J Comput Aided Mol Des (2011) 25:699–708
DOI 10.1007/s10822-011-9447-8

PERSPECTIVE

The reductionist paradox: are the laws of chemistry and physics sufficient for the discovery of new drugs?

Gerald M. Maggiora

Received: 18 May 2011 / Accepted: 13 June 2011 / Published online: 23 June 2011

Abstract Reductionism is alive and well in drug-discovery research. In that tradition, we continually improve experimental and computational methods for studying smaller and smaller aspects of biological systems. Although significant improvements continue to be made, are our efforts too narrowly focused? Suppose all error could be removed from these methods, would we then understand biological systems sufficiently well to design effective drugs? Currently, almost all drug research focuses on single targets. Should the process be expanded to include multiple targets? Recent efforts in this direction have lead to the emerging field of polypharmacology. This appears to be a move in the right direction, but how much polypharmacology is enough? As the complexity of the processes underlying polypharmacology increase will we be able to understand them and their inter-relationships? Is "new" mathematics unfamiliar in much of physics and chemistry research needed to accomplish this task? A number of these questions will be addressed in this paper, which focuses on issues and questions not answers to the drug-discovery conundrum.

Keywords Biological reductionism · Emergent properties · Hierarchy · Drugs

G. M. Maggiora (✉)
Department of Pharmacology & Toxicology, College of Pharmacy and BIO5 Institute, University of Arizona, Tucson, AZ 85721, USA
e-mail: gerry.maggiora@gmail.com

G. M. Maggiora
Translational Genomics Research Institute, 445 N. Fifth Street, Suite 600, Phoenix, AZ 85004, USA

Introduction

Reductionism is alive and well in modern biological and drug research. In that tradition, we continually improve experimental and computational methods for studying smaller and smaller aspects of biological systems. This may be a symbol of the ultimate success of the reductionist approach, which has dominated science for many years, but the success has come at a price; namely, functional relationships among the components have become obscured. Ironically, the more detail of the molecular components that is revealed, the more confounding the mechanisms relating these components. Are our efforts becoming too narrowly focused? Suppose all error could be removed from these methods, would we then understand biological systems sufficiently well to design more effective drugs? The answer we shall see is most likely no.

Biological systems have a hierarchical structure as indicated in blue type on Fig. 1. On the "Reductionist" side of the figure, dramatic increases in the capabilities of molecular, cellular, and structural biology have enabled us to elucidate the structural details of a growing number of macromolecular entities until we are now able to solve the structure of proteins with no known function(s) [1]. However, as we move down the hierarchy much of the biological environment within which these entities function is "edited" out.

While this detailed structural knowledge has revolutionized the drug-discovery process, it has come at a cost. For example, proteins representing esoteric drug targets can be expressed in various stable cell types or in some cases in cell-free systems. Whether these environments are representative of the biological environment in which the target typically functions is, in many cases, problematic. Nevertheless, the relative simplicity of the systems has

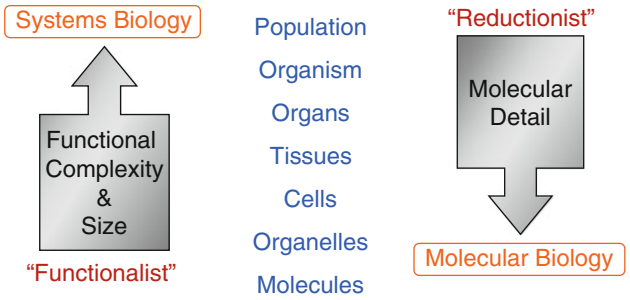

Fig. 1 Hierarchical nature of biological systems

given rise to numerous high-throughput screening (HTS) methods that, coupled with extensive compound collections, have enabled scientists to mine extensive regions of chemical space for biologically active compounds. This brute force approach, while effective in identifying active compounds, has not significantly increased the number of truly new chemical entities or drugs. In fact, during the period from 1988 to 2004 the number of approved new molecular entities (NMEs) stayed relatively constant at slightly more than 20, about half the number of non-NMEs and repurposed drugs. In contrast, research expenditures have risen sharply from 10 to around 40 million USD [2].[1]

