A Beginning of the End of the Holism versus Reductionism Debate?

Molecular Biology Goes Cellular and Organismic

PETER SCHUSTER

Peter Schuster, Editor in Chief of Complexity, *is at the Institüt für Theoretische Chemie der Universität Wien, A-1090 Wien, Austria; E-mail: pks@tbi.univie.ac.at*

About 20 years ago the famous biologist and scholar of evolutionary game theory, John Maynard Smith commented on the holism versus reductionism debate between organismic and molecular biologists with the following phrase [1]: "… Holists are, I think, in a weaker position, if only because progress has been so much faster from the bottom up than from the top down. Yet, I do share their conviction that there are laws that can only be discovered by research on whole organisms, and on populations of organisms. Almost all my own work has been done at those levels. What should be the attitude of a biologist working on whole organisms to molecular biology? It is, I think foolish to argue that we are discovering things that disprove molecular biology. It would be sensible to say that there are phenomena that they will one day have to interpret in their terms. …". As I shall try to outline later in this essay the expectations of John Maynard Smith have today become (almost) true in the emerging field of systems biology.

The holism versus reductionism debate is an old theme in philosophy and science. Sometimes holism is even traced back to Aristotle's "Metaphysics," which contains the famous sentence: "The whole is more than the sum of its parts." The problem here is the usage of the term "sum." If sum implies a simple arithmetic sum, the statement expresses nothing more and nothing less that the parts of the whole do interact. For every property $F(\Xi)$ of an ensemble $\Xi = \{X_1, X_2, \ldots, X_n\}$ we can write

$$F(\Xi) = \sum_{i \in \Xi} f^{(1)}(X_i) + \sum_{i \in \Xi} \sum_{i>j \in \Xi} f^{(2)}(X_i, X_j) + K + f^{(n)}(X_1, X_2, \ldots, X_n).$$

The terms in this cluster expansion corresponds to one-body, two-body, and higher body interactions up to the n body term. In the naive interpretation Aristotle's sentence says that there are nonvanishing two or more body terms and therefore algebraic additivity is violated, and this is almost always the case except in ideal gases.

Apart from trivial algebraic non-additivity Aristotle's sentence has three deeper interpretations that can be cast into the following questions:

 Vol. 13, No. 1
DOI 10.1002/cplx.20193
Published online 8 October 2007 in Wiley InterScience
(www.interscience.wiley.com)

1. Are new properties emerging in the progression from a lower to a higher hierarchical level?
2. Can we describe the phenomena on a higher hierarchical level by means of fundamental laws operating on the lower level or do we need radically new laws of Nature that become operational in the form of specific forces only on the higher level?
3. Are there limits in the predictability of complex systems that cannot be compensated by improved knowledge on the parts of the system?

The answer to question 1 will almost always be "yes." We consider, for example, the world of atoms and the hierarchically higher world of molecules being composed of atoms. The chemical bond is part of the notions needed to describe the properties of molecules but does not exist in the world of atoms: The chemical bond is an emergent property of molecular physics or chemistry.

Question 2 asks whether the postulation of special laws and forces such as the notorious vital force—*vis vitalis*—behind living organisms is indispensable. Such additional forces causing essentially novel regularities on the higher level, which need new fundamental laws to describe them, are much harder to argue and to verify. In the life sciences this second version of holism has become very unpopular and the majority of scientists would currently agree that it is extremely unlikely to discover new fundamental forces in biology, psychology, or sociology. In other words, there is a common belief that neither biology nor psychology nor sociology will lead to observations that contradict contemporary physics.[1]

Question 3 addresses so-called scientific holism and finds its confirmation in the existence of principle reasons such as quantum physical uncertainty or deterministic chaos and technical limitations, for example, incomplete information on initial and boundary conditions.

