

Finding wood among the trees

Is the reductionist ambition for molecular biology in danger of being thwarted by the volume of the data it produces, or even by the absorbing interest of its collection?

Is molecular biology running into a dead end? Is it possible that the exciting programme to make biology a branch of physical chemistry is at risk of frustration by the sheer volume of the data now being acquired, by the complexity of some of the problems whose solution seems essential for the success of the reductionist programme — and by the honest fun of manipulating the techniques now available?

It is a curious situation, perhaps engendered by the plethora of discovery in the past thirty years and more. The genetic code, messenger RNA and its function, reverse transcriptase and (quickly) the techniques of genetic manipulation came treading on each other's heels. Nobody can cavil at what has been accomplished.

Most vividly, it would not have been possible to form a view of the cause of AIDS had it not been for the previous quarter of a century's molecular biology. More important, people have a sense that things will eventually be understood. Hardly a week passes without the publication of the amino-acid sequence of yet another cell-membrane protein, usually a receptor of some kind. Protein-DNA interactions, generally believed to determine the regulation of genes, are being dissected. There are now seven interleukins. People are also describing in previously unimagined detail the means by which information is relayed within cells, from membrane to nucleus and in the other direction, as in the processing of proteins *en route* from ribosomes to effector sites.

But now, while nobody can fail to marvel at the variety of the data accumulated, the volume of it is its most impressive attribute. The databanks storing nucleotide and protein sequences would be even further behind if laboratory people were as diligent as they should be in communicating their sequences quickly. Soon there will be comparable problems in storing authentic data from structural studies of biomolecules. If there is a cause for disappointment, it is that the reductionist prospectus should so slowly deliver its promises.

Naturally, it would be folly to ask that biology should become a branch of mathematics; physical chemistry is a sufficiently tall order. Even in the physical sciences, the explanations derived from first principles have not driven out experiment; the whole of hydrodynamics has not been reduced to the solution

of the Navier-Stokes equations. But, outside particle physics and cosmology, there is a sense in which the reductionist recipe works. Students, for example, can be taught the principles and helped to understand when and why they will not suffice.

Sadly, it remains a proper question of molecular biology whether the recent torrent of description amounts to understanding. Future historians may think it odd that so much should have been learned about the molecules on which life depends while so little has been understood about their function, or about life itself. The obvious parallel is with spectroscopy in the 1920s: technique had made it possible to fill catalogues with the frequencies of spectral lines, the search for empirical explanations was vigorously conducted, but only the coming of quantum mechanics made sense of all the data. The obvious answer is that the parallel may be false, and that biology has necessarily complicated explanations.

An example is the problem of protein folding. Should it not by now be almost a problem for undergraduates to solve, to ask what will be the conformation (tertiary structure) of a protein whose amino-acid sequence is known, or has been specified? That, alas, is not yet the case. There are many elaborations of the empirical rules which are the achievement of the past few decades, but few memorable generalizations.

Much the same is true of DNA itself. At least three conformations of the double helix (A, B and Z) are recognized, but there is only the foggiest understanding of how the environment determines the structure of A and/or B, or why particular nucleotide sequences should predispose towards Z instead. Like the protein problem, of course, these are complicated questions, not least because of the certainty that, in molecules in which intramolecular hydrogen bonding is supposed to be a crucial determinant of structure, interactions with extrinsic hydrogen-bonded water molecules are at least as important, but are not yet described.

For lack of the means to tackle these complications, people tackle instead the problems of anhydrous proteins and anhydrous DNA. It is like the drunk who, asked why he searched for his lost latch-key beneath a street-lamp, explained that "It's too dark where I dropped it".

Reflection shows that much the same is

happening elsewhere within the ambit of the programme to make molecular biology the means of explaining life. There is, for example, energy transduction. The molecular structure of muscle fibres is exquisitely described, as is the role of calcium ions, but who can give a simple account of the thermodynamics of the muscle cell? Is molecular biology in that phase corresponding to the introductory phases of infantry training when people are required to learn the names of the pieces of a machine-gun, but are not told how they function or what the whole assembly is for?

It is not essential that the reductionist programme should say how life began, but its successes would be more widely appreciated if it could provide a kind of existence theorem. But instead there seems to have been a succession of fashions. In the 1950s, electrical discharges in mixtures of ammonia, methane and water were shown to yield modest amounts of glycine; in the 1960s there were speculations that the polymerization of cyanogen followed by its hydrolysis would yield even more remarkable products; and there is now a general opinion that the first self-replicating molecules must have been RNA.

Fashion also shapes what is understood of the working of the brain. In less than a quarter of a century, we have been offered the promise that neurotransmitters would tell how the brain works and then that small peptides would do the trick. Now, neural networks have come into their own. Of course, the recognition of the role of dopamine in the substantia nigra has helped in the treatment of parkinsonism (and even, perhaps, of schizophrenia), while neural networks are the most illuminating metaphor of the brain so far. But how well do we understand how the brain works?

The most conspicuous gaps in the molecular description of life are the answers to simple questions such as that. Or, perhaps more rudimentary, what provokes a cell to divide?

Is there a danger, in molecular biology, that the accumulation of data will get so far ahead of its assimilation into a conceptual framework that the data will eventually prove an encumbrance? Part of the trouble is that excitement of the chase leaves little time for reflection. And there are grants for producing data, but hardly any for standing back in contemplation.

John Maddox