Systems Biology

How far has it

Athel Cornish-Bowden (Unité de Bioénergétique et Ingénierie des Protéines, Centre National de la Recherche Scientifique, Marseilles, France)

As a phrase, 'systems biology' was probably invented in 1968 by Mesarović[1], but for the first three decades it had essentially no impact: *PubMed* lists just 19 publications that mention it in all the years before 2000. However, the first output from the human genome project stimulated an explosive growth, and the number of publications referring to 'systems biology' reached about 500 in 2005 alone, the year when the Biochemical Society organized a Focused Meeting entitled 'Systems Biology: will it work?'. After that it became too much effort to continue counting, and the number is certainly much higher now.

One might now hope to see some fruits of all this activity, but scepticism is widespread. Sydney Brenner describes systems biology as "low-input, high-throughput, no-output biology". He has said that the claims of radical systems biology cannot be met, and that a weaker version is just a new name for physiology. It takes a certain amount of chutzpah for a molecular biologist (whose field was memorably defined by Erwin Chargaff as "the practice of biochemistry without a licence") to make that particular criticism, but no matter: it is certainly true that achieving the more ambitious objectives of systems biology will take a long time and will require a greater capacity for analysing the results – and putting them in the context of clearly formulated questions – than is apparent today.

Likewise, in a recent entry on his blog David Colquhoun said that "it seems to me that that most attempts at system biology have been disappointing (please correct me if I'm wrong)". I won't try to correct him, because I don't think he is wrong, notwithstanding some spectacular successes (which he recognized), such as the heart model that Denis Noble and his colleagues developed over many years, starting long before anyone spoke of systems biology. Nonetheless, Hans Westerhoff gave his contribution to the Biochemical Society Focused Meeting of 2005 the resounding title of 'Yes!', and Noble entitled his 'The heart is already working'. (My own was more hesitant: 'Systems biology may work when we learn to understand the parts in terms of the whole'). In a limited sense, Westerhoff and Noble were both right, and if all of the researchers currently pursuing what they call systems biology were doing so at the same high intellectual level as they are, they would probably be right in a more general sense. Unfortunately, however, that is not at all clear, and one may cynically suggest that the main purpose of 'systems biology' is to serve as a convenient tag for labelling all sorts of different kinds of research in the hope that it may attract research grants. For many people, it seems to mean little more than doing the same sort of things that they have always done, but on an ever-increasing (and already gigantic) scale in the hope, and perhaps belief, that some meaning will emerge from the mountain of data that is emerging from the various types of '-omics'. There needs to be a balance between data production and analysis in depth, especially if multiomics is involved.

For Westerhoff and Noble, and for me, systemic thinking means more than just collecting a huge amount of data, and more than just moving from a focus on the elements of a system towards a focus on the interactions between them; it means seeking a vision of the system as a whole: not just its components, not just the connections between them, but a vision of how the whole system works. As Henrik Kacser commented, "One thing is certain: if you want to understand the whole you must study the whole." Multiomics can be regarded as a step towards this, but only if the various branches are fully integrated. Kacser himself wanted to step outside the black box, to define general principles about how metabolic systems behave, rather than to study each enzyme in mechanistic detail. However, many of those working today in the tradition that he created are combining these general principles with experimental information by computer modelling of metabolic systems

come?

with the much larger amount of kinetic data that is now available. The same principle applies also to all of the '-omics' fields, because in all of these, the experiments ought to be designed to answer specific questions. The right questions, which can only come from a broad and deep biological knowledge, can lead to a global view of a system as a whole.

The number of real systems modelled with real experimental data remains surprisingly small. The JWS-Online database curated by Jacky Snoep now contains about 90 models, an impressive number until one realizes that most of the models in it are not metabolic models built with experimentally determined kinetic parameters. Of the ones that are, there are not many more than there were 10 years ago, and they refer to a restricted group of systems, including glycolysis in the erythrocyte and the bloodstream form of *Trypanosoma brucei*, amino acid metabolism in *Escherichia coli* and *Arabidopsis thaliana*, and sucrose production in sugarcane. These studies have a number of important points in common: they were done on the basis of data collected for the purpose (or at least with the possibility of subsequent modelling kept in mind), in conditions as close to physiological as possible, and normally obtained by a single research group. In these circumstances, a computer model can produce results very close to those observed experimentally, as Hans Westerhoff and his colleagues have shown, so there is some basis for believing that the predictions it makes for conditions that have not been studied.

Unfortunately, these are exceptional. Despite the huge amount of experimental kinetic information apparently available about a wide array of enzymes in a wide array of organisms (albeit with a strong bias towards *E. coli*, *Saccharomyces cerevisiae* and some mammals), most of this information has not yet been embodied in usable models, and probably never will be. The coverage of enzymes is too haphazard, the conditions are too arbitrary (from the physiological point of view), often failing to take account of such important features as the reverse reaction or the presence of effectors. That is why most of the groups involved in the studies mentioned chose to start again from the beginning, using only data collected for the purpose. That is also why the number of studies is not very large.