The common form of reductionism contrasting holism of type 2 is often called hierarchical reductionism [3]. It is related to the idea of a unity of science and states that complex systems can be described by a hierarchy of levels in which each form of organization is described in terms of the objects of the next lower level. Within science, hierarchical reductionism is expressed, for example, by the statements: Fundamental chemistry is based on physics, fundamental biology is based on chemistry, psychology is based on biology, sociology is based on psychology, and, eventually, political science, anthropology, and economics are based on sociology. A majority of scientists is accepting the first two reductions, chemistry $\Rightarrow$ physics and biology $\Rightarrow$ chemistry, but the other reductions are often strongly opposed by many researchers. Examples are the controversial discussions of the 1970s on the interpretations of observations from sociobiology or evolutionary psychology [4]. More recently, the ongoing debate has reached mind and brain—taking up older issues—because neurobiologists are now localizing complex psychological phenomena, for example, free will or religious behavior, through activities in certain areas of the brain [5, 6].

Methodological reductionism or the reductionists' program, on the other hand, is *the* method of handling problems in science if one aims at going beyond pure narrative descriptions. Physics is the discipline that has most experience with reduction but everywhere in science experimental exploration of regularities requires reduction in the sense of simplification and constant environments, in particular, in many variable systems. Even for understanding how and why the whole is more than the sum of its parts, profound knowledge of the parts is indispensable. It seems useful to end the academic holism versus reductionism debate by referring to the famous biologist John Maynard Smith [1], who had a rather pragmatic view on the subject. He compares macroscopic biologists as pursuing a holistic strategy by means of a top-down approach to describe the phenomena observed in biology with the reductionists' program of molecular biologists, who perform a bottom-up approach to interpret biological phenomena by the methods of chemistry and physics. As said above, he is sympathetic with the holistic approach but rejects holistic arguments boiling down to the claim that, because we do not understand some phenomena at present, there must be some special (vital) force that is responsible for them. He says: "… As it happens, I do not understand how modern sewing-machines work, but this does not lead me to suppose that the laws of topology have been broken: Indeed, I feel confident I could find out if someone would let me take one into pieces. …". Since the beginnings of physiological chemistry—later called biochemistry—in the 19th century, chemists, biochemists, and molecular biologists are taking organisms into pieces. It seems to me that molecular life sciences have now reached a point at which macroscopic properties of cells and organisms can be interpreted and understood by this bottom-up approach.

The most spectacular progress in molecular genetics of the second half of last century was the invention of new DNA sequencing techniques by Walter Gibert and Frederik Sanger who were both awarded the Nobel Prize in 1980. These protocols allow for fast sequencing of large DNA stretches and eventu-

[1] *A related but more radically formulated view comparing top-down explanations with deus ex machina solutions or skyhooks is found under the heading "Skyhooks or cranes?" in Dennett [2].*

DOI 10.1002/cplx

ally of whole genomes.[2] Since then a large number of genomes has been sequenced, and questions previously restricted to small pieces of DNA can now be raised and answered on the level of the genetic information of the entire cell. Reconstruction of phylogenies from molecular data was raised to another level, and questions of horizontal gene transfer[3] can now be answered precisely. The sequence of a genome, however, is not the solution of understanding cells and organisms; it is just the determination of the chemical formula of the DNA of a whole cell. Nevertheless, *genomics* initiated the most fundamental developments in molecular life sciences. The next natural step was to identify and analyze all cellular proteins encoded by the DNA. Again, new techniques started *proteomics*, the discipline aiming at a complete list of proteins in the cell and their interactions. At the same time it became clear that even the highly complex network of protein interaction tells very little about the regulation of cell activities, in particular, in higher organisms. Earlier estimates considered about 97% of mammalian DNA as "junk" in the sense that it was not translated into protein. Now, highly conserved sequence stretches outside regions coding for proteins were detected [7] and even more important for practical purposes, single nucleotide mutations in such DNA regions were found to be characteristic for certain forms of disease, in particular, diabetes type two (see, e.g., [8]). Within the last 20 years it became more and more clear that the prokaryotic view of gene regulation and processing is only part of the story in higher organisms. RNA, originally considered only as the messenger of genetic information from DNA to protein, was found out to be a central molecule for regulation of cellular dynamics. Several completely unexpected functions of small RNA molecules were discovered; as an example we mention small interfering RNAs [9, 10], which were found to have an important role in the regulation of gene expression.

