PHILOSOPHICAL PSYCHOLOGY, VOL. 15, NO. 4, 2002

Routledge
Taylor & Francis Group

BioComplexity: a pluralist research strategy is necessary for a mechanistic explanation of the "live" state

F. J. BRUGGEMAN, H. V. WESTERHOFF & F. C. BOOGERD

ABSTRACT *The biological sciences study (bio)complex living systems. Research directed at the mechanistic explanation of the "live" state truly requires a pluralist research program, i.e. BioComplexity research. The program should apply multiple intra-level and inter-level theories and methodologies. We substantiate this thesis with analysis of BioComplexity: metabolic and modular control analysis of metabolic pathways, emergence of oscillations, and the analysis of the functioning of glycolysis.*

1. Introduction

The discovery of the structure of DNA by Watson and Crick in 1953 initiated the so-called molecular biological revolution in biology (Watson & Crick, 1953a,b). It culminated in the deciphering of the entire nucleotide sequence of the human genome in 2001 (International Human Genome Consortium, 2001; Venter *et al.*, 2001). At present, the sequencing of entire genomes of both prokaryotes and eukaryotes progresses continuously. Therefore, in principle, all the genes, and their corresponding proteins, present in a large number of different species are known or will be known in due time. The realization of this aim of functional genomics will be a fabulous scientific achievement. Moreover, especially in the last 30 years of the last century, a multitude of different experimental and computational techniques has been developed in genetics, biochemistry, physiology, biophysics, and structural biology. Combining the information obtained through studies on functional genomics and the aforementioned techniques should make a head start for cell biology possible in the 21st century (cf. Westerhoff, 2000, 2001, in press).

One of the major challenges of cell biology is how to deal with the overwhelm-

F. J. Bruggeman, Molecular Cell Physiology, Biocentrum Amsterdam, de Boelelaan 1085 NL-1081 HV, Amsterdam, The Netherlands, email: frankb@bio.vu.nl; H. V. Westerhoff, Molecular Cell Physiology, Biocentrum Amsterdam, de Boelelaan 1085 NL-1081 HV, Amsterdam, The Netherlands; Mathematical Biochemistry, Biocentrum Amsterdam, de Boelelaan 1085 NL-1081 HV Amsterdam, The Netherlands; Stellenbosch Institute for Advanced Study, South Africa, email: hw@bio.vu.nl; F. C. Boogerd, Molecular Cell Physiology, Biocentrum Amsterdam, de Boelelaan 1085 NL-1081 HV, Amsterdam, The Netherlands, email: fcb@bio.vu.nl

ISSN 0951-5089/print/ISSN 1465-394X/online/02/040411–30 © 2002 Taylor & Francis Ltd
DOI: 10.1080/0951508021000041996

ing complexity of unicellular life. Any living cell consists of a densely packed cytoplasm (Goodsell, 1991) harboring thousands of molecules and macromolecules that are in constant motion; they regulate and catalyze reactions. The macromolecules either function through direct physical interaction with each other, or through more indirect interactions (e.g. by catalyzing conversions of small molecules to which they and others respond). Hereby, a complex, dynamic biochemical network is formed that operates as a sort of unit—the living cell. Furthermore, current studies in cell biology indicate that in some (though certainly not all) aspects, living systems are spatiotemporally organized. This influences the properties of their constituent (macro)molecules and processes, and complicates experimental research and conceptual understanding of the "live" state. An example is the compartmentation of metabolites into separate pools for each of the classical intracellular compartments (organelles), with transport between the compartments, making measured average concentrations meaningless. However difficult to grasp as they may be, such 4D biochemical networks allow living systems to display their emergent "live" character.

One candidate strategy to arrive at the elucidation and mechanistic understanding of the "live" state might be to continue doing research via the usual reductionist approaches. This will result in an enormous body of physicochemical data on the entities and processes that constitute living organisms. Let us suppose that one combines the metabolic stoichiometric network—as inferred from annotated genome sequences and/or obtained through classical biochemical or physiological methods—with all kinetic and thermodynamic data on (macro)molecular interactions and catalytic processes, largely obtained through *in vitro* "wet" biochemistry, for a particular (unicellular) organism (Covert *et al.*, 2001; Westerhoff, 2000, 2001). Then, one has essentially all the physicochemical information that underlies the functioning of a "live" cell at one's disposal. Now, one may wonder whether: (1) this data set provides a sufficient description of the organism, and (2) (cell) biology ends here or just begins. In other words, are the reductionist research programs of the life sciences sufficient to explain the "live" state (mechanistically)?

We will argue that there is more to (cell) biology than the mere completion of the physicochemical data set of all processes that take place within a particular unicellular organism through a reductionist research program. Such a data set would not be sufficient on its own to explain the "live" state mechanistically. Rather, one may argue that it constitutes the start of a "true" biological science (Westerhoff, 2000, 2001). It will involve the replacement of the conventional "hard-reductionist" research strategy (as advocated particularly in the molecular biological revolution) by a pluralist research strategy (including reductionist and holistic features), aimed at explaining "BioComplexity." To be sure, we do not argue that reductionism as such is futile. On the contrary, we argue that it is an essential component of an ultimate synthesis with its antithetic holism. Without the former we will not be able to obtain the necessary physicochemical data set in the first place.

The general outline of this paper is as follows. First, we will describe the reductionist and anti-reductionist positions taken within (cell) biology. This description only serves as a preamble to our BioComplexity argument; it is not a survey of

the philosophy of biology. Then, after discussing *in silico* replicas of biochemical networks we will arrive at a position of explanatory pluralism within the philosophy of biology. Next, we will illustrate our multilevel approach by applying various intra-level and inter-level theories and methodologies to the understanding of aspects of the living cell. In particular, we draw attention to the potential of inter-level concepts. We do this in the explication of some recent examples of BioComplexity in biochemical systems (metabolic and modular control; emergence of oscillations) and in discussing glycolysis in the light of various integrative conceptual and experimental approaches.

2. (Anti-)reductionist strategies in cell biology

Reductionism and anti-reductionism in cell biology are based on the assumption that living systems are arranged hierarchically on multiple levels; smaller systems (i.e. members of the lower-level) harboring physical entities and processes mutually interact and together form larger (i.e. higher-level) systems. Subsequently, the latter systems may function in their turn as subsystems for the formation of still larger systems, and so on. The organizational hierarchy in cell biology starts with the properties of individual organic molecules (e.g. amino acids, nucleotides), and goes via oligomers and polymers (e.g. enzymes, structural proteins, DNA, RNA, fats) to cell organelles (e.g. mitochondria, plastids, nuclei) and finally to the whole cell, which then interacts with its environment. Apart from a hierarchy based on the size and nature of cellular constituents, other useful hierarchies can be formulated. Cellular function can also be divided in functional units that group together large numbers of processes. For instance, a broad division in transport, anabolism, catabolism, and maintenance is sometimes useful. Alternatively, one may distinguish between replication, transcription, translation, and metabolism. The borders of the various levels in the biological organizational hierarchy are not strictly defined.

2.1. Reductionist strategies

Reductionist explanations of biological phenomena at a particular level of biological organization (e.g. metabolic pathway) are invariably sought in terms of properties of constituents of lower levels of organization (e.g. metabolites and enzymes). This research strategy led to the successful methodology of chopping up biosystems into smaller fragments (e.g. enzymes and organelles), and studying these parts *in vitro*.

Reductionism can be envisaged as a continuum; it ranges from radical to pragmatic forms. Reductionist biologists will take—more often implicitly than explicitly—a certain stance within the continuum. Some biologists are convinced that cellular processes will eventually be explained in terms of fundamental theories from physics. The following citations illustrate their point of view: "An organism is essentially nothing but a collection of atoms and molecules" (Crick, 1966); "There is just one science, physics: the rest is just social work" (Watson from Rose, 1997); "The highest object at which the natural sciences are constrained to aim, but which they will never reach, is … in one word, the reduction of all the phenomena of nature

to mechanics" (Kirchhoff from Nagel, 1979). Such statements have the connotation that life is devaluated to atoms and molecules. There are no such things as living atoms or molecules. Wimsatt refers to such statements as examples of "vulgar reductionism" (Wimsatt, 1997). In the end, very few biologists, if any, adhere to such ultimate forms of reductionism in practice, although they may still expect that reduction of biology to physics will occur in due time.

A large group of molecular biologists and biochemists work on the level of isolated subsystems (e.g. enzymes, reconstituted subsystems, or organelles), and tacitly assume that knowing the properties of such parts in isolation is sufficient for understanding them when they are embedded in living systems; for instance, knowing the (kinetic) enzyme properties *in vitro* would entail understanding of its *in vivo* functioning in the cell. Accordingly, such reductionists do not emphasize the need for reconstruction of the system or for appreciation of the properties of the cytoplasm in any way, and to them "the cell is just a bag of enzymes" (Mathews, 1993). In the extreme, this point of view amounts to the following claims: (1) each biological function is coupled to a single enzyme—genetic determinism—(Morange, 2001), (2) *qualitative* information on biochemical systems is sufficient for their understanding, and (3) there are no (emergent) system properties. Apparently, these scientists advocate reductive elimination: higher-level theories (concepts) are to be reduced to and replaced by lower-level theories (concepts). So, in the end, cell biology (physiology) should be replaced by biochemistry (enzymology).

Another group of biologists consider reductionism to be a methodological concept; reduction then is a kind of mechanistic explanation that does by no means entail the elimination of a higher-level description of a certain phenomenon. On the contrary, the ontological security of higher-level phenomena and concepts is actually reinforced whenever they appear in a theory that shows how those phenomena and concepts result from networks of constitutive, lower-level processes. For example, physiological reasoning might indicate why a specific metabolic pathway has certain regulatory features (Hofmeyr *et al.*, 2000; Monod, 1972; Rose, 1997).

Frequently, a wider perspective (i.e. wider than merely reductionist) is needed; many phenomena will only become understandable or observable when perceived in a wider context, e.g. the (whole) systemic environment.

2.2. Anti-reductionist strategies

Like reductionism, anti-reductionism starts with the notion that living organisms are hierarchically organized on various levels. To be sure, all levels have to comply with the laws of physics and chemistry; there is no supramaterial essence; holism should not be confused with vitalism (*élan vital*, *entelechy*, *Bildungstrieb*). Contrary to the reductionist, however, the anti-reductionist emphasizes the role of functional (not necessarily causal) and integrative explanations; explanations in which parts perform some function in a surrounding system in their systemic context. Thus, for the anti-reductionist cell biologist, the cell is more than and different from the sum of the properties of its enzymes (cf. Nagel, 1979, on the various meanings of similar

holistic slogans). This implies that properties on the level of the whole cell are not simply predictable from the properties of the cellular constituents.

The (emergent) systemic properties of a biosystem are rooted in the non-linear interactions between its constituents and are quantitatively dependent on the physiological state. Whereas in the reductionist view emergent properties are nothing but illusions that will vanish if one looks deeper into the phenomenon, they really exist in the view of the anti-reductionist. Therefore, anti-reductionist methods stress the necessity of integrating all (experimental) data concerning the processes that constitute the living cell. However, to resynthesize the whole (cell) from its parts (enzymes and metabolites) and their interactions is by no means straightforward. Somehow, the collective behavior of the parts in the coordinating cell and the physicochemical properties of the cytoplasm need to be taken into account (organicism; see Gilbert & Sarkar, 2000). To put it differently, since the state of a subsystem, and therefore its functional capacity, depends upon the state of other subsystems it interacts with (through direct interactions and possibly even through their effect on the physicochemical milieu of the cytoplasm), subsystems should be investigated in their "natural habitat." Because of these (non-linear) interactions, the properties of a subsystem in isolation do not show up necessarily or to the same degree in its properties *in vivo*. This is an essential element of BioComplexity, and one that is deeply rooted in non-linearity, as will be illustrated below through our examples of BioComplexity.

