hoped to be hardly alters the fact that there is a vast body of theory underlying physics that allows the results of many experiments to be predicted. The theory of biology is far more restricted, being essentially limited to the theory of natural selection. This offers a very convincing mechanism for evolution and makes some predictions about the characteristics of a previously unknown species, but has essentially nothing to say about the origin of life, the moment when organized systems learned how to maintain their organization, in other words how to stay alive.

Biology does of course depend on many fragments of theory, such as the understanding of enzyme mechanisms that comes from studies of organic reaction mechanisms, or, most notably, the whole area of energy management known as bioenergetics, which depends on the laws of thermodynamics. However, none of these can be regarded as theories of biology at the same level as the theory of natural selection, because they apply to non-living systems no less (or more) than they do to living systems. Indeed, a major part of our understanding of biochemistry came from the overthrow of vitalism by Buchner (1897)—the recognition that living systems obey the same laws of chemistry and physics as non-living systems. Volcanic activity depends on the laws of thermodynamics just as much as a living organism does, but one would not call thermodynamics a theory of volcanoes. The distinction we are making here (made, of course, by Schrödinger (1944) before us) is that although no one now doubts that adherence to the laws of physics is necessary for life, it is much less clear that the currently known laws are sufficient.
Does biology need a theory of life?

The question of how far systemic ideas have influenced systems biology as it is currently understood can be answered at a fairly simple level, in terms of the ideas of metabolic control analysis developed from the seminal contributions of Kacser and Burns (1973) and of Heinrich and Rapoport (1974), or at a much more profound (and difficult) level, in terms of the theory of biological organization developed by Rosen (1991). At the simpler level this influence already exists: metabolic control analysis already forms a significant part of systems biology, though not as large a part as it probably should. At the more profound level it is probably fair to say that Rosen’s ideas have had no impact at all on ordinary practice. So far as most biologists are concerned there is no theory of the whole organism, and for most the lack of one has no importance, as they would agree with Medawar (1977) that discussing the nature of life represents “a low level in biological conversation”.

Medawar’s comment may have had some validity when he made it, as it could be argued that in the absence of the “omics” technology that we now have there was little that a theory of life or of the whole organism could have contributed, but that is no longer true, and to advance significantly further (other than in the accumulation of yet more detailed information) biology will need to integrate the information that already exists into a whole. It needs what Woese (2004) has called a guiding vision, because “without an adequate technological advance the pathway of progress is blocked, and without an adequate guiding vision there is no pathway, there is no way ahead.”

A pessimistic view would liken systems biology to cybernetics: in the middle of the 20th century this was confidently predicted to offer solutions to all problems of organization and regulation. However, apart from giving biochemists the idea of feed-back inhibition, it has largely vanished from biological consciousness, after failing to deliver on its early promise. What, then, ought the guiding vision of systems biology to be? In the deepest sense, Rosen’s (M,R) systems may provide this (Cornish-Bowden et al., 2007), but it will be a long time before these have any practical application, except in the negative sense that recognizing that some current objectives are impossible to realize may avoid some futile effort. In the shorter term, the less profound systemic ideas involved in metabolic control analysis are already applicable to current biotechnology, and may offer easier and better ways to success than the brute-force approach that has dominated the field since genetic manipulation became possible. Analysis of metabolic pathways and networks has a great potential in biotechnology and medicine, and constitutes a powerful tool in drug research (Eisenthal and Cornish-Bowden, 1998; Ramos-Montoya et al., 2006).

Machines and organisms

Rosen’s theory of (M,R) systems (Rosen, 1991) treats the fundamental properties of living organisms as metabolism and repair, though replacement expresses better than repair the intended meaning (Letelier et al., 2006). It is natural to suppose that “repair” includes the fundamental biological idea of reproduction, but this is not the case, because Rosen was little concerned with reproduction in the usual sense or with the
other central idea of modern biology, evolution. For most biologists these will seem to be such crucial omissions that they deprive Rosen’s theory of any interest it might have. However, the point is that Rosen was interested in life at a more fundamental level: until the early organisms had succeeded in staying alive, i.e. in maintaining their organization for a significant period, there was no question of either reproduction or evolution. It follows, therefore, that understanding how organisms stay alive is more fundamental than understanding how they reproduce or evolve.