In 2004 an international consortium of researchers started the ENCODE (Encyclopedia of DNA elements) project [11]. The goal is to study all transcripts from genomes and create a complete mapping their functions. Today, roughly 300 scientists are involved in this global undertaking. They published the results of their pilot study covering 30 Megabases (Mb) or about 1% of the human genome, in June 2007 [12, 13]. The major findings of the consortium were a surprise: (i) The human genome is pervasively transcribed and the majority of its nucleotides are associated with at least one primary transcript, (ii) large numbers of novel non-protein-coding transcripts have been identified, (iii) numerous previously unrecognized transcription start sited have been identified, (iv) a total of 5% of the nucleotides in the genome can be confidently identified as being under detectable evolutionary constraints in mammals, (v) different functional elements vary greatly in their sequence variability, and (vi) surprisingly, many functional elements are seemingly unconstrained across mammalian evolution indicating the existence of large pools of variable elements. Complete functional analysis of DNA transcription products is indicating usage of the entire genome for regulatory and housekeeping purposes, and not only 3% as in the old days of "Junk DNA". Neutrality on the sequence level is a fact, and, it is much greater than expected. We still have to learn why some regions of the DNA are under severe constraints, whereas others can vary almost unrestrictedly. For the first time in molecular biology up-scaling of the study seems to aim at a complete mapping of all DNA-related functions in the cell. A factor 100 is large and it remains to be seen whether or not the ENCODE consortium is sufficiently strong to derive the *transcriptome* of the entire human genome. Surely other groups will join in, if necessary, and in a few years we shall have the full information of DNA function on a genome wide basis.

Meanwhile systems biology is already starting to do simulations on the dynamics of complete cells or even entire organisms [14, 15]. Still there is a long way to go before such an approach can make reliable predictions. The networks of cellular metabolism and genetic regulation are huge: Some 20,000 genes and proteins set lower limits to the numbers of molecular players. Moreover, analysis of gene expression in different tissues revealed that the same stretch of DNA can be translated into different proteins in differently differentiated cells. Modern molecular biology confronts theorists, mathematicians, and computer scientists with hitherto unknown complexity, but the path leading toward a bottom-up understanding of life is already visible.

Eventually coming back to our initial holism versus reductionism debate, I make my somewhat provocative statement: In order to understand Nature we can neither dispense from the reductionists' program and its results nor can we totally abolish the holistic view. I shall make my point by means of referring to two examples: the quantum defect in atomic spectroscopy and Mendel's rules of genetics.

In the second half of the 19th century, the scholars of early spectroscopy discovered the basic regularities of term energies in the spectral lines of atoms—mainly hydrogen, alkaline metal atoms, and earth alkaline metal atoms. Apart from very minor deviations the hydrogen spectra fulfilled a $(-1/n^2)$-law for the term energies. Empirical observation showed that the spectra of alkaline metal atoms could still be interpreted by means of the hydrogen formula, provided a quantum defect $\delta_{n,l}$, was introduced, leading to term energies of $(-1/(n - \delta_{n,l})^2)$, where n and l are the principal and the angular momentum quantum numbers, respectively. Individual values for $\delta_{n,l}$ were determined empirically, the atomic spectra were reproduced successfully, but a deeper understanding of atomic structure and

---

[2] *The genome or the genotype of an organism is the complete genetic information that is stored in its DNA.*

[3] *Horizontal gene transfer is transmittance of genetic information between contemporary organisms of the same or different species.*

DOI 10.1002/cplx

spectra was not possible. Attempts were made to interpret the origin of quantum defects by deviations of the orbits for electron movement from circles. All extensions to spectra of atoms with three or more valence electrons failed. The only solution of the problem came from a then new bottom-up theory, quantum mechanics.