Also, anti-reductionism represents a continuum that ranges from mild to extreme forms. Biologists with strong anti-reductionist convictions will maintain that research should take place only at the level of the cell. In their arguments, they mostly refer to the active (downward causal) role played by the organization of the cell "soup" (e.g. cytoskeleton, organelles) in determining constituent process properties (Srere, 2000; Welch, 1992, and references therein). Accordingly, all properties measured in subsystems in isolation from the cell are considered incorrect (e.g. all *in vitro* enzyme properties are necessarily different *in vivo*). Thus, the emergent properties of the cell as a whole are considered irreducible to and unpredictable from the properties of the cellular constituents. Consequently, only non-invasive methods that leave the entire cell intact are thought to yield reliable experimental data about cell functioning.

Other less radical anti-reductionists adhere to a form of anti-reductionism where the subsystems operate semi-autonomously—coordination by the whole system not necessarily being dictatorial. In their view, emergent properties are not necessarily irreducible to the *in vitro* properties of lower-level components, provided that all the interactions between these components are taken into account. According to them, some values of *in vitro* measured properties might be different from their respective *in vivo* values, but not all. In other words, as a first approximation one can assume that *in vitro* properties do not differ (significantly) from their *in vivo* counterparts, and subsequently experimentally test this.

Testing can be done in two ways: via *in vitro* reconstitution or *in silico* reconstitution. In the former method, the subsystems isolated from the living cell are put together again to yield an *in vitro* system at a higher level. For example, after

isolation and characterization of the glycolytic enzymes they can be added back to a test tube under controlled conditions to reconstitute the glycolytic pathway *in vitro*. By comparing reconstituted glycolysis with *in vivo* glycolysis, one can determine whether the properties of the enzymes as measured in isolation together suffice to yield the properties of *in vivo* glycolysis. If not, apparently some essential interactions are lost during isolation. Then one can investigate systematically what should be added to the tube to arrive at a fully reconstituted glycolytic pathway that mimics its *in vivo* behavior.

The above experimental strategy is sound and promising, but it is also laborious and not easy to accomplish. *In silico* reconstitution of the subsystem at hand then serves not only as an attractive alternative, but also provides a means to go back and forth between the *in vitro* and the *in silico* subsystems. Moreover, many cellular subsystems cannot be reconstituted *in vitro* at all. In those cases, *in silico* modeling is the only option to study such subsystems.

In silico replicas of biochemical systems indicate that, at least for some systems, *in vitro* properties are sufficient to explain *in vivo* behavior (Bakker *et al.*, 1997; Kremling & Gilles, 2001; Kremling *et al.*, 2001; Rohwer *et al.*, 2000; Teusink *et al.*, 2000; http://www.jjj.bio.vu.nl/). In this view emergent properties are not claimed to be irreducible in all cases; these properties just emerge from the interactions between the components and, therefore, can be explained mechanistically and obtained through detailed calculations of *in silico* replicas (Westerhoff, 2001).

In the next section, we will illustrate the problems encountered in studying biochemical systems by adopting pure reductionist or pure holistic viewpoints.

3. Is reductionism or anti-reductionism sufficient for explaining BioComplexity?

3.1. The living cell—a network of biochemical processes displaced from equilibrium

The living cell is the smallest unit of life. Living systems are (semi-)open systems that prevail in non-equilibrium states; life is unsustainable in thermodynamic equilibrium, as Schrödinger already noted in his monograph *What is life?* in 1945 (Schrödinger, 1992). To keep itself removed from thermodynamic equilibrium, any system requires a continuous dissipation of Gibbs free energy. Through metabolism of nutrients, organisms continuously "feed on negentropy" as dictated by the second law of thermodynamics (Schrödinger, 1992; Westerhoff & van Dam, 1987). For the proliferative and synthetic aspects that are so characteristic for life, this continuous dissipation of Gibbs free energy does not suffice, however. Organized life depends on the coupling of Gibbs free-energy harvesting and free-energy captivating processes. And this coupling requires intricate gadgets, developed in biological evolution, at the level of complex macromolecules such as the H^+-ATPase, or at the network level such as in the chemiosmotic coupling mechanism (Westerhoff & van Dam, 1987). If, however, the two coupled processes were suddenly decoupled, the Gibbs free-energy captivating process would immediately start to dissipate Gibbs free energy as well. The latter is again dictated by the second law of thermodynam-

ics. Now that we have established that a living cell is essentially a network of partly coupled enzyme-catalyzed and transport processes, we can start trying to understand the required gadgets.

The reductionist biochemical research programs have been, and still are, vital for disentangling the network of biochemical processes. Chopping up the network in parts and investigating the subsystems in isolation have provided us with a data set containing many of the physicochemical properties of the constituent processes of a particular organism. Such *in vitro* data sets contain, for example, the types of enzymes, the affinities of enzymes for their substrates, products and effectors, the equilibrium constants of the reactions, and the diffusion constants for all (macro)molecules. This set of data is already rather large and is expected to encompass enzymological knowledge about all cellular processes in due time. On the other hand, this data set is many orders of magnitude smaller that the complete data set of biology, which should one day include all spatial coordinates of all amino acids in all proteins. The next question that surfaces is: "What can we do with such a physicochemical data set?"

3.2. In silico *replicas of living systems*

The answer is that it is sufficient for an *in silico* reconstruction of all mass flow, i.e. conversion and transport of (macro)molecules, that takes place in a particular (unicellular) organism or in any part of that cell (Endy & Brent, 2001; Heinrich & Rapoport, 1977; Westerhoff, 2000, 2001). More information on the description of living systems (subsystems) as a large network of mass flow and its relationship with the physicochemical data sets is given in Table 1.

In other words, in principle the set of differential equations (Eq. 1, Table 1), describing the change in the concentration of all metabolites in a cell with time as a function of all the rates of processes the metabolites are engaged in, can be constructed solely from data derived from reductionist methods. Thus, Eq. 1 constitutes the first step towards an *in silico* replica of a particular organism, say, the enterobacterium *Escherichia coli*. Next, one may wonder to what extent this *in silico* version of *E. coli*, reconstructed from the results obtained with *in vitro* reductionist methods, mimics the behavior of its organic counterpart. Would the *in silico* behavior be identical to the *in vivo* behavior of *E. coli*?

The answer is presumably "no." The replica could prove to be incorrect for a number of reasons having to do with the tacit assumption, made by most biochemists and molecular biologists who work *in vitro*, that a cell is just a bag of enzymes. Here, we will elaborate on two important reasons, which are supported by experimental evidence, why such an assumption might prove wrong: the existence of (1) macromolecular crowding (Ellis, 2001a,b) and (2) channeling (cf. Ovadi, 1995; Welch, 1986). Both phenomena are arguments in favor of perceiving living cells as integrative systems (see Section 3.3). Macromolecular crowding refers to the phenomenon that association frequencies of specific macromolecules (say A and B) increase with an increasing concentration of other (macro)molecules (... X, Y, Z; see Zimmerman & Minton, 1993). Since the macromolecules *in vivo* occupy some 30%

TABLE 1. Mathematical description of mass flow in living systems

In living systems, all processes take place either in (or at) membranes or in an aqueous environment (cytoplasm). Assuming that diffusion of (macro)molecules is sufficiently fast, such that the cell is essentially homogenously "stirred," a concise description of the time evolution of all m (macro)molecules, $X_i (i = 1..m)$, in a cytoplasm with r reactions, i.e. processes, is given by the following equation (Heinrich & Schuster, 1996):

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}\,(t,\mathbf{p})_{mx1} = \mathbf{N}_{mxr}\cdot\mathbf{v}(\mathbf{x}(t,\mathbf{p}),\mathbf{p})_{rx1}$$

where bold symbols denote matrices and their subscripts indicate the number of rows and columns, e.g. mxr. The vector $\mathbf{x}(t, \mathbf{p})$ is the mx1 metabolite vector constituting concentrations of enzymes, metabolites, mRNA, transcription factors, and the like. The mxr-matrix $\mathbf{N}$ is the so-called stoichiometry matrix which has as its (i,j) entry the stoichiometric coefficient of the ith metabolite of $\mathbf{x}(t, \mathbf{p})$ in the jth reaction of the rate vector $\mathbf{v}(\mathbf{x}(t, \mathbf{p}), \mathbf{p}$. The $r \times 1$ rate vector $\mathbf{v}(\mathbf{x}(t, \mathbf{p}), \mathbf{p})$ constitutes the rates or velocities of the r reactions, e.g. enzyme catalyzed conversions or transport and their units are concentration per unit time. The entries in $\mathbf{v}(\mathbf{x}(t, \mathbf{p}), \mathbf{p})$ are rate equations which are generally non-linear functions of (some) metabolite concentrations, e.g. for the ith entry of $\mathbf{v}(\mathbf{x}(t, \mathbf{p}), \mathbf{p})$,

$$v_i = \frac{k_{cat,i}\cdot X_i}{1+\dfrac{X_{12}}{K_{i,12}}}\cdot\frac{\dfrac{X_{62}}{K_{i,62}}\cdot\dfrac{X_{15}}{K_{i,15}}\cdot\left(1-\dfrac{X_{73}\cdot X_{22}}{K_{eq,i}}\right)}{\left(1+\dfrac{X_{62}}{K_{i,62}}+\dfrac{X_{73}}{K_{i,73}}\right)\cdot\left(1+\dfrac{X_{15}}{K_{i,15}}+\dfrac{X_{22}}{K_{i,15}}\right)}$$

This equation describes the rate of the reaction $X_{62} + X_{15} \overset{v_i}{\longleftrightarrow} X_{73} + X_{22}$ catalyzed by enzyme i with concentration X_i which is a function of the concentration of the 12th metabolite (X_{12}) in $\mathbf{x}(t, \mathbf{p})$ and the physicochemical parameters: $K_{i,12}$, $K_{i,62}$, $K_{i,15}$, $K_{i,73}$, $K_{i,22}$, $K_{eq,i}$, and $k_{cat,i}$. Both the precise form of the rate equation and the magnitudes of its physicochemical parameters result from the reductionist methods of the biochemist, i.e. determination of enzyme kinetic mechanism in isolation. The identity of the enzyme-catalyzed processes in $\mathbf{v}(\mathbf{x}(t, \mathbf{p}), \mathbf{p})$ results from reductionist methods assisted by metabolic network reconstruction methods from genome nucleotide data (Covert *et al.*, 2001).

of the cell volume and *in vitro* experiments are mostly carried out under diluted conditions, the implication is that the *in vitro* measured process properties may differ from their *in vivo* counterparts (Ellis, 2001a,b). Channeling is the phenomenon that some enzymes directly transfer—channel—their product to a second enzyme to serve as a substrate (Ovadi, 1995). Consequently, these (macro)molecular substrates and products do not enter the bulk solution—usually the cytoplasm—as unbound metabolites. This way, the *in vivo* kinetic mechanism of an enzyme may be different from its kinetic behavior *in vitro*, since the enzyme *in vivo* can no longer be thought of as independent of the other enzymes with which it has a channeling relationship. To make it even more *biocomplex*, the two phenomena may influence each other as well.

Just these two examples already indicate that there might be a substantial difference between the observed *in vitro* behavior of a part outside the cell and the *in vivo* behavior of the same part inside the cell. If so, it would considerably affect the degree to which the *in silico* version mimics the behavior of *E. coli*. The issue of the difference between results obtained *in vitro* and *in vivo* is highly relevant and asks

for experimental tests to assess the degree of difference and for quantitative *in vivo* experimentation.