This essential property that allows an organism to stay alive is metabolic closure, which allows them to preserve their integrity of organization and to be autonomous. We shall discuss later what this implies, but we note at the outset that no machine has any property equivalent to metabolic closure, and the fallacy of the machine metaphor for organisms defeats most attempts to understand them in their entirety. It may have been philosophically tenable for Descartes to hold that organisms are essentially machines, but it is not tenable today. There is a vast gulf between organisms and machines, and at the moment we cannot see how it might be bridged, even in principle, let alone in practice. A sufficiently detailed study of a machine may allow a competent engineer to produce another machine with the same functionality, but we are totally incapable of designing an organism today, and it may even be naive to think that it will ever be possible to design an organism \textit{ab initio}. Everything that we know about organisms confirms Rosen’s contention that an organism is not a machine.

{}\begin{center}
\includegraphics[width=\textwidth]{figure1.png}
\end{center}

\textit{Figure 1}. A schematic representation of metabolism and replacement. The conventional idea of metabolism is as a set of chemical reactions, represented here by the steps from $S_1$ to $P$, catalysed by a series of specific enzymes, $E_1$ to $E_4$. However, these enzymes are not supplied from outside and are not indefinitely stable (as represented by the arrows labelled Decay). Accordingly they need to be synthesized (“replaced”) by chemical reactions that use products of metabolism as starting materials.

In January 2007 the popular comic strip \textit{Dilbert} contained a conversation in which the question “Your sales representative told us that the product heals itself. Is that true?” received the answer “It’s totally true … that he said that.” Why was this amusing? It would not have been amusing if the conversation had been placed at an agricultural fair and the product had been a disease-resistant breed of pig; it would then have been an uncontroversial claim, because we all know that animals are capable of recovering from injuries and illnesses without external intervention. No, it was amusing because the cartoonist knew that his readers know perfectly well that machines cannot recover from
damage without external help. We know this, but we are tempted to ignore it in over-optimistic projections of where current biological engineering will lead.

![Diagram of metabolism and replacement](image-url)

*Figure 2. A more abstract representation of metabolism and replacement. All of the chemical reactions of metabolism are represented by the single arrow from A to B; thus A (reactants) and B (products) must be regarded as sets, not as individual metabolites. Catalysis by the set of enzymes needed for the metabolism is represented by the dashed arrow from f, representing the set of metabolic enzymes, acting on A. Replacement is represented by the arrow from B to f, and is catalysed by a replacement system (another set of enzymes) Φ. The diagram raises (but does not answer), the question of how metabolic closure is achieved, i.e. of how Φ is replaced. Discussion of how the question might be answered may be found elsewhere (Letelier et al., 2006; Cornish-Bowden et al., 2007).*

Philosophically the difference between organisms and machines lies in the different kinds of causation described by Aristotle. Machines and organisms both are open to material causation, as both are constructed from external materials, and both release used materials into their environments. They differ, however, in final causation, because all machines are made for a purpose, to fulfil particular functions, whereas organisms have no final causes. More important, they also differ in efficient causation, because in making a machine the essential decisions about which parts are installed in which locations are external, whereas in an organism the catalysts that decide how the organism is to be constructed are themselves products of the same organism; they are not supplied from outside, and, in Rosen’s words (Rosen, 1991), an organism is closed to efficient causation. The essential idea is illustrated in Figure 1, which represents the enzymes that catalyse the metabolic reactions as being themselves products of the same metabolism. However, there is a serious difficulty with this representation, because synthesis of the enzymes also requires catalysts, which are not shown in Figure 1. The problem is illustrated by the more abstract representation of metabolism and replacement in Fig. 2. This illustrates the problem but does not attempt a solution, as this requires a deeper analysis (Letelier et al., 2006; Cornish-Bowden et al., 2007).