On the functionalist side of Fig. 1, moving up the biological hierarchy, function is "reintroduced" and the size and complexity of the systems tend to increase. This represents a more realistic state of affairs that may be difficult to study in some cases since all of the factors comprising the system are generally not known or fully understood. However, this is not meant to imply that the functionalist approach must be applied from the bottom up. Such approaches are not, in general, possible because complex systems exhibit emergent properties (vide infra) that cannot be inferred from knowledge of the individual components alone, unless a theoretical framework (i.e., a theory) exists that describes the function(s) of the components [3].[2] In fact, many of the models that have been constructed are based on a top-down approach that includes enough functional detail to provide a consistent explanation of the biological phenomena under study [4].

Because of the current attention being paid to studying biological systems at higher levels of the hierarchy, there appears to be renewed interest in classical pharmacology, a field that has been on the wane for some years [5]. Moreover, viewing biology, pharmacology, and medicine from a more systems perspective, opens up new possibilities for drug therapy that go beyond the classical one target-one drug model (vide infra) that has dominated drug research for many years [6, 7].

This paper is not about the challenges posed by the development and application of chemical and physical methods for studying biological systems. Rather, it is concerned with issues associated with the nature of these systems and why the traditional approach to pharmaceutical research has resulted in a diminishing number of new pharmaceuticals. And, as a consequence of this, it is also concerned with the changing role(s) of computational methods in biological and drug research, how they might help provide new approaches, and why they have essential roles to play in future research.

Biological reductionism

Basically, the reductionist approach seeks to decompose biological systems into their constituent parts in an effort to understand the biology induced by these parts [8]. Philosophers of science have defined three distinct, albeit closely related, types of biological reductionism [9].[3] Methodological reductionism is closest to what is used here.

Discussions on biological reductionism have gone on for many years, where much of the focus has been on whether the laws underlying chemistry and physics are sufficient to explain biology. Michael Polanyi, who has been one of the most outspoken and eloquent critics of biological reductionism, succinctly states [10]: "When I say that life transcends physics and chemistry, I mean that biology cannot explain life in our age by the current workings of physical and chemical laws". Early on, such an observation implied a vitalistic view of biology, namely, that some type of "life force", not understandable in terms of chemistry and physics, was responsible for the formation, reproduction, and sustainability of biological systems, but as will be seen in the sequel, such assumptions are not required to understand biology.

No doubt the reductionist approach of dissecting biological systems into their constituent parts has been effective in explaining the chemical and physical basis of

[1] Note that this estimate is from the Pharmaceutical Research and Manufacturers of America (PhRMA). An estimate by the National Science Foundation put the change at slightly more than half of the PhRMA estimate. All estimates are in constant dollars, inflation being removed.

[2] An example applicable to thermodynamic systems may be appropriate here. The theoretical framework of statistical mechanics provides a means (i.e. a theory) for determining the values of macroscopic variables (e.g., enthalpy, free energy, pressure, etc.) from the microscopic variables of individual molecules.

[3] As might be expected, philosophers of science provide a much more fine-grain approach to reductionism in biology, defining three principal types of reduction—ontological, methodological, and epistemic. Methodological reductionism comes closest to that described here, although bits of all three may implicitly be included. An excellent and thorough discussion is given in the section on Reductionism in Biology in the Stanford Encyclopedia of Philosophy, which is freely accessible on the Internet—See Ref. [9].

living systems. Many biologists now, however, realize that this approach has reached its limit because biological systems are highly complex and possess *emergent properties* [8] (vide infra). And it is the notion of an "Organizing Principle" (or theory of biology) as described by Polanyi [10] (cf. [3]) that provides a means for characterizing and understanding the relationship of the components of biological systems to the properties that emerge from them (vide infra).

Emergent properties

The function(s) of systems that possess emergent properties cannot be explained, or even predicted, by studying their individual parts; higher-level knowledge (e.g. a theory) of what function(s) these parts are to perform is needed. For example, take a clock. If one is presented with the collection of parts from a totally dismantled clock and is not told their function, it would be difficult to reconstruct the clock. A more dramatic example of emergent properties is seen in Fig. 2. On the left of the figure is a pile of colored Legos™, while on the right is a *model* of a Babbage difference engine constructed by Andy Carol out of Legos™ that evaluates the function $f(x) = ax^2 + bx + c$ for $x = 0$, 1, 2, 3,…,n to three-digit accuracy [11]. The left directed arrow represents the direction of reductionism; the right directed arrow represents the direction of functionalism. Clearly, dismantling the Babbage difference engine is a relatively simple task. However, reconstructing it from its constituent parts is definitely not a simple task unless some knowledge of its desired function is known, and even then it might be quite difficult.