The phosphotransferase system (PTS) in *E. coli* may serve as an example here (Postma *et al.*, 1993; Rohwer *et al.*, 1998, 2000). It consists of four enzymes that transport glucose from the environment across the plasma membrane into the interior of the cell and concomitantly phosphorylate glucose (at the expense of phosphoenolpyruvate), yielding glucose-6-phosphate. If all four PTS enzymes operate independently, i.e. do not function through channeling, then *in vitro* kinetic studies of the individual enzymes would be sufficient; the *in silico* replica of PTS would reproduce the *in vivo* functioning of PTS. However, detailed biochemical experimentation has shown that PTS enzymes directly transfer their products. Hence an *in silico* replica of the PTS-system reconstructed from (isolated) enzyme kinetic data would most likely result in fluxes that are inconsistent with their *in vivo* values. Additionally, the channeling properties of the system, mechanistically encoded in the kinetic properties of the PTS enzymes, might depend upon the total (or local) concentration of macromolecules in *E. coli* (i.e. upon macromolecular crowding; see Rohwer *et al.*, 1998). At present, we are experimentally determining the extent of the contribution of channeling and crowding to the functioning of the PTS-system (Rohwer *et al.*, 1998; C. Francke, personal communication).

The PTS-system indicates that reductionist methods *per se* do not suffice for a mechanistic explanation of the "live" state. In explanatory pluralism, the importance of anti-reductionist strategies is acknowledged, as is experimental testing of the significance of the *in vivo* environment.

3.3. Identification of the type of biosystem

Part of the continuing dispute among biologists on the basis of their differing (anti-)reductionist methodologies (Bock & Goode, 1998) may well originate from a lack of appreciation in contemporary biology as to the type of physicochemical organization underlying living matter. We will illustrate this using Bechtel and Richardson's (1993) terminology on types of systems. These authors discerned two types of system—*aggregative* and *composit*e systems.

System properties of aggregative systems depend linearly on component properties (Bechtel & Richardson, 1993; Wimsatt, 1997). Or, to put it differently, the organization of the system is *not* a major determinant of system functioning (Wimsatt, 1997). This implies that *in vitro* analysis of isolated components should suffice to explain system behavior. However, as fundamental as it may seem when considering the conservation laws (concerning mass, energy, and momentum) as formulated in non-relativistic physics, linearity is an exception in biology (Wimsatt, 1997).

Composite systems come in two forms—*component* and *integrative* systems (Bechtel & Richardson, 1993). The organization of a component system does not affect the properties of the components. In other words, if a typical unicellular organism is a composite system, knowledge of (1) the kinetic properties of its enzymes obtained through *in vitro* methods (e.g. maximal rates and affinities for their substrates, products, and effectors) and (2) its regulatory and mass-flow

network topology should suffice for a mechanistic explanation of its system properties. However, since component systems are non-linear, as opposed to aggregative systems, mechanistic explanations of their systemic properties will entail system reconstruction (*in silico* or *in vitro*). On the other hand, in integrative systems, "systemic organization is significantly involved in determining constituent functioning" (Bechtel & Richardson, 1993). This means that through the physicochemical properties of the cytoplasm and the interactions between components, some component properties are altered such that system functioning can no longer be predicted from *in vitro* knowledge of component properties. Component properties that might be affected are, for example, substrate, product, or effector affinities, binding properties for macromolecular complex formation, and diffusion constants for (macro)molecules.

Simon made a similar distinction (Simon, 1969). In his perception of decomposability, component and integrative systems would be, respectively, *nearly decomposable* and *minimally decomposable* (Bechtel & Richardson, 1993; Simon, 1969). A nearly decomposable system is one that can be almost completely understood through application of the *in vitro* determined component properties, whereas a system is only minimally decomposable when *in vitro* determined properties are only applicable to a limited degree (see also Bechtel & Richardson, 1993).

Thus, choosing the appropriate methodology to arrive at mechanistic explanations of phenomena of complex living systems invariably depends upon whether such systems are aggregative, component, or integrative systems. For instance, to understand the latter systems, primarily *in vivo* methods will be required; alternatively, living systems might prove to be integrative only with respect to some, not all, of their characteristics. In the next sections, we will discuss this problem more thoroughly and contend that biochemical systems are integrative (minimally decomposable) systems at least regarding some system properties.

Figure 1 visualizes in a qualitative and speculative manner the contributions of reductionist and anti-reductionist methodologies to the mechanistic explanation of the "live" state as a function of the degree of BioComplexity. It also illustrates the notions of linearity and non-linearity, and of aggregative, component and integrative systems in their relations to cell biological phenomena. The figure is meant to suggest that the contribution of a reductionist or an anti-reductionist research program to the explanation of a particular biocomplex phenomenon progressively decreases or increases, respectively, with increasing BioComplexity. To clarify the meaning of the degree of BioComplexity, a number of biological phenomena are ranked in order of increasing BioComplexity: (1) enzyme catalysis, (2) metabolic pathway, (3) macromolecular crowding, (4) channeling, (5) metabolons, (6) local regulation and control, (7) local adaptation and plasticity, (8) global coordination of transcription and translation, (9) cell division, (10) global regulation and control, and (11) global adaptation and plasticity.

Figure 1 indicates that enzyme catalysis and metabolic pathways are classified as component systems. Kinetic descriptions of individual enzymes (Eq. 2, Table 1) are classically determined under *in vitro* conditions. Descriptions of metabolic pathways (Eq. 1, Table 1) based on *in vitro* determined kinetics of individual

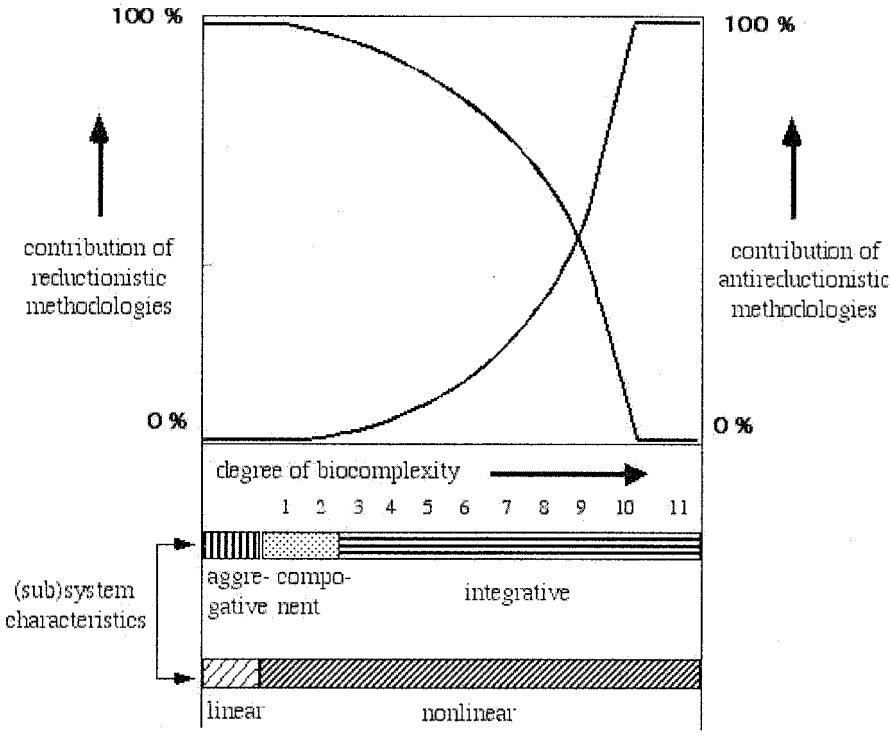

FIG. 1: The contribution of reductionist and anti-reductionist methodologies to the explanation of increasing biocomplex phenomena. The lower x-axis depicts the type of process (linear or non-linear). The middle x-axis gives the type of system the processes are part of (aggregative, component or integrative). The upper x-axis ranks the degree of BioComplexity. The left and right y-axis depict the contribution (in percentages) of reductionist and anti-reductionist methodologies, respectively, to a BioComplexity research program that aims at the mechanistic explanation of a biological phenomenon characterized by a certain degree of BioComplexity. Numbers 1–11: biological phenomena (see text) ranked in order of increasing BioComplexity. Although the order of the phenomena is meaningful, the exact position on the x-axis is more or less arbitrary; the listed phenomena only serve to give an indication about what it meant by increasing BioComplexity.

enzymes do not take into account biocomplex phenomena. Phenomena such as channeling and macromolecular crowding can alter kinetic properties of enzymes and hereby render the system integrative. Appreciation of BioComplexity necessitates enzyme kinetic studies with enzymes embedded in their natural habitat, i.e. the metabolic pathway in the cytoplasm. Metabolic pathways with channeling (and possibly macromolecular crowding) have been termed metabolons (Srere, 2000). Fluxes through metabolons are regulated and controlled locally through multiple mechanisms that affect the activity of the constituent enzymes, e.g. through gene expression and changes in levels of effector metabolites (see section on metabolic control analysis below). Regulation and control of fluxes and concentrations can lead to emergent properties of biochemical networks enabling them to respond adequately to changes in their physicochemical environment (e.g. plasticity through adaptation). Global coordination of transcription and translation is mediated by

transcriptional regulators that affect transcription of large sets of enzymes involved in a variety of cellular functions. Cell division and its coordination is a globally regulated orchestration of a large number of enzyme-catalyzed processes. As a whole, global regulation can lead to major physiological adaptive decisions—global plasticity of individual cells or populations (e.g. cell dormitivity through sporulation and programmed cell death).

4. Multiple intra-level and inter-level theories

4.1. Explanatory pluralism

In the preceding sections, we discussed some of the limitations of reductionist and anti-reductionist methodologies in explaining the "live" state. We argued that the type of (physicochemical) organizational structure of (unicellular) life would determine the success and failure of these strategies. Yet three general observations point to the importance of integrative aspects of biosystems: (1) they show non-linear dynamics—non-linearity seems to be the rule rather than the exception, (2) they comprise a democratic control structure—there is no genomic (or metabolic) supremacy over other cellular processes, and (3) they possess a "unique" intracellular environment—some, perhaps many, *in vivo* properties are not observable or quantitatively describable *in vitro*. These characteristics are still under extensive scientific investigation, but they may play a major role in determining the functioning of living organisms. In each individual case, it should be experimentally confirmed to what extent reductionist methods are sufficient for understanding the *in vivo* functioning of a particular biosystem, since this might depend on the phenomenon and/or the type of species (or tissue). This analysis, although certainly not exhaustive, suggests that if one wants to understand (mechanistically) how (unicellular) life emerges from its constituent processes, reductionist and anti-reductionist methods should go hand in hand; a pluralist approach is required.

Thus, our pluralist research strategy attempts to adopt the best of both reductionism and anti-reductionism. Accordingly, we agree on the anti-reductionist dogma that levels of organization are ontologically all of value, i.e. there is no dictatorial but rather democratic control of processes at higher levels on processes on lower levels. We may not agree however that they are always of precisely equal value. In fact, we have methods available that can quantify the relative importance of the different levels experimentally (ter Kuile & Westerhoff, 2001; Snoep *et al.*, 2002). Intra-level theories and methodologies may and should be developed autonomously at every level. They will shed light on the functioning of a cell at a given level and on its relationship with its immediate neighboring levels (Westerhoff *et al.*, 1990). Nevertheless, the construction of inter-level theories, and not eliminative theory reduction, yields the most promising and rewarding perspective linking adjacent levels of cellular organization, e.g. metabolic pathways and enzymes. In this respect, reductionism and anti-reductionism are relative; they should be related to a certain given level.