We must not forget, of course, that life must satisfy not only metabolic closure but also bioenergetics, which implies that a living organism is an open system in the thermodynamic sense, with a flux of matter and energy. There is no conflict here, however, because the two statements involve different levels of causation: closure to efficient causation does not imply closure to material causation, which would be absurd.
This excursion into philosophy may appear superfluous in a discussion of biological engineering, but it is absolutely essential if one is to understand the limitations on what biological engineering can do, not just at present with present levels of technology, but in principle and for ever. If engineering just means tinkering with existing organisms then that is, of course, perfectly possible, and that is what genetic engineering is. But the difference between that and designing an entirely new organism is more than just a difference of scale; it is absolutely fundamental and probably unbridgeable. If we accept Rosen’s position it will never be solved; if we think he is mistaken we must still recognize the magnitude of the gap, and we still need to understand the nature of the difficulty before there can be any hope of surmounting it. Simply gathering more data about the details of gene products will not lead to a solution.

This conclusion is pessimistic, and is perhaps wrong. However, if it is wrong it needs to be shown to be wrong: it will not be sufficient to continue with today’s reductionist approach in a pious hope that the tremendous advances that this approach brought in the 20th century will be matched by even greater advances in the 21st. If it is right it has practical consequences, just as much as the understanding of chaotic dynamics that came from the studies of 25 years ago had practical consequences: it may have failed to solve the problem of long-range weather forecasting, but it still convinced most meteorologists not to waste more time and effort in pursuit of an unattainable objective.

Another practical implication of Rosen’s theory comes from the new light that it sheds on the occurrence and role of multiple functions of proteins. It has been known for many years that some of the lens proteins of the vertebrate eye are identical to glycolytic enzymes, and examples of such “moonlighting” proteins are being discovered with increasing frequency (Tipton et al., 2003). Until now they have been regarded as an interesting feature of life, albeit one that can greatly complicate the diagnosis and treatment of monogenic diseases (Sriram et al., 2005), but without deep significance. However, analysis of how organisms can be closed to efficient causation indicates that metabolic closure would be impossible without multifunctionality (Letelier et al., 2006; Cornish-Bowden et al., 2007). This in turn implies that the instances of multifunctional proteins that are known today will prove to be a small proportion of those that exist.

The message from metabolic control analysis

If the message from the hard systemic approach of Rosen is the rather negative one that it is useless to continue trying to achieve the impossible, the message from the softer systemic ideas embodied in metabolic control analysis is more positive. There is little doubt that many of the early difficulties in interpreting the first genome data were due to failure to incorporate the lessons from 25 years of metabolic control analysis, and little doubt that these lessons will need to be understood better in the future.

The hope when genome data first became available was that determining the functions of the genes first identified by examining DNA sequences would be relatively easy: it would be sufficient to examine what function was lost—or what other deleterious effects appeared—when the gene was deleted. To the surprise of many, though not of those who had studied metabolic control analysis, the most common effect of such
knockout experiments was that there was no effect: many genes proved to be “silent” and had no obvious phenotype. In the case of *Saccharomyces cerevisiae*, the probability is about 80% that deleting a randomly chosen gene will have no effect on growth or on any easily measurable metabolic flux (Cornish-Bowden and Cárdenas, 2001b). This should not have been a surprise, as it follows directly from the summation properties that control coefficients obey (Kacser and Burns, 1973): these imply, in simple terms, that flux control is shared (unequally) among all the enzymes in the system, and as the number of enzymes is typically large the share that each one has is typically small. Although this idea is usually applied to specific metabolic pathways, it also applies to gross fluxes like rates of growth, where thousands of enzymes may share control. Eliminating an enzyme entirely may still have an effect—part of the reason why the probability in yeast is only about 80%, rather than 100%—if the reaction catalysed by the deleted enzyme is absolutely essential for growth in the culture conditions used; but organisms typically have more than one way of satisfying an essential function, and eliminating one of the ways will not necessarily therefore eliminate the function. The simplest way of having more than one way to fulfil a function is to have multiple enzymes that catalyse the same reaction, or isoenzymes, but more complicated ways exist as well, and all of them imply that even deleting an enzyme that catalyses an essential function may be phenotypically silent, at least if its effects are measured solely in terms of fluxes.

The second lesson to be learned from control analysis is that metabolite concentrations are typically much easier to perturb than fluxes, with the result that a gene that appears to have no phenotype can often be revealed by studying the effect on metabolite concentrations of deleting it, even though the deletion has no effect on fluxes, and especially not on a gross flux like growth. When this is done, even genes for isoenzymes cease to be silent (Cornish-Bowden and Cárdenas, 2001b; Raamsdonk et al., 2001).