This also brings up the concept of analog systems—systems that carry out the same function(s) but are constructed of different parts. Figure 3 shows other implementations of the Babbage difference engine. Since they carry out the same function they are analogs of the one shown in Fig. 2. Going back to the clock example, one can certainly think of a number of clock analogs that look quite different, but nonetheless tell time. As will be seen in the sequel, the concept of analog systems is somewhat akin to that of degenerate systems in biology.

A last example involves function estimation. As is well known, many mathematical functions can be represented in terms of linear combinations of other functions, called basis functions. Each of these sets of functions is different in form from the others as shown in the expression below,

$$f = \sum\nolimits_k c_k^{\alpha} \varphi_k^{\alpha} \Leftarrow \begin{cases} \{\varphi_1^1, \varphi_2^1, \varphi_3^1, \ldots\} \\ \{\varphi_1^2, \varphi_2^2, \varphi_3^2, \ldots\} \\ \vdots \\ \{\varphi_1^n, \varphi_2^n, \varphi_3^n, \ldots\} \\ \vdots \end{cases}$$

where the sets of functions to the right of the large curly bracket are the basis functions used to represent the function f. All of the different basis sets are capable of representing f, although some are more efficient than others (i.e., they can provide a good estimation of the function using fewer terms than less efficient sets of basis functions). This example represents two points noted in the above discussion. First, a set of basis functions is ineffective unless one knows what function is to be estimated. And second, the different basis sets can be considered to be similar to the analog representations of, for example, the Babbage difference engine described above. These examples illustrate the fact that systems that behave in a similar manner need not be similar in the exact manner in which they are constructed.

To summarize the discussion so far: (1) there are many ways to represent the same functionality, (2) a given functionality may be deconstructed ("reduced") at some level into its constituent parts, and (3) the reverse process is not, in general, possible unless a theory or functional principles are known.

The laws of nature

The laws of the physical sciences are basically simple, for example,

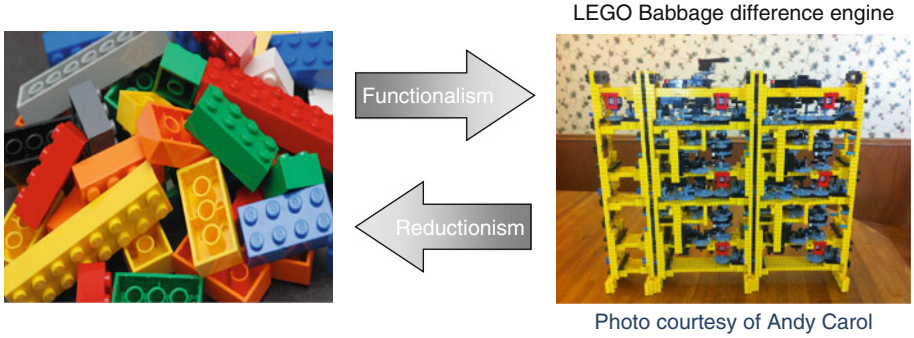

Fig. 2 Diagram illustrating an emergent property (*viz.* the ability to compute) based on the construction of a Babbage difference engine entirely out of individual Legos™ [9]