To illustrate our point of explanatory pluralism, we will discuss how multiple

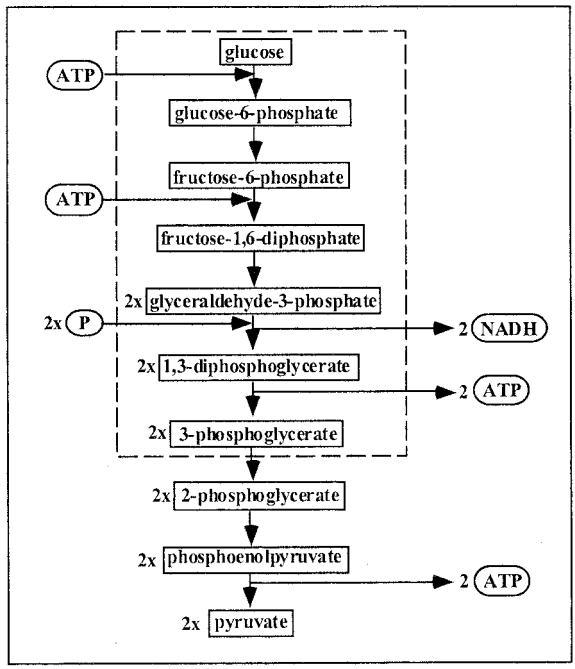

FIG. 2: Scheme of the glycolytic pathway. The sequential conversion of glucose into pyruvate is shown in a simplified version of glycolysis. Usually, all reactions take place in the cytosol. Glycolysis in *Trypanosoma brucei* forms an exception to this rule: the dashed box indicates the part of glycolysis that occurs in the glycosome in this parasite (NB glucose is also present in the cytosol).

intra-level and inter-level theories/methodologies are contributing to the current understanding of glycolysis in yeast (*Saccharomyces cerevisiae*) and the parasite that causes sleeping sickness (*Trypanosoma brucei*). In the cellular hierarchy, glycolysis belongs to the level of metabolic pathways whereas its catalytic constituents belong to the (lower) level of macromolecules. Glycolysis is a metabolic pathway that oxidizes glycolysis into pyruvate in a stepwise manner (Figure 2).

Intra-level research on both levels and the construction of inter-level theories have proven useful in understanding the properties of glycolysis. Table 2 displays typical examples of intra-level and inter-level theories/methodologies that have been applied in understanding glycolysis. Since intra-level and, especially, inter-level theories and methodologies play such an important role in our argument, we will pay more attention to them in the next sections.

4.2. Intra-level theories and methodologies

At the level of enzymes, the properties of the glycolytic enzymes are investigated with methods from structural biology and enzymology. Structural biology yields information on the 3D structure of enzymes as inferred from crystallographical, nuclear magnetic resonance, and electron-microscopical measurements and from computer-based 3D-structure modeling. On the other hand, enzymology provides us with the

TABLE 2. Pluralistic research on glycolysis: examples of intra-level and inter-level theories and methodologies that have contributed to the present understanding of the glycolytic enzymes, of glycolysis as a whole, and of their mutual dependence

Intra-level theories/methodologies		Inter-level theories/methodologies
Glycolytic enzymes	Glycolysis	Glycolytic enzymes/glycolysis
Structural biology	Physiology	Mathematical modeling
Enzymology	BioEnergetics	Metabolic control analysis

kinetic data on the catalytic properties of the glycolytic enzymes. For example, V_{max} (the maximal rate of the enzyme-catalyzed reaction) and the Michaelis–Menten constant K_m (the substrate concentration at which an enzyme is saturated for 50%) are useful concepts in enzyme kinetics (a theory describing the rate of an enzyme as a function of the concentrations of substrates, products and effectors). Importantly, such intra-level enzyme concepts do not refer to the properties of the *in vivo* system (glycolysis) the glycolytic enzymes are embedded in. Nevertheless, these physicochemical data are necessary for an *in silico* reconstruction of glycolysis in order to study its systemic properties (see below).

Also, at the level of metabolic pathways, multiple intra-level theories/methodologies can be discerned. For instance, physiology is dealing with the functioning of glycolysis as a whole in cellular catabolism and anabolism. Its aims are to measure the *in vivo* flux through glycolysis, the intracellular concentrations of its intermediate metabolites, the importance of shunts and branches, and the distribution of intermediates among catabolism and anabolism. Bioenergetics is engaged with the role played by glycolysis in redox and free-energy balances of the cell. Its goal is to find out how net free energy is made available (as ATP) and how the glycolytic redox reactions fit in the overall cellular redox balance. These intra-level pathway concepts do not refer to properties of the lower-level constituents (the individual glycolytic enzymes) of the pathway.

4.3. Inter-level theories and methodologies

Inter-level theories/methodologies are becoming increasingly important in (cell) biology now that we understand the physicochemical properties of an increasing number of constituents of biochemical systems. Inter-level theories link properties of systems to properties of their subsystems. Enzymes and metabolic pathways are an example of such a part–whole relationship. Mathematical biochemistry relates properties of metabolic pathways to enzyme properties and hereby facilitates quantitative understanding of biochemical systems (Westerhoff, 2000, 2001). Two related examples, both regarding the so-called "turbo-design" of glycolysis, will illustrate the power of the methodology of kinetic mathematical modeling of pathways such as glycolysis.

5. The inter-level approach

5.1. Mathematical modeling

In most organisms, including the yeast *S. cerevisiae*, the glycolytic enzymes reside in one compartment, the cytoplasm. The "turbo-design" means that the glycolytic pathway is designed such that ATP is invested in the first reactions and a surplus of ATP is harvested in the lower part of glycolysis (Figure 2) in order to speed up the upper part of glycolysis (Teusink *et al.*, 1998). However, with the aid of a mathematical model of glycolysis based upon *in vitro* kinetic data for glycolytic enzymes (and constructed as explained in Table 1) it was shown that the "turbo-design" also contains an inherent risk. If the ATP-consuming first part is not tightly regulated with respect to the ATP-producing second part, glycolytic intermediates may accumulate to levels that are detrimental to cellular functioning. So the function of the tight regulation of the first part of glycolysis is explained by showing that cell functioning would go awry in case this regulation is absent or deficient.

A related example concerns glycolysis in the trypanosome *Trypanosoma brucei*, the parasite that causes African sleeping disease. Here, the first part of glycolysis takes place in a specialized organelle, the glycosome (Figure 2). Also, here a mathematical model has been constructed from *in vitro* determined kinetic enzyme properties of glycolysis in *T. brucei* (Bakker *et al.*, 1997). A realistic replica is based upon precise kinetic knowledge of the molecular components. This is a special type of computer modeling in which an *in silico* experiment can be done as a substitute for an *in vivo* experiment that is not feasible (removing the glycosomal membrane). First, the replica was shown to mimic the relevant *in vivo* behavior of the biosystem under study. Then the "impossible" experiment was performed *in silico* and the results were compared to the *in silico* experiment with the "wild type." In this way, it was tentatively established that the function of compartmentation is to protect the parasite from the negative effects of the "*turbo*" design of glycolysis (Bakker *et al.*, 2000).

Mathematical biochemistry has also resulted in a number of important inter-level theories/concepts. These include: (1) metabolic control analysis (Fell, 1997; Heinrich & Rapoport, 1974; Kacser & Burns, 1973; Westerhoff & van Dam, 1987) and (2) flux analysis (Schilling *et al.*, 1999, and references therein). Metabolic control analysis (MCA) relates properties of enzymes to changes in fluxes through metabolic pathways and metabolite concentrations upon changes in particular parameters. MCA bases the distribution of control of biological functioning in enzyme properties. Flux analysis studies the distribution of (steady-state) fluxes through different branches of steady-state biochemical networks. So far, it culminated in the concept of elementary modes (Schuster *et al.*, 2000). In the next section, we will elaborate on MCA.

5.2. Metabolic control analysis

5.2.1. Local elasticity and global control. MCA embodies a rigorous mathematical framework for studies of control and regulation of biochemical systems (Fell, 1997;

Heinrich & Rapoport, 1974; Kacser & Burns, 1973; Westerhoff & van Dam, 1987). MCA relates changes in systemic properties of the biochemical network—metabolite concentrations, flux or any function thereof—to the causal changes in physicochemical properties of their components, e.g. enzymes or modules of multiple enzymes. The component (local) property is referred to as elasticity coefficient (notation: $\varepsilon_{X_i}^{v_j}$ for the elasticity of the (enzymatic) process v_j to an effector metabolite X_i). It is defined as the immediate fractional change in the rate of an enzyme-catalyzed reaction upon a concentration change in only one of its effector metabolites while the whole network remains in a steady state. A large elasticity coefficient indicates that this particular process changes its rate to a large extent upon a change in the concentration of only X_i. Conversely, a small elasticity coefficient indicates that the process is "saturated" with X_i.

The systemic (global) properties are referred to as the control coefficients of a particular process, e.g. an enzyme-catalyzed reaction, on any metabolite, flux or function thereof. A flux control coefficient is defined as the fractional change in steady-state flux, after the whole system has relaxed to a new steady state of reference, upon a fractional change in enzyme activity (concentration). MCA has been derived for steady and oscillating reference states of biochemical networks (Demin *et al.*, 1999; Heinrich & Schuster, 1996; Kholodenko *et al.*, 1997; Reijenga *et al.*, 2002). It is a mathematical framework that connects two adjacent levels in the biological organizational hierarchy—the level of the metabolic pathway to the level of the constituent processes (e.g. enzyme-catalyzed reactions and transport)—which makes it an inter-level theory.

In modular control analysis, an extension of MCA, an analogous mathematical connection is made between a large biochemical network and the metabolic pathways it consists of. The pathways composed of multiple enzymes (modules) are treated here as if they were single enzymes (Schuster *et al.*, 1993).

5.2.2. Control is a systemic property. Now, we will illustrate how one can calculate the change in systemic properties in terms of the underlying changes of componential properties for the simple linear metabolic pathway depicted in Figure 3A. Assuming that the network operates in a (structurally stable) steady state, i.e. $v_1 = v_2$ (which equals the steady-state pathway flux $\mathcal{J}$), it will relax to a new steady state upon a (infinitesimally) small perturbation in the activity (or equivalently concentration) of the first enzyme. The new steady state differs from the initial reference steady state such that the concentration of X_1 and the flux $\mathcal{J}$ are slightly perturbed. The difference between the two steady states in terms of flux is dictated by the control coefficients of MCA through application of the summation (here: $C_1^{\mathcal{J}} + C_2^{\mathcal{J}} = 1$) and connectivity (here: $C_1^{\mathcal{J}} \cdot \varepsilon_x^1 + C_2^{\mathcal{J}} \cdot \varepsilon_x^2 = 0$) theorems for flux control coefficients, which yields in this case for enzyme 1:

$$C_1^{\mathcal{J}} = \frac{1}{1 - \dfrac{\varepsilon_{X_1}^1}{\varepsilon_{X_1}^2}} \qquad [1]$$

Equation 1 illustrates the important notion that a system property (the control of an enzyme on the flux in the steady state) can indeed be expressed solely in terms of

subsystem properties (kinetic properties of the enzymes constituting the system). Moreover, Eq. 1 also shows that the flux control coefficient of the first enzyme is not only dependent upon the kinetic properties of this enzyme (as expected) but also on the kinetic properties of the second enzyme (unexpectedly, perhaps). As a matter of fact, its value depends on the ratio of the elasticities of enzyme 1 and enzyme 2 with respect to X_1, which are (often) <0 and >0, respectively. Thus, $C_1^{\mathcal{J}}$ is a monotonously decreasing function of $-\frac{\varepsilon_{X_1}^1}{\varepsilon_{X_1}^2}$ for as long as the parameters of the two enzymes remain unchanged. If, however, enzyme or environmental parameters are changed somehow, then elasticities will change as well and consequently the control distribution among the two enzymes may also change.