**Metabolic regulation**

Some of the more hostile attitudes to metabolic control analysis, such as that of Atkinson (1990), derive from a perception that it has no use for such central concepts in biochemical regulation as feedback inhibition by end products (Umbarger, 1956; Yates and Pardee, 1956) and cooperative and allosteric interactions (Monod et al., 1965; Koshland et al., 1966). It is true that these concepts are are not always very evident in discussions of control analysis, but they are not ignored, though their classical interpretation has required some revision. Contrary to the usual idea, these mechanisms are almost irrelevant to flux control, but they are crucial for concentration control: fluxes can be controlled efficiently without using any of the classic mechanisms, but only at the cost of huge variations in metabolite concentrations (Hofmeyr and Cornish-Bowden, 1991; Cornish-Bowden et al., 1995; Cornish-Bowden and Cárdenas, 2001c).

The essential points, which apply when the flux through a metabolic pathway is determined by the need for its product, can be summarized in terms of the law of supply and demand: the classic mechanisms exist so that when a change in flux occurs in response to a change in demand this is accompanied by minimal changes in the concentrations of the intermediates in the pathway (Hofmeyr and Cornish-Bowden,
This law derives, of course, from economic theory, but it works much more efficiently in biochemistry than it does in economics, as long as one remembers that not all pathways need to be regulated by demand. The mammalian liver, for example, phosphorylates glucose not primarily to meet its own relatively modest need for energy, but to prevent hyperglycaemia; in other words, most glucose phosphorylation in the liver is supply-driven, and hexokinase D ("glucokinase"), the principal enzyme involved, shows none of the characteristics that one would expect an enzyme that responds to demand to have (Cárdenas, 1995; Cornish-Bowden and Cárdenas, 2004; Cornish-Bowden and Nanjundiah, 2006).

Although much of our knowledge of metabolic regulation was originally derived from studies in bacteria, such as the classic work on aspartate metabolism in Escherichia coli (Stadtman et al., 1961; Patte et al., 1967), the principles also apply to higher organisms in which different organs fulfil specialized tasks. This is well illustrated by the hexokinase isoenzymes in the human, where different isoenzymes predominate in different tissues and have properties appropriate for the needs of the tissues concerned (Cárdenas, 1995). As we have mentioned, hexokinase D, the characteristic isozyme of liver, has the properties expected for an enzyme regulated by substrate supply: lack of saturation at physiological glucose concentrations, a cooperative response to glucose, and insensitivity to product inhibition by glucose 6-phosphate. In complete contrast, brain has a need for glucose that must be satisfied even when other tissues cannot be supplied, and hexokinase A, which predominates in brain, is saturated by very low concentrations of glucose, and is inhibited by glucose 6-phosphate: the rate of glucose phosphorylation thus changes with changes in demand but not with changes in the availability of glucose. Muscle also has a need for large amounts of glucose, but it is less crucial to satisfy it than it is for the brain, and hexokinase B, predominant in muscle, has similar properties to hexokinase A with the important difference that it is unsaturated at low glucose concentrations that would be saturating for hexokinase A. This example could be pursued in greater detail, but the point is already clear, that the principles of metabolic regulation derived from studies of prokaryotes apply also to higher organisms.

The practical biotechnological importance of this is that before one can modify the metabolic behaviour of an organism for biotechnological ends one must understand the organism’s regulatory mechanisms and their functions, which have evolved in order to satisfy its own needs, not those of biotechnology. This last may seem an obvious point, and indeed it ought to be, but it is sometimes forgotten. When one reads, for example, that the ability of a virus to resist a drug is due to "careless transcription" (Smith and Simons, 2004) a line has clearly been transgressed. Whatever intentions one may meaningfully impute to a virus, a willingness to die in the cause of human health is certainly not one of them. More generally, many attempts to increase yields of desirable metabolites by overexpressing supposedly rate-determining enzymes have failed because of failure to recognize the existence of regulatory mechanisms that have evolved precisely to prevent accumulation of metabolites that the organism does not need. As a particular example, efforts to increase yields of ethanol in Saccharomyces cerevisiae have been thoroughly analysed (Niederberger et al., 1992), but the principle applies widely (Fell, 1997). Only in the unusual case of a pathway regulated by supply, such as glycogen synthesis in the mammalian liver, does one find a rate-limiting enzyme
that one can overexpress in order to increase the flux (Cárdenas, 1995; Cornish-Bowden and Cárdenas, 2004; Cornish-Bowden and Nanjundiah, 2006; Agius et al., 2004).