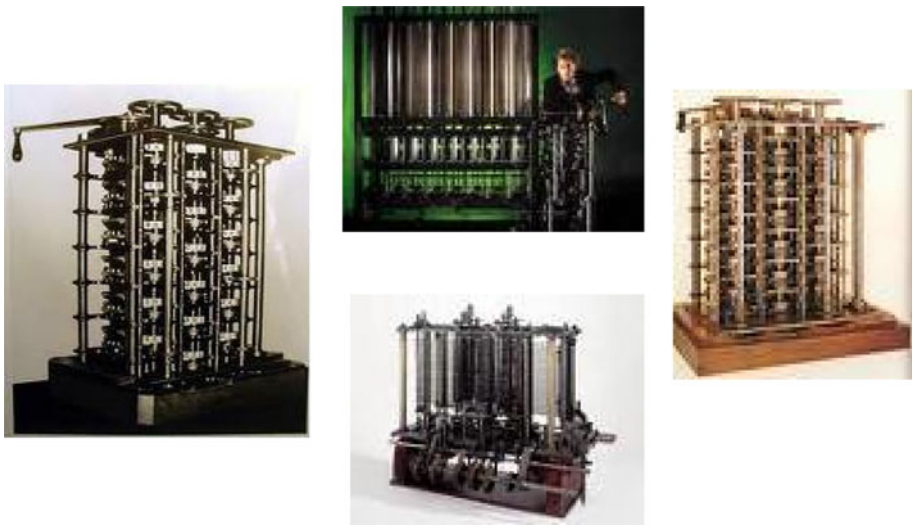

Fig. 3 Other realizations of the Babbage difference engine

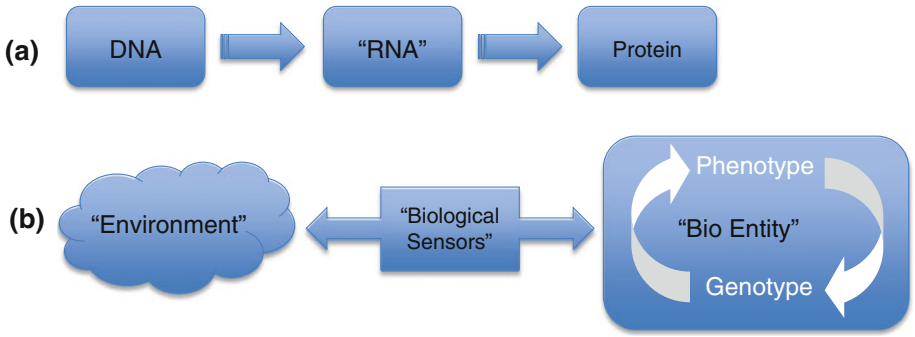

Fig. 4 **a** The diagram depicts the linear relationship exemplified by the "Central Dogma of Molecular Biology," namely that the flow of genetic information goes from DNA → RNA (Several forms) → Protein. **b** The diagram depicts a more realistic relationship between the "Environment" and a "Biological Entity" (organism, organ, cell,…) that exists within it. A variety of "Biological Sensors" mediates the interactions. The overall biological system is a highly complex, open system with significant non-linear behavior and feedback controls. Adapted from [12]

Mechanics: $F \sim ma$

Electrostatics: $F \sim \frac{qq\prime}{r^2}$

Relativity: $E = mc^2$

There are no equivalently simple laws in the biological sciences. "Laws of Biology" if they exist, are inherently complex. For example, the Central Dogma of Molecular Biology that DNA → RNA → Protein does not accord with "biological reality." Figure 4 illustrates this point in more detail. From part (a) of the figure it is clear that the linear relationship among DNA, RNA, and proteins is an extreme oversimplification. Part (b) depicts a more realistic model of the relationship of a biological system's genotype and its phenotype. As shown, a cell's, or more generally an organism's, phenotype is influenced by genomic and epigenetic phenomena, the latter being linked to a variety of biological sensors that are able to sense their environments and influence the function(s) the system can carry out [12]. Thus, knowledge of a cell or organism's genome alone is clearly insufficient if one is to understand its biology from early stages of existence, through development to maturity, and ultimately to death. In fact, even knowing its complete proteome or any of its many possible "-ome's" is still insufficient [13–15]. What is needed, minimally, is a more systems oriented view that integrates all of the various omics components within a hierarchical framework that depicts the interrelationships of a cell or organism's molecular, macromolecular, and supramolecular components. This expands on an earlier quote by the renowned biomathematician Nicolas Rashevsky [16]:

> As we have seen, a direct application of the physical principles, used in the mathematical models of biological phenomena, for the purpose of building a theory of life as an aggregate of individual cells is not likely to be fruitful. We must look for a principle

> [that] connects the different physical phenomena involved and expresses the biological unity of the organism and the organic world as a whole.