Before the development of MCA, physiologists and biochemists used to discuss flux control from the perspective of "the rate-limiting step"; implying that one can always find an enzyme, say enzyme i, in a metabolic pathway consisting of j enzymes that has all control of the flux (i.e. $C_i^{\mathcal{J}} = 1$ and $C_{j \neq i}^{\mathcal{J}} = 0$), irrespective of the steady state. MCA, however, showed that "the rate-limiting step" has no *a priori* validity. It may exist but only when the elasticities assume special values. In practice, it is an exception rather than the rule and the control of flux is usually distributed among all pathway enzymes.

From a philosophical point of view, Eq. 1 also indicates that a certain value of the flux control coefficient of enzyme 1 is multiply realizable. In other words, the same value for the flux control coefficient can be obtained for different values of the elasticity coefficients as long as the ratio of the latter coefficients does not change. Hence, in principle the first enzyme can exert the same systemic control on the flux through the pathway when the physicochemical properties of both enzymes are altered in proportion.

5.2.3. Subsystem control distribution changes upon embedment. The control exerted by an enzyme on the flux through a pathway changes when this pathway is embedded in a larger metabolic network. We shall use this change in control to illustrate the *in vitro* versus the *in vivo* behavior of the metabolic pathway at hand (Figure 3A). Figure 3B shows another simple pathway that consists of three enzymes instead of two and two variable intermediates instead of one. In Figure 3C the metabolic pathways from Figures 3A and 3B are combined to form a larger network through the regulatory effects of X_1 and X_3 on the rates of enzyme 3 and enzyme 1, respectively. The control exerted by enzyme 1 on its steady-state flux change (i.e. on $\mathcal{J} = v_{1SS} = v_{2SS}$) changes upon embedment in the larger network. The extent of change can be determined with modular control analysis (Bruggeman *et al.*, 2002; Hofmeyr & Westerhoff, 2001; Westerhoff & Kahn, 1993; Westerhoff *et al.*, 1990), and in this case results in the following equation:

$$G_1^{\mathcal{J}} = \left(1 + \frac{1}{1 - r_{X_1}^{X_3} \cdot r_{X_3}^{X_1}} \cdot \frac{1}{C_1^{X_1}}\right) \cdot C_1^{\mathcal{J}} \qquad [2]$$

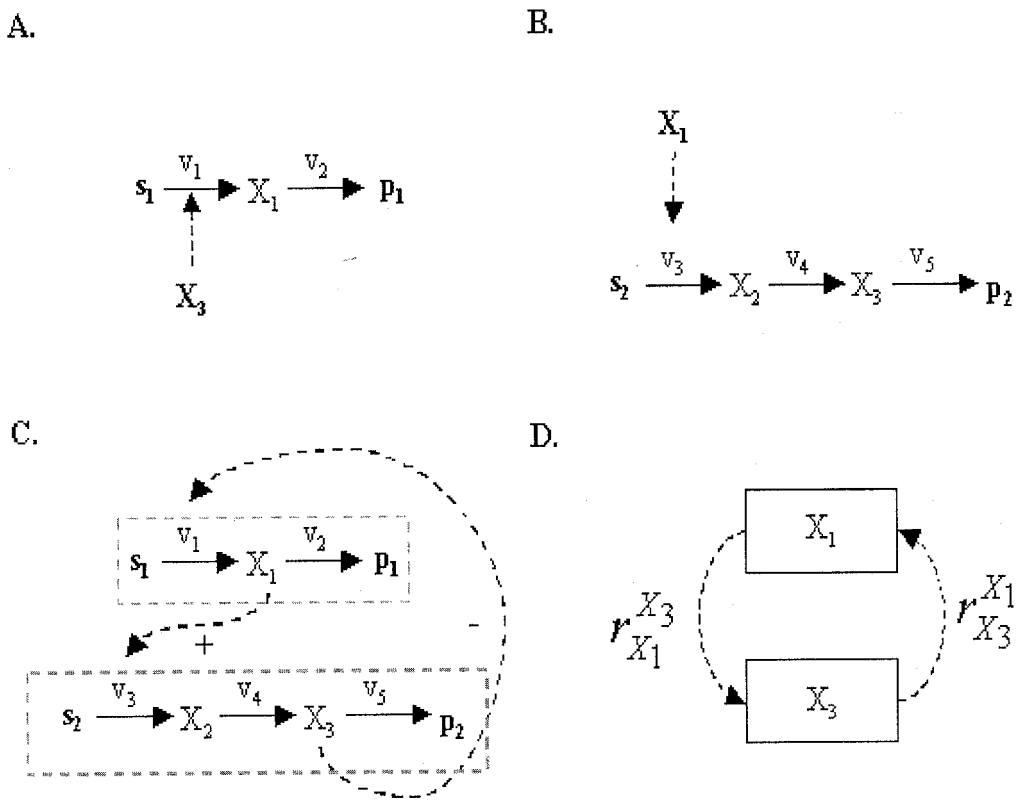

FIG. 3: Four different aspects of a multilevel metabolic network: (A) a linear metabolic pathway that consists of two reactions, operating at rates v_1 and v_2, that produce and consume metabolite X_1; (B) a linear pathway that consists of three sequential reactions, v_1, v_2, and v_3, that convert s_2 into p_2 through the intermediates X_2 an X_3; (C) a metabolic network that consists of the linear metabolic pathways of (A) and (B) (dashed boxes) coupled to each other through the metabolites X_1 and X_3; (D) a modular representation of the biochemical network of (C). The boxes contain the metabolic pathways of (A) and (B) and are denoted by X_1 and X_3, respectively. $r_{X_1}^{X_3}$ and $r_{X_3}^{X_1}$ are inter-modular response coefficients (see text). In (A)–(C) all reaction symbols v_i indicate the rate of reactions i in amount of metabolite converted per unit time. The metabolites in boldface were held constant during the simulations.

$G_1^{\mathcal{J}}$ is the control coefficient of enzyme 1 when pathway 1A is connected to pathway 1B. The coefficients $r_{X_1}^{X_3}$ and $r_{X_3}^{X1}$ are inter-modular response coefficients (Bruggeman *et al.*, 2002; Hofmeyr & Westerhoff, 2001). $C_1^{X_1}$ and $C_1^{\mathcal{J}}$ are control coefficients of enzyme 1 on X_1 and $\mathcal{J}$, respectively, in case the concentrations in the metabolic pathway of Figure 3B are held constant.

Equation 2 illustrates that the control of enzyme 1 on its own steady-state flux changes upon embedment in a larger metabolic network. It is equal to the control coefficient of enzyme 1 in the smaller system multiplied by an extra bracketed term representing the effect of embedding the smaller system in the large network. More specifically, the extra term is due to the change in boundary conditions of both pathways upon embedment. For, in the larger metabolic network X_3 is no longer a parameter (but a variable) for the pathway shown in Figure 3A and the same holds

for X_1 and the pathway depicted in Figure 3B. Alternatively, one may say that the change in control is due to additional interactions and to changes in pre-existing interactions.

6. Emergence

6.1. Oscillating biochemical networks

The aforementioned change in boundary conditions—addition or removal of dynamic interactions—may also result in qualitatively new behavior in a system relative to the behavior of its constituents. This behavior, an example of an emergent property, is shown in Figure 4. If we define the metabolic pathways of Figures 3A and 3B as the constituents of the biochemical network depicted in Figure 3C, oscillations are observed in the total network (Figure 4C). These are not observed in the individual sub-networks (Figures 4A and 4B). Dynamic systems theory on Hopf bifurcations shows that the reductionist approach of treating the individual processes of the system as if in isolation is impotent here; knowledge of the components and stoichiometry only is not sufficient. The condition for the Hopf bifurcation—for the emergence of an oscillation—is determined by the properties of the components, the stoichiometry of the network and the values for all variables under steady-state conditions for the total network [1]. The requirement of complex eigenvalues of the Jacobian matrix at the fixed point can only be met when the system contains at least two independent variables, which implies that it must be a system, not a single process system: system reconstruction is then necessary to calculate the eigenvalues of a matrix that must at least be two by two. To put it differently, at least two degrees of freedom of the system have to be reconstructed and appreciated in a steady-state systemic context. Apparently, after analysis (mechanistic-explanatory studies) follows synthesis (reconstruction of the whole from the parts and their interactions). Parenthetically the analysis need not involve the solution of a *non-linear* system of equations (Nicolis & Prigogine, 1989; Wagner, 1999). Linearized equations can give rise to complex eigenvalues. However, the linearized equations must change as the system moves away from the fixed points, for the system to engage in limit cycle oscillations rather than to explode, and for the system to be consistent with thermodynamics (Cortassa *et al.*, 1991).

Referring to oscillations as emergent properties might come as a surprise, since we just calculated the oscillations from the properties of the component parts and their interactions. Why then do we classify oscillations as emergent properties and *not* merely as system properties? An important prerequisite for calling oscillations an emergent property of the large biochemical network is that the two constituent biochemical pathways do not display oscillations in their respective phase spaces when they are isolated from each other. In addition, its appearance is not simply deducible from the properties of the constituents due to non-linearity; it is necessary to reconstruct the system. If all enzymes obey reversible Michaelis–Menten kinetics,

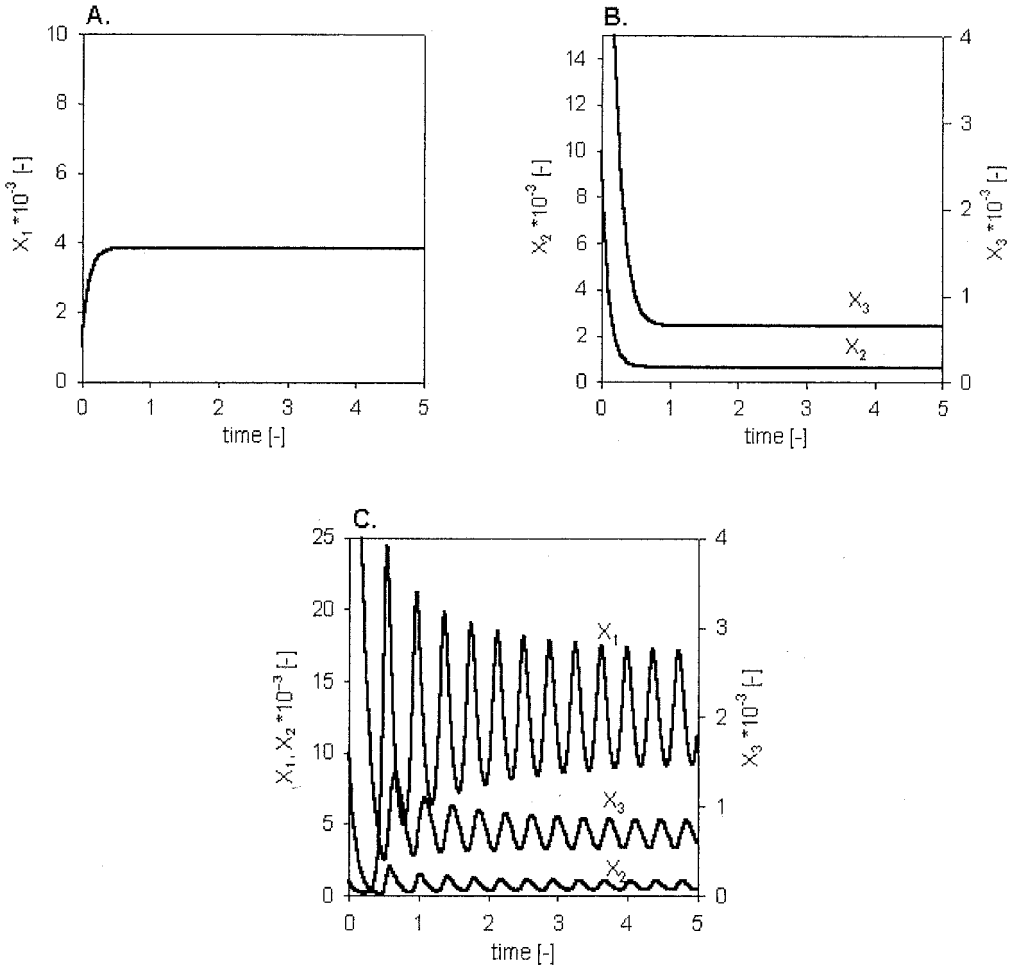

FIG. 4: Oscillations in a metabolic network: (A) concentration of the intermediate metabolite X_1 versus time for the system shown in Figure 3A; (B) concentration of intermediate metabolites X_2 and X_3 versus time for the system shown in Figure 3B; (C) oscillating concentrations of X_1, X_2, and X_3 in the biochemical network depicted in Figure 3C. The system of differential equations can be found in the Appendix.

it can be shown that the two pathways can never display oscillations in isolation (Kholodenko & Denim, 1993; Palsson & Lightfoot, 1985).