Mendel derived his rules of genetics in sexual reproduction from empirical observations of statistical nature. He postulated the existence of indivisible packages of inheritance and could explain the observation that he had made in careful experiments with some properties of peas. The fact that the observed ratios of variants in the progeny were idealized fractions only, and many other properties did not obey Mendel's rules at all, were explained in experimental genetics and population genetics by "skyhooks" or *dei ex machina* called "linkage of genes," "pleiotropy," and "epistasis." The only satisfactory explanations for all hereditary phenomena came with the discovery of DNA, the exploration of the mechanism of recombination at the molecular level, and the identification of genes with Mendel's postulated packages of inheritance. The empirical knowledge of macroscopic biology, of course, was required in order to know what it was that had to be explained. Similarly, other skyhooks summarized as "epigenetics" were used as kind of a garbage bag for all phenomena that escaped an explanation by conventional genetics. By now they too found straightforward explanations at the molecular level. Interestingly, the notion of genes is becoming more and more obscure the closer one looks at the complex dynamics of cellular regulation. Many problems of medicine and pharmacology seem to have their ultimate origin in our still too simple view of cells and organisms. We need to extend the understanding of large networks with complex structures and highly specific nonlinear elements. I believe there will be a day when genes are reduced to what they were at the beginning: a theoretical or formal concept to interpret the empirical rules of Mendelian genetics.

An excellent monograph on the history of physical thought has the title "The How and the Why" [16]. I would like to extend this heading by adding "The What," because notions, properties, purposes, and goals—if there are any—are added to the successful reductionism by the holistic view of Nature. Synergism between the reductionistic approach to carry out decisive experiments and the holistic view of problems clarifying the properties and functions waiting to be explained is required for further progress in our understanding of Nature.

REFERENCES

1. Maynard Smith, J. The Problems of Biology; Oxford University Press: Oxford, UK, 1986.
2. Dennett, D.C. Darwin's Dangerous Idea. Evolution and the Meanings of Life; Simon & Schuster: New York, 1995.
3. Dawkins, R. The Blind Watchmaker. Longman Scientific & Technical: Harlow, Essex, UK, 1986.
4. Caplan, A.L., Ed. The Sociobiology Debate. Readings on Ethical and Scientific Issues; Harper & Row: New York, 1978.
5. Pinker, S. How the Mind Works; W.W. Norton & Co.: London, 1997.
6. Dennett, D.C. Freedom Evolves; Simon & Schuster: New York, 1995.
7. Dermitzakis, E.T.; Reymond, A.; Antonarakis, S.E. Conserved non-genic sequences—An unexpected feature of mammalian genomes. Nat Rev Genet 2005, 6, 151–157.
8. Scott, L.G. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007, 16, 1341–1345.
9. Hüttenhofer, A.; Schattner, P. The principles of guiding by RNA: Chimeric RNA-protein enzymes. Nat Rev Genet 2006, 7, 475–482.
10. Moffat, J.; Sabatini, D.M. Building mammalian signaling pathways with RNAi screens. Nat Rev Mol Cell Biol 2006, 7, 177–187.
11. ENCODE Project Consortium. The ENCODE project. Science 2004, 306, 636–640.
12. Greally, J.M. Encyclopedia of humble DNA. Nature 2007, 447, 782–783.
13. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007, 447, 799–816.
14. Palsson, B.Ø. Systems biology. Properties of reconstructed networks; Cambridge University Press: New York, 2006.
15. Klipp, E.; Herwig, R.; Kowald, A.; Wierling, C.; Lehrach, H. Systems biology in practice. Concepts, implemetation, and application; Wiley-VCh: Weinheim, DE, 2005.
16. Park, D. The How and the Why. An essay on the origins and development of physical theory. Princeton University Press: Princeton, NJ, 1988.

DOI 10.1002/cplx