**Concluding remarks**

Despite the great and rapid advances in biological knowledge that have come from the sequencing of genomes and from the new fields that have grown out of this, our attitude in this article has not been entirely positive about the usefulness of the progress in these fields. On the one hand, we cannot deny our belief that there has been far too much emphasis on the accumulation of data and far too little on the questions that can be answered by means of the data accumulated, and this belief can only be strengthened by a recent article (Abecasis et al., 2007) that lays great stress on hopes that “new, inexpensive sequencing methods [will] provide even higher-throughput capabilities and allow more detailed analysis of individual genomes” but almost none on why this might be desirable. On the other hand, some suggestions about how systemic ideas should influence how research is actually done may be helpful. Clearly researchers cannot wait for a complete or even a partial understanding of life at the sort of level sought by Rosen, because this may not come in this century, if at all. Nonetheless, realization of the limitations of the machine analogy may help to avoid some wasted effort. It is more important also to realize that data accumulation is not useful in itself; it is or may be useful in relation to long-term questions that the data may help to answer. In other words research needs to be hypothesis-driven and not just technique-driven.

**References**


Cárdenas, M. L., 1995. “Glucokinase”: its regulation and role in liver metabolism. R. G. Landes, Austin, TX, USA.

Table 1. “Omics” terminology in the biological literature.¹

<table>
<thead>
<tr>
<th>Search term</th>
<th>Total</th>
<th>2000–2007 (%)</th>
<th>First use</th>
</tr>
</thead>
<tbody>
<tr>
<td>“genome”</td>
<td>128703</td>
<td>56</td>
<td>1953</td>
</tr>
<tr>
<td>“genomics”</td>
<td>25675</td>
<td>75</td>
<td>1987</td>
</tr>
<tr>
<td>“systems biology”</td>
<td>2027</td>
<td>99</td>
<td>1988</td>
</tr>
<tr>
<td>“proteome” or “proteomics”</td>
<td>15720</td>
<td>98</td>
<td>1995</td>
</tr>
<tr>
<td>“transcriptome” or “transcriptomics”</td>
<td>2575</td>
<td>99.6</td>
<td>1997</td>
</tr>
<tr>
<td>“metabolome” or “metabolomics”</td>
<td>612</td>
<td>99.7</td>
<td>1998</td>
</tr>
<tr>
<td>“metabonome” or “metabonomics”</td>
<td>196</td>
<td>99.5</td>
<td>1999</td>
</tr>
<tr>
<td>“fluxome” or “fluxomics”</td>
<td>23</td>
<td>96</td>
<td>1999</td>
</tr>
<tr>
<td>“metagenomics” or “metagenome”</td>
<td>173</td>
<td>99.6</td>
<td>1998</td>
</tr>
<tr>
<td>or “metagenomic”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“interactome” or “interactomics”</td>
<td>161</td>
<td>99.4</td>
<td>1999</td>
</tr>
<tr>
<td>“omics” (whole word)</td>
<td>351</td>
<td>100</td>
<td>2002</td>
</tr>
</tbody>
</table>

¹The Table shows information obtained from the PubMed database (PubMed, 2007) as checked in March 2007. The word “genome” has been in use at least since 1953, but “genomics” as the name of a field of study was invented by the founding editors of the journal Genomics: “For the newly developing discipline of mapping/sequencing (including analysis of the information) we have adopted the term GENOMICS. We are indebted to T. H. Roderick of the Jackson Laboratory, Bar Harbor, Maine, for suggesting the term. The new discipline is born from a marriage of molecular and cell biology with classical genetics and is fostered by computational science” (McCusick and Ruddle, 1987). Percentages in the third column are approximate, as all of the totals in the second column change rapidly. The various “omics” terms should not be confused with other terms, such as glycosome, spliceosome, trypanosome, etc., that contain the unrelated root -some.