Nevertheless, oscillations may be considered a "weak" form of emergence and there exist other more pertinent examples of emergent properties, e.g. macromolecular crowding or channeling effects on enzyme-catalytic process property, as discussed in preceding sections.

6.2. Bacterial signaling networks: phosphoneural networks?

The example of metabolic oscillations illustrated that emergence can be calculated

from mechanistic-biochemical network models, and subsequently traced back to individual properties of enzymes (i.e. sensitivities of enzymes to metabolites), through application of the condition for Hopf bifurcations. However, one might argue that the emergence of metabolic oscillations in biology is just a curious phenomenon that might not be functional in itself. For oscillations, this holds true perhaps (however, this needs to be validated experimentally in each case), but for the autonomous oscillations that drive the cell cycle, function is beyond doubt. In addition, biochemical networks show other types of emergent behavior that may prove highly functional for the cell. The understanding of these phenomena may be important if one wants to mechanistically explain cellular functioning (BioComplexity). We will illustrate this point with the phosphoneural network hypothesis for the bacterium *Escherichia coli* (Hellingwerf *et al.*, 1995).

The phosphoneural network hypothesis entails that the many phosphoryl transfer pathways present in *E. coli* together may function as a neural network (Hellingwerf *et al.*, 1995). A subset of these pathways consists of a particular type of (bacterial) signaling network, a so-called two-component system. A two-component system is composed of a pair of proteins, a sensor and a regulator. The sensor autophosphorylates itself upon receiving a characteristic signal and subsequently transfers the phosphoryl group to its cognate regulator, which in its phosphorylated form engages its regulatory role in the metabolism of the cell. In principle, the prerequisites to comply with a neural network seem to be fulfilled (Hellingwerf *et al.*, 1995): (1) *E. coli* contains 32 such systems, which operate in parallel, (2) some of the two-component systems show the phenomenon of autoamplification, i.e. their (constituents') concentration depends upon previous signaling, and (3) at least *in vitro* they seem to engage in cross-talk. Cross-talk means that a particular sensor phosphorylates a non-cognate regulator.

If a phosphoneural network is indeed operative, then *E. coli* might show the typical characteristic of a neural network, like learning, training, conditioning, LTP, and memory (Bray, 1995; Bruggeman, *et al.*, 2000; Hellingwerf *et al.*, 1995). Intelligence-like behavior (conditioning and learning) was observed *in silico* for the glutamine synthetase regulatory cascade in response to ammonia shortage (Bruggeman *et al.*, 2000). Hoffer *et al.* have experimentally shown that the phosphate-responsive signal transduction network in *E. coli* responded faster and more extensively when it had experienced previous phosphate limitation (Hoffer *et al.*, 2001). Moreover, the shorter response time correlated with autoamplification of the corresponding two-component system. However, careful (*in vivo*) experimentation indicated that cross-talk of two-component signal transduction chains at the level of cross phosphorylation of response regulators does not seem to be significant under physiological conditions (Verhamme *et al.*, 2002). At the physiological level, there are indications of cross-talk between regulons. The mechanisms of some of these may be metabolic or gene-expression rather than signal-transduction-based (Bruggeman *et al.*, 2000; Snoep *et al.*, 2002). More mechanisms remain to be elucidated. This example makes a case for mechanistic studies on emergent properties of biochemical networks *in vivo*.

7. BioComplexity research: bottom–up and top–down analysis of biosystems

7.1. Bottom–up analysis with integrative bioinformatics

Traditionally, cellular regulation was assumed to be essentially dictatorial; it was supposed to be dominated by regulation of gene expression. A linear sequence of information transfer from DNA to mRNA to protein to metabolic function was implied. If this were true, then measuring the transcriptome—the total population of mRNA species of a particular cell—should suffice to unravel regulation at the cellular level. However, studies on glycolysis showed that cellular regulation could be democratic; the glycolytic flux in a number of protists was rarely regulated solely by gene expression (ter Kuile & Westerhoff, 2001). Hierarchical control analysis (HCA), a branch of modular control analysis (see section on modular control analysis) that takes gene expression into account, deals with the distribution of control in such systems (Westerhoff *et al.*, 1998). Accordingly, for a full understanding of glycolysis, the genome, the transcriptome, the proteome, and the metabolome all need to be studied [2]. Ideally, one would like to have a method by which the concentrations of all cellular substances at a certain moment in time and under certain physiological conditions could be measured, i.e. so that $\mathbf{x}(t, \mathbf{p})$ as given in Table 1 could be determined. The acquisition of this data set could be called the aim of X-omics.

Integrative Bioinformatics then should be the discipline that gathers all the kinetic parameters of these cellular constituents, determines their networking, by calculating the above-mentioned functions, and compares the outcome with the results of X-omics. Upon additional experiments suggested by this comparison, the *in silico* model should be improved. Through an iteration of this process, the level of computational "understanding" of the functioning of the live cell should increase. At present, there is no data set that yields a complete reconstruction of the state of a particular (unicellular) organism, but we speculate that such information might become available within 10 years. This should be an enormous achievement in biology, comparable to the discovery of the structure of DNA or the sequencing of the human genome, since this data set allows one finally to reconstruct *in silico* the total state of a living (most likely prokaryotic) organism. Once we have obtained this version of the living cell through a laborious iteration between *in silico* and *in vivo–in vitro* experimentation, life has been explained.

However, will we (namely, not just our computers) then necessarily understand life? Unfortunately, we have merely reconstructed it *in silico*, which was a necessary test of our knowledge. The next step should be to do research on the final *in silico* version itself and one possible route to take would be to profit from its modularity (Brown *et al.*, 1990; Hofmeyr & Cornish-Bowden, 2000; Hofmeyr & Westerhoff, 2001; Westerhoff & van Dam, 1987).

This section aimed at illustrating that the combination of cell physiology (holistic), enzymology (reductionist), and mathematical biochemistry/integrative bioinformatics (inter-level) embodies a necessarily pluralist approach that one day may explain the "live" state. We use the word "necessary" to emphasize that it is the non-linear nature of biosystems that makes every mechanistic explanation a systemic

explanation; to what extent a holistic perspective is required depends on the phenomenon to be explained, but pure reductionist reasoning will not be sufficient.

7.2. Top–down analysis with modular descriptions of cellular networks

Instead of reconstructing the whole "step by step" from its parts (bottom–up), one could also attempt to start from the other end and to use the higher-level perspective to discriminate the relative importance of parts for the cell to function as a whole (top–down). In such a modular perspective, the parts are not enzymes, but instead represent distinguishable portions of cellular metabolism ("modules"), which consist of multiple enzymes and metabolites. Modular control analysis does this for metabolic control analysis. It treats a group of enzymes as if it were a single supra-enzyme, which is allowed when certain conditions are met. With modular control analysis the control of the activity of a certain module with respect to a specific systemic (higher-level) steady-state property—flux, concentration or any function thereof—can be determined (Heinrich & Schuster, 1996; Schuster *et al.*, 1993; Westerhoff & van Dam, 1987). The entire glycolytic pathway might be envisaged as a (multifunctional) module and, for example, its control on the anabolic flux, i.e. growth rate of the cell, may be determined (Koefoed *et al.*, 2002). These modules are probably important also functionally for the cell.

Modular control analysis also gives the mathematical conditions required for a subset of enzymes to operate as a monofunctional unit. Each enzyme in a monofunctional unit has the same co-response coefficient [3] with respect to all fluxes and metabolites outside the unit (Kholodenko *et al.*, 1995; Rohwer *et al.*, 1996). In this sense, modular control analysis is also an inter-level theory with respect to physiology and biochemistry enabling one to question the modular character of glycolysis and to propose alternatives. This then is important for the physiological approach that assumes a specific modular description of the system in its top–down view.

The question how a particular cell should be decomposed into modules has still not been answered definitely. One way would be to use higher-level concepts as a guide. For instance, if one chooses to study free-energy metabolism, the cell could be decomposed into multiple modules that either produce or consume free energy (Westerhoff & van Dam, 1987). Subsequently, the causal role of each module in determining the energy state of the cell could be established. In this respect some interesting studies were done (Ainscow & Brand, 1999a,b; Hofmeyr & Cornish-Bowden, 2000; Jeneson *et al.*, 2000). All these studies indicate that, even though the whole consists of numerous processes, modules can be discerned that perform a certain causal role in the regulation of the free-energy state of cells. This role can be assessed experimentally.

8. System biology: a holistic view on the "live" state

With the advent of X-omics, part of cell biology seems to have changed from a hypothesis-driven (deductive) to an inductive science where the data supposedly speaks for itself (Allen, 2001). At present, bioinformaticians attempt to elucidate

trends in huge data sets gathered with X-omics using sophisticated clustering algorithms. Typical studies involve comparison of large sets of mRNA levels pertaining to two different physiological conditions, e.g. a low and a high level of a particular nutrient or signal. Hereby, subsets of mRNA that portray similar behavior, or subsets of mRNA that respond to only one of the two conditions, are distinguished. Surprisingly perhaps, such studies usually do not apply any existing biochemical knowledge in their data analysis. This type of bioinformatics focuses on only one aspect of living system research, holistic data acquisition.

Some research groups—so-called system biologists (Ideker *et al.*, 2001; Kitano, 2001)—take a more pluralist view on living systems by focusing on both reductionist and holistic information at the same time. System biological studies may be more promising than the "pure number crunching" studies of the bioinformaticians. For instance, Ideker *et al.* investigated a small "local" subsystem (galactose metabolism) in yeast in relation to "global" mRNA expression profiles (Ideker *et al.*, 2001). They then showed how the experimental observations in X-omics and pre-existing knowledge of the system could be "induced" to a more definitive understanding of how the system actually works. Indeed, system biology encompasses more than X-omics and integrative bioinformatics. It stands for integrative studies that combine reductionist molecular-biological and holistic-physiological data by tying them together through quantitative computer modeling (the construction of silicon cell parts). Biology is—even for the simplest organism—still far from mechanistically understanding the working of and interactions among all cellular components *in vitro* or *in vivo*. Yet the aforementioned system biological studies, which endeavor to combine reductionist, holistic and inter-level methods, constitute a great leap forward in explaining the BioComplexity of living systems.