Although the disadvantages of haphazard data collection are most obvious in models of whole systems, they extend also to many experiments carried out with purely mechanistic objectives. How many papers on enzymology have we all seen that do not adequately specify the assay pH, the components of the buffer system, the method of purifying the enzyme, or, nowadays, the post-translational modifications? The Beilstein-Institut has long been familiar to organic chemists in the name of one of the world's first databases. In recent years it has become increasingly active in more biological domains, has created the STRENDA (**St**andards for **R**eporting **E**nzymology **D**ata) Commission to make recommendations for reporting results. The current recommendations have now been adopted by most of the leading journals of biochemistry, although not yet by the *Biochemical Journal*. If they become widely followed, they will enhance the usefulness of databases such as BRENDA and KEGG, allowing the data found in them to be used in the construction of much better models of metabolic systems. A planned web form is currently at the stage of inviting expert criticism[2]: this is intended to make it easier for authors and editors to check whether necessary information is included in papers submitted for publication. The commentary in the same issue by Palsson and Zengler on the integration of multiomic datasets[3] is also pertinent to points that I have mentioned above.

At a different level, systemic thinking is thinking not just about parts of organisms, such as the heart or the glycolytic pathway, but about the organism as a whole: how is it organized? How does it make itself? How does it maintain itself in the face of changes in its environment? What exactly is it that allows us to say that it is alive? The definition of life is far more subtle and difficult than is usually admitted. Crystals can grow and reproduce themselves, but they are not alive. Mules can certainly be alive, but they cannot reproduce themselves. Any naive definition of life leads immediately to contradictions, of which these are just two examples. This is a question that biochemists, among the most reductionist of all

biologists, have preferred not to think about. The field has been developed instead by a heterogeneous collection of non-biochemists: Robert Rosen regarded himself as a mathematical biologist, but most would regard him as a mathematician interested in some recondite questions of biology; Humberto Maturana and Francisco Varela were neurophysiologists; Tibor Gánti is a chemical engineer; Stuart Kauffman perhaps comes closest to being a biochemist, but is far from being an off-the-shelf typical biochemist. With such a disparate collection of people, one should perhaps not be surprised that, despite considerable overlap in some of their ideas, they ignored one another virtually completely: none of their principal works makes any mention at all of any of the others. Only now is some effort being made to make a synthesis in which the similarities and differences are clearly identified and analysed.

In this context, it is also unsurprising that people claiming to be systems biologists can make grandiose claims about understanding organisms and creating artificial organisms without ever needing to concern themselves with what a living organism is. Farmers have been creating new organisms since the dawn of agriculture, but these don't count because they are exercises in selection. For the last few decades, geneticists have been creating new organisms by genetic manipulation, but that also doesn't count because the great majority of genes are just the natural ones that were there all along. There has, however, been a much publicized claim that replacing the genome of a bacterium with a synthetic genome does, at last, constitute creation of a new organism. I want to explain why this is no more than an incremental advance on what farmers have been doing for centuries and geneticists for 30 years. As a step towards artificial life, it is such a trivial step that it doesn't count either.

Suppose I were to announce at a meeting of the Biochemical Society that I had a three-part plan to come to the next one using the power of my muscles alone. This would no doubt be greeted with scepticism (given that I live a long way away from anywhere the next meeting is likely to be held), so I would say that I had already achieved the first two stages, and the third was only a matter of time. The first stage involved walking to where I had parked my bicycle, and I had solved that part of the problem long ago. The second involved using my bicycle to reach the airport. That was more difficult, but I'm happy to say that I've now solved that as well. The third part still needs to be worked out, so for the moment I'm inserting myself into an airliner that can bring me to Gatwick, and from there I can reach the site of the meeting by public transport. If I claimed that this was a significant step towards solving the problem as stated I would be ridiculed, and the idea that significant progress has been made towards synthesizing an organism should also be ridiculed. Synthesizing an entire genome was, of course, an impressive technological feat, but it was based entirely on what was already known. Inserting it into an existing organism is simply using the bacterium's expertise to solve the fundamentally difficult part of the problem.

Many practitioners of systems biology will doubtless see this as a very negative assessment of their activities, but we should recall Robert Rosen's comment:

Quite early in my professional life, a colleague said to me in exasperation, "The trouble with you, Rosen, is that you keep trying to answer questions nobody wants to ask." This is doubtless true. But I have no option in this; and in any event, the questions themselves are real, and will not go away by virtue of not being addressed. This attitude, I know, has estranged me from many of my colleagues in the scientific enterprise, and has put me far from today's 'main stream'.

Quite so, and as long as biologists continue to think that studying systems means collecting huge amounts of data in the absence of a global view of the organism, they will not be progressing towards a real understanding of how organisms stay alive. ■

References

1. Mesarović, M.D. (1968) in Systems theory and biology vol. 351 (Mesarović, M.D., ed.) pp. 59–87, Springer-Verlag, New York
2. Apweiler, R., Armstrong, R., Bairoch, A., et al. (2010) Nat. Chem. Biol. **6**, 785
3. Palsson, B. and Zengler, K. (2010) Nat. Chem. Biol. **6**, 787–789

Athel Cornish-Bowden is Directeur de Recherche Émérite in the CNRS Unit of Bioenergetics and Protein Engineering in Marseilles. His main interests are in enzyme kinetics in general, and their application to multi-enzyme systems in particular. email: acornish@ifr88.cnrs-mrs.fr

Note: The Biochemical Journal *has not made the STRENDA recommendations mandatory but the section on reporting enzyme kinetic activity in its Instructions to Authors refers authors to these recommendations for additional suggestions of how to report and interpret kinetic data.*