While the more detailed description of biological systems espoused above is a step in the right direction, it omits some elements that are crucial to a full understanding. Thus, treating such systems in complete detail is a daunting challenge to say the least. Fortunately, it is not necessary or, perhaps, even desirable to do so in all cases. Rather, the challenge is in developing models of sufficient detail [17] to describe processes that occur at hierarchical levels associated with the phenomena one wants to investigate. Depending on the circumstances, models constructed at these levels need not be described in *completely* molecular terms. Rather, they can be developed using "black box" or "grey box" formalisms [18], where each of the boxes simulates a particular biological function (or functions) that is (are) activated by a suitable set of inputs to deliver an appropriate set of outputs.

Why are biological systems so difficult to study?

Biological systems are difficult to study for the following reasons: (1) they are complex in terms of the number and variety of their components, the types and strengths of interactions that connect them, and the hierarchical structure (vide supra) within which they are embedded, (2) they possess emergent properties (vide supra), (3) they exhibit redundancy and degeneracy, (4) they exhibit modularity in terms of their subsystems, (5) they are open, stable non-equilibrium systems, and (6) they exhibit significant non-linear behavior.

A number of these features are well known, and some have already been alluded to. The concept of degeneracy, however, which is quite different from that of redundancy, generally is not well known [19]. Redundancy in biological systems refers to multiple copies of the same subsystem, so that if one becomes inactivated others can still carry out the required function(s). Degeneracy, on the other hand, refers to the condition where multiple, but *different* subsystems, can effectively carry out the same function. This, as noted earlier, is closely related to the concept of analog systems. In fact, the existence of degeneracy in biological systems means that subsystems, which may be unknown to us, could exist that can carry out analogous functions to those subsystems that we are aware of. Thus, directly modifying gene function using knockout, knockin, or conditional mutant methods [20] or modifying it using siRNA or related methods [21] may not always yield the expected result(s). This may in certain cases represent a serious challenge to the development of effective models in systems biology.

What is required to understand biological systems? According to biochemist and systems biologist Leroy Hood [22]

> To really understand [complex, biological] systems, you have to collect global datasets from each of these levels [of the hierarchy] and then integrate them together if you're to get a coherent understanding of the system.

To obtain the massive amounts of data required by such an approach suggests the need for a highly mechanized high-throughput approach to biology. Although great strides have been made in the development of suitable assays, significant factors exist that lessen the desirability of such an approach. These include its high initial and ongoing costs, its primarily data and not hypothesis driven nature (cf. [15]), and its implicit use of what I would call the "there must be a pony in here somewhere" point of view.[4] In more explicit terms this latter point refers to the notion held by many, especially in the Pharma community, that if enough data are gathered we are bound to find something interesting.

Not surprisingly, there are hidden costs of high-throughput, data-driven research, namely, the industrialization of biology and ultimately the resulting move towards assembly line science. Such a trend, if it continues for too long, runs counter to the requirements needed to ensure an effective research environment and will have long-term consequences on research productivity. There are also a number of related issues associated with how much biology we can understand from such mega-experiments, which by their very nature generate large amounts of noisy data that can confound any analysis, whether data-driven or not.

Biology has now become a "data-rich" science [13, 15, 23]. Is so much data a good or bad thing? Do the environments in which the experiments are carried out adequately represent the actual environments of the biological systems being modeled? Most of the data collected are time independent; how important are time-dependent effects

[4] While there are many versions of the story, it basically goes as follows. The parents of six-year-old twin boys are worried that they are developing extreme personalities. One appears to be a total pessimist and the other a total optimist. So the parents take them to a psychiatrist. First, the psychiatrist treats the pessimist. He takes him into a room filled with all sorts of toys to lift his spirits, but instead of being delighted the boy begins to cry. Asked why he was crying, the boy replies "If I play with the toys I'll probably just break them." The psychiatrist then treats his brother, trying to dampen his outlook by taking him into a room filled with horse manure. Unlike his brother, the boy is entirely delighted and begins digging through the manure with his bare hands. Somewhat taken aback, the psychiatrist asks the boy what he is doing, to which the boy replied "With all of the manure in this room, there must be a pony in here somewhere".