9. Discussion

Most (cell) biologists should be confronted with the same basic problem: "How to isolate a particular biosystem or subsystem from its environment to perform controlled and defined experiments with only a limited number of variables." This *ceteris paribus* clause has culminated in the well-known reductionism–holism debate that from time to time re-emerges. In the heat of the debate, reductionists are described as scientists that deny the importance of systemic relationships, whereas holists would deny the importance of components and their intrinsic properties. Here, we tried to pinpoint the different ingredients of this discussion in biology, namely, that the reductionist and the holist have to cope with different, but mutually dependent, problems, i.e. the reductionist with functional (emergent) system properties and organization of the cytoplasm, the holist with the (experimental and conceptual) power of mechanistic explanations. Importantly, both problems can be tackled experimentally, either in the laboratory or *in silico*. However, the extent of both problems depends upon the physicochemical organization of the living system under study, constituting a component or an integrative system. This embodies the scientific substantiation of the popular notion that the whole is more than the sum of its parts.

There is an increasing awareness among scientists that there are clear limitations to the reductionist program (Bock & Goode, 1998; Nurse, 1997; Williams, 1997). The (re)synthesis of the whole from its parts fails most of the time—synthesis is not simply the reverse of analysis (Anderson, 1972; Wilson, 1999). "One can not recover a chicken from chicken soup" (Kell & Welch, 1991). This became especially evident upon completion of the human genome sequence which indicated that the number of genes in the human genome only differed, for example, from that in the fly genome by a factor of two, despite the obvious difference in human and fly intelligence. This surprisingly small numerical difference illuminates the major shortcoming of genetic determinism (Morange, 2001). Apparently, there is more to biology and presumably even to molecular cell biology than just genes. Consequently, most biologists will stick to more or less moderate forms of reductionism. This point of view is not often explicated or recognized as such. It remains part of a set of tacit presuppositions.

In our opinion, moderate forms of reductionism as such still do not suffice either. Instead, we advocate a pluralist approach, the study of phenomena with multiple theories and methodologies on and between various levels. Here, both reductionist and holistic strategies are employed simultaneously and their relative contributions to explaining the phenomena depend on the degree of BioComplexity of the latter. The merits of the reductionist strategy are gladly incorporated. At the same time its shortcomings are acknowledged and the autonomy and usefulness of higher-level and inter-level explanations in physics, chemistry and biology are emphasized (Laughlin & Pines, 2000; Laughlin *et al.*, 2000; Srere, 2000; Strohman, 1997a,b; Westerhoff, 2000, 2001).

In this article, we have shown that emergence in cell biology, such as oscillations and well-managed turbo-design, is nothing exotic or non-scientific. It is a property of (non-linear) systems that have survived biological evolution. It is observable throughout the different organization levels of those systems. Through interactions between subsystems, properties can emerge that do not exist at the level of the individual components. Moreover, modeling of (sub)systems indicates that emergent properties can be calculated from the kinetic properties of components when the latter are embedded *in silico* in the topology of the biochemical network. The emergence of oscillations indicates that mechanistic explanations of emergent properties are possible when the systemic context is explicitly taken into account; through calculations one can determine whether a certain biochemical system has the tendency to display oscillations or not. We have intended to make the concept of emergence in cell biology operational.

A metabolic oscillation though is just a modest form of an emergent property. Macromolecular crowding and channeling are perhaps stronger examples of emergence. The strongest examples, of course, are to be found at the level of the whole cell. Higher-level concepts invaded cell biology already a long time ago (e.g. catabolism and anabolism, free-energy status). Yet, terms like adaptation, conditioning, memory, or possibly even learning as properties of specific biochemical networks are still not recognized as important in cell biology (Bruggeman *et al.*, 2002; Jonker *et al.*, 2002). A reason for anticipating the existence of such higher-level

functional properties is the notion that the genomes of primitive living unicellular organisms are substantially larger than is deemed necessary for basal metabolism and gene expression (Westerhoff, in press). This might indicate that microorganisms have more potential then first assumed. The sophisticated regulatory networks that function in microorganisms like *E. coli* may testify to this. The fact that the actual regulation of cellular functions is much more complex than what is needed for simple function suggests that the function of this complexity may be defined in terms of some sort of "macromolecular intelligence" (Bruggeman *et al.*, 2000; Hellingwerf *et al.*, 1995; Hoffer *et al.*, 2001).

Contemporary cell biology is growing from a qualitative to a quantitative science that progresses through precise and integrative (pluralist) experimental approaches. An increasing number of scientists emphasize the importance of recognizing the systemic functioning of living organisms and they are collectively referred to as system biologists (Ideker *et al.*, 2001; Kitano, 2001; Westerhoff, 2000, 2001). Since the genome sequencing programs for the yeast and human genomes, systemic research programs appear to be here to stay. Although the reductionist approach has undoubtedly delivered indispensable tools and data, including the genome sequence, it could now impede progress if it were to resort in the refusal to acknowledge that there is more to cell biology than just proteins.

Notes

[1] A Hopf bifurcation entails that the real part of an eigenvalue equals zero under the conditions that its imaginary part is unequal to zero (Guckenheimer & Holmes, 1990). Furthermore, eigenvalues can be expressed in terms of (unscaled) elasticities, which indicates that Hopf bifurcations can be mechanistically explained in terms of enzyme sensitivities (for their substrates, products and effectors) calculated in a specific steady state of reference.

[2] Genome: the identity and the location of all genes on the DNA. Transcriptome, proteome, and metabolome represent the individual concentrations of all mRNAs, proteins, and metabolites, respectively.

[3] A co-response coefficient is defined as the ratio of two control coefficients concerning one particular process with respect to two systemic properties (flux and/or metabolite concentration).

References

AINSCOW, E.K. & BRAND, M.D. (1999a). Top–down control analysis of ATP turnover, glycolysis and oxidative phosphorylation in rat hypatocytes. *European Journal of Biochemistry*, *263*, 671–685.

AINSCOW, E.K. & BRAND, M.D. (1999b). Internal regulation of ATP turnover, glycolysis and oxidative phosphorylation in rat hypotocytes. *European Journal of Biochemistry*, *266*, 737–749.

ALLEN, J.F. (2001). Bioinformatics and discovery: induction beckons again. *Bioessays*, *23*, 104–107.

ANDERSON, P.W. (1972). More is different: broken symmetry and the nature of the hierarchical structure of science. *Science*, *177*, 393–396.

BAKKER, B.M., MICHELS, P.A., OPPERDOES, F.R. & WESTERHOFF, H.V. (1997). Glycolysis in bloodstream form Trypanosoma brucei can be understood in terms of the kinetics of the glycolytic enzymes. *Journal of Biological Chemistry*, *272*, 3207–2315.

BAKKER, B.M., MENSONIDES, F.I., TEUSINK, B., VAN HOEK, P., MICHELS, P.A. & WESTERHOFF, H.V. (2000). Compartmentation protects trypanosomes from the dangerous design of glycolysis. *Proceedings National Academy of Science USA*, 97, 2087–2092.

BECHTEL, W. & RICHARDSON, R.C. (1993). *Discovering complexity: decomposition and localization as strategies in scientific research.* Princeton, NJ: Princeton University Press.

BOCK, G.R. & GOODE, J.A. (1998). *The limits of reductionism in biology*. Novartis Foundation Symposium 213, Chichester: Wiley.

BRAY, D. (1995). Protein molecules as computational elements in living cells. *Nature*, *376*, 307–212.

BROWN, G.C., HAFNER, R.P. & BRAND, M.D. (1990). A top–down approach to the determination of control coefficients in metabolic control theory. *European Journal of Biochemistry*, *188*, 321–325.

BRUGGEMAN, F.J., VAN HEESWIJK, W.C., BOOGERD, F.C. & WESTERHOFF, H.V. (2000). Macromolecular intelligence in microorganisms. *Biological Chemistry*, *381*, 965–972.

BRUGGEMAN, F.J., WESTERHOFF, H.V., HOEK, J.B. & KHOLODENKO, B.N. (2002). Modular response analysis of cellular regulatory networks. *Journal of Theoretical Biology*, *218*, 507–520.

CORTASSA, S., AON, M.A. & WESTERHOFF, H.V. (1991). Linear non-equilibrium thermodynamics describes the dynamics of an autocatalytic system. *Biophysical Journal*, *60*, 794–803.

COVERT, M.W., SCHILLING, C.H., FAMILI, I., EDWARDS, J.S., GORYANIN, II, SELKOV, E. & PALSSON, B.O. (2001). Metabolic modeling of microbial strains *in silico. Trends in Biochemical Sciences*, *29*, 179–186.

CRICK, F.H.C. (1966). *Of molecules and man*. Seattle, WA: Washington University Press.

DEMIN, O.V., WESTERHOFF, H.V. & KHOLODENKO, B.N. (1999). Control analysis of stationary forced oscillations. *Journal of Physical Chemistry*, *B103*, 10,695–10,710.

ELLIS, R.J. (2001a). Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences*, *26*, 597–604.

ELLIS, R.J. (2001b). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current Opinions in Structural Biology*, *11*, 114–119.

ENDY, D. & BRENT, R. (2001). Modeling of cellular behavior. *Nature*, *409*, 391–395.

FELL, D.A. (1997). *Understanding the control of metabolism.* London: Portland Press.

GILBERT, S.F. & SARKAR, S. (2000). Embracing complexity: organicism for the 21st century. *Developmental Dynamics*, *219*, 1–9.

GOODSELL, D.S. (1991). Inside a living cell. *Trends in Biochemical Sciences*, *6*, 203–206.

GUCKENHEIMER, J. & HOLMES, P. (1990). *Non-linear oscillations, dynamical systems, and bifurcations of vector fields*. Berlin & Heidelberg: Springer–Verlag.

HEINRICH, R. & RAPOPORT, T.A. (1974). A linear steady-state treatment of enzymatic chains: general properties, control and effector strength. *European Journal of Biochemistry*, *42*, 89–95.

HEINRICH, R. & RAPOPORT, T.A. (1977). Metabolic regulation and mathematical models. *Progress in Biophysical Molecular Biology*, *32*, 1–82.

HEINRICH, R. & SCHUSTER, S. (1996). *The regulation of cellular systems*. New York: Chapman & Hall.

HELLINGWERF, K.J., POSTMAN, P.W., TOMMASSON, J. & WESTERHOFF, H.V. (1995). Signal transduction in bacteria: phosphoneural network(s) in *Escherichia coli*? *FEMS Microbiology Reviews*, *16*, 309–321.

HOFFER, S.M., WESTERHOFF, H.V., HELLINGWERF, K.J., POSTMAN, P.W. & TOMMASSEN, J. (2001). Autoamplification of a two-component regulatory system results in "learning" behavior. *Journal of Bacteriology*, *183*, 4914–4917.

HOFMEYR, J.H.S. & CORNISH-BOWDEN, A. (2000). Regulating the cellular economy of supply and demand. *FEBS Letters*, *476*, 47–51.

HOFMEYR, J.H.S. & WESTERHOFF, H.V. (2001). Building the cellular puzzle: control in multi-level reaction networks. *Journal of Theoretical Biology*, *208*, 261–285.

IDEKER, T., GALITSKI, T. & HOOD, L. (2001). A new approach to decoding life: systems biology. *Annual Review of Genomics and Human Genetics*, *2*, 343–372.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860.

JENESON, J.A., WESTERHOFF, H.V. & KUSHMERICK, M.J. (2000). A metabolic control analysis of kinetic controls in ATP free energy metabolism in contracting skeletal muscle. *American Journal of Physiological and Cellular Physiology*, *279*, C813–C832.

JONKER, C.M., SNOEP, J.L., TREUR, J., WESTERHOFF, H.V. & WIJNGAARDS, W.C.A. (2002). Putting

intentions into cell biochemistry: an artificial intelligence perspective. *Journal of Theoretical Biology*, *214*, 105–134.

KACSER, H. & BURNS, J.A. (1973). The control of flux. *Symposium for the Society of Experimental Biology*, *27*, 65–104.

KELL, D.B & WELCH, G.R. (1991). No turning back: reductionism and biological complexity. *Times Higher Educational Supplement*, 9 August, 15.

KHOLODENKO, B.N. & DEMIN, O.V. (1993). Product stabilization and dynamic stability in unbranched metabolic pathways: what are the optimal structures? *Biochemistry*, *58*, 270–281.

KHOLODENKO, B.N., SCHUSTER, S., ROHWER, J.M., CASCANTE, M. & WESTERHOFF, H.V. (1995). Composite control of cell function: metabolic pathways behaving as single control units. *FEBS Letters*, *368*, 1–4.

KHOLODENKO, B.N., DEMIN, O.V. & WESTERHOFF, H.V. (1997). Control analysis of periodic phenomena in biological systems. *Journal of Physical Chemistry*, *B101*, 2070–2081.

KITANO, H. (2001). *Foundations of systems biology*. Cambridge, MA: MIT Press.

KOEFOED, S., OTTEN, M., KOEBMANN, B.J., BRUGGEMAN, F.J., BAKKER, B.M., SNOEP, J.L., KRAB, K., VAN SPANNING, R., VAN VERSEVELD, H., JENSEN, P.R. & WESTERHOFF, H.V. (2002). A turbo engine with automatic transmission? How to marry chemicomotion to the subtleties and robustness of life. *Biochimica Biophysica Acta*, *1555*, 75–82.

KREMLING, A. & GILLES, E.D. (2001). The organization of metabolic reaction networks II: signal processing in hierarchical structured functional units. *Metabolic Engineering*, *3*, 138–150.

KREMLING, A., BETTENBROCK, K., LAUBE, B., JAHREIS, K., LENGELER, J.W. & GILLES, E.D. (2001). The organization of metabolic reaction networks III: application for diauxic growth on glucose and lactose. *Metabolic Engineering*, *3*, 362–379.

TER KUILE, B.H. & WESTERHOFF, H.V. (2001). Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters*, *500*, 169–171.

LAUGHLIN, R.B. & PINES, D. (2000). The theory of everything. *Proceedings National Academy of Science*, *97*, 28–31.

LAUGHLIN, R.B., PINES, D., SCHMALIAN, J., STOJKOVIC, B.P. & WOLYNES, P. (2000). The middle way. *Proceedings National Academy of Science*, *97*, 32–37.

MATHEWS, C.K. (1993). The cell: bag of enzymes or networks of channels? *Journal of Bacteriology*, *175*, 6377–6381.

MONOD, J. (1972). *Chance and necessity: an essay on the natural philosophy of modern biology*. London: Collins.

MORANGE, M. (2001). *The misunderstood gene*. Cambridge, MA: Harvard University Press.

NAGEL, E. (1979). *The structure of science*. Indianapolis, IN: Hackett.

NICOLIS, G. & PRIGOGINE, I. (1989). *Exploring complexity: an introduction*. New York: Freeman.

NURSE, P. (1997). The ends of understanding. *Nature*, *387*, 657.

OVADI, J. (1995). *Cell architecture and metabolic channeling*. New York: Springer–Verlag.

PALSSON, B.O. & LIGHTFOOT, E.N. (1985). Mathematical modeling of dynamics and control in metabolic networks IV: local stability analysis of single biochemical control loops. *Journal of theoretical Biology*, *113*, 261–277.

POSTMA, P.W., LENGELER, J.W. & JACOBSON, G.R. (1993). Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiological Reviews*, *57*, 543–594.

REIJENGA, K.A., WESTERHOFF, H.V., KHOLODENKO, B.N. & SNOEP, J.L. (2002). Control analysis of autonomously oscillating biochemical networks. *Biophysical Journal*, *82*, 99–108.

ROHWER, J.M., SCHUSTER, S. & WESTERHOFF, H.V. (1996). How to recognize monofunctional units in a metabolic system. *Journal of Theoretical Biology*, *179*, 213–228.

ROHWER, J.M., POSTMA, P.W., KHOLODENKO, B.N. & WESTERHOFF, H.V. (1998). Implications of macromolecular crowding for signal transduction and metabolite channeling. *Proceedings National Academy of Science*, *95*, 10,457–10,552.

ROHWER, J.M., MEADOW, N.D., ROSEMAN, S., WESTERHOFF, H.V. & POSTMA, P.W. (2000). Understanding glucose transport by the bacterial phosphoenolpyruvate glycose phosphotransferase system on the basis of kinetic measurements in vitro. *Journal of Biological Chemistry*, *275*, 34,909–34,921.

ROSE, S. (1997). *Lifelines: biology, freedom, and determinism.* London: Penguin Books.

SCHILLING, C.H., SCHUSTER, S., PALSSON, P.O. & HEINRICH, R. (1999). Metabolic pathway analysis: basic concepts and scientific application in the post-genomic era. *Biotechnology Progress*, *15*, 296–303.

SCHRÖDINGER, E. (1992). *What is life?* Cambridge: Cambridge University Press.

SCHUSTER, S., KAHN, D. & WESTERHOFF, H.V. (1993). Modular analysis of the control of complex metabolic pathways. *Biophysical Chemistry*, *48*, 1–17.

SCHUSTER, S., FELL, D.A. & DANDEKAR, T. (2000). A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, *18*, 326–332.

SIMON, H. (1969). *The sciences of the artificial.* Cambridge, MA: MIT Press.

SNOEP, J.L., VAN DER WEIJDEN, C.C., ANDERSEN, H.W., WESTERHOFF, H.V. & JENSEN, P.R. (2002). DNA supercoiling in *Escherichia coli* is under tight and subtle homeostatic control, involving gene-expression and metabolic regulation of both topoisomerase I and DNA gyrase. *European Journal of Biochemistry*, *269*, 1662–1669.

SRERE, P.A. (2000). Macromolecular interactions: tracing the roots. *Trends in Biochemical Sciences*, *25*, 150–153.

STROHMAN, R.C. (1997a). Profit margins and epistemology. *Nature Biotechnology*, *15*, 194–200.

STROHMAN, R.C. (1997b). The coming Kuhnian revolution in biology. *Nature Biotechnology*, *15*, 1224–1225.

TEUSINK, B., WALSH, M.C., VAN DAM, K. & WESTERHOFF, H.V. (1998). The danger of metabolic pathways with turbo design. *Trends in Biochemical Sciences*, *23*, 162–169.

TEUSINK, B., PASSARGE, J., REIJENGA, C.A., ESGALHADO, E., VAN DER WEIJDEN, C.C., SCHEPPER, M., WALSH, M.C., BAKKER, B.M., VAN DAM, K., WESTERHOFF, H.V. & SNOEP, J.L. (2000). Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *European Journal of Biochemistry*, *267*, 5313–5329.

VENTER, J.C. *et al.* (2001). The sequence of the human genome. *Science*, *291*, 1304–1351.

VERHAMME, D.T., ARENTS, J.C., POSTMA, P.W., CRIELAARD, W. & HELLINGWERF, K.J. (2002). Investigations of *in vivo* cross-talk between key two-component systems of *Escherichia coli*. *Microbiology*, *148*, 69–78.

WAGNER, A. (1999). Causality in complex systems. *Biology and Philosophy*, *14*, 83–101.

WATSON, J.D. & CRICK, F.H.C. (1953a). Genetical implications of the structure of Deoxyribonucleic acid. *Nature*, *171*, 964–967.

WATSON, J.D. & CRICK, F.H.C. (1953b). Molecular structure of nucleic acids—a structure for deoxyribosenucleic acid. *Nature*, *171*, 737–738.

WELCH, G.R. (1986). *The fluctuating enzyme.* New York: Wiley.

WELCH, G.R. (1992). An analogical "field" construct in cellular biophysics: history and present status. *Progress in Biophysical Molecular Biology*, *57*, 71–128.

WESTERHOFF, H.V. (2000). Perspectives: live perspectives of biochemistry. *The Biochemist*, *22*, 43–47.

WESTERHOFF, H.V. (2001). The silicon cell, not dead but live! *Metabolic Engineering*, *3*, 207–210.

WESTERHOFF, H.V. (in press). BioComplexity for starters; getting our acts together. *BioComplexity.*

WESTERHOFF, H.V. & KAHN, D. (1993). Control involving metabolism and gene expression: the square matrix method for modular decomposition. *Acta Biotheoretica*, *41*, 75–83.

WESTERHOFF, H.V. & VAN DAM, K. (1987). *Thermodynamics and control of free energy transduction.* Amsterdam: Elsevier.

WESTERHOFF, H.V., KOSTER, J.G., VAN WORKUM, M. & RUDD, K.E. (1990). On the control of gene expression. In A. CORNISH-BOWDEN & M.L. CARDENAS (Eds) *Control of metabolic processes* (pp. 399–412). New York: Plenum Press.

WESTERHOFF, H.V., JENSEN, P.R., SNOEP, J.L. & KHOLODENKO, B.N. (1998). Thermodynamics of complexity: the live cell. *Thermochimica Acta*, *309*, 111–120.

WILLIAMS, N. (1997). Biologists cut reductionist approach down to size. *Science*, *277*, 476–477.

WILSON, E.O. (1999). *Consilience: the unity of knowledge.* New York: Random House.

WIMSATT, W.C. (1997). Aggregativity: reductive heuristicss for finding emergence. *Philosophy of Science*, *64*, S373–S384.

ZIMMERMAN, S.B. & MINTON, A.P. (1993). Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual Review of Biophysics and Biomolecular Structure*, *22*, 27–65.

Appendix

The set of autonomous differential equations that governed the evolution of the metabolites of the biochemical network depicted in Figure 4C were:

$$\frac{\mathrm{d}X_1(t)}{\mathrm{d}t} = \frac{V_1}{1 + \left(\frac{X_3(t)}{K_{1X_3}}\right)^n + \frac{X_1(t)}{K_{1X_1}}} - V_2 \cdot X_1(t)$$

$$\frac{\mathrm{d}X_2(t)}{\mathrm{d}t} = \frac{V_3}{1 + \left(\frac{K_{3X_1}}{X_1(t)}\right)^{n3} + \frac{X_2(t)}{K_{3X_2}}} - \frac{V_4 \cdot \frac{X_2(t)}{K_{4X_2}}}{1 + \frac{X_2(t)}{K_{4X_2}} + \frac{X_3(t)}{K_{4X_3}}}$$

$$\frac{\mathrm{d}X_3(t)}{\mathrm{d}t} = \frac{V_4 \cdot \frac{X_2(t)}{K_{4X_2}}}{1 + \frac{X_2(t)}{K_{4X_2}} + \frac{X_3(t)}{K_{4X_3}}} - V_5 \cdot X_3(t)$$

where $V_1 = 2.5$, $K_{1X_3} = 0.00025$, $K_{1X_1} = 0.1$, $n_1 = 3$, $V_2 = 10$, $V_3 = 10$, $K_{3X_1} = 0.15$, $K_{3X_2} = 0.1$, $V_4 = 10$, $n_3 = 3$, $K_{4X_2} = 1$, $K_{4X_3} = 1$, $V_5 = 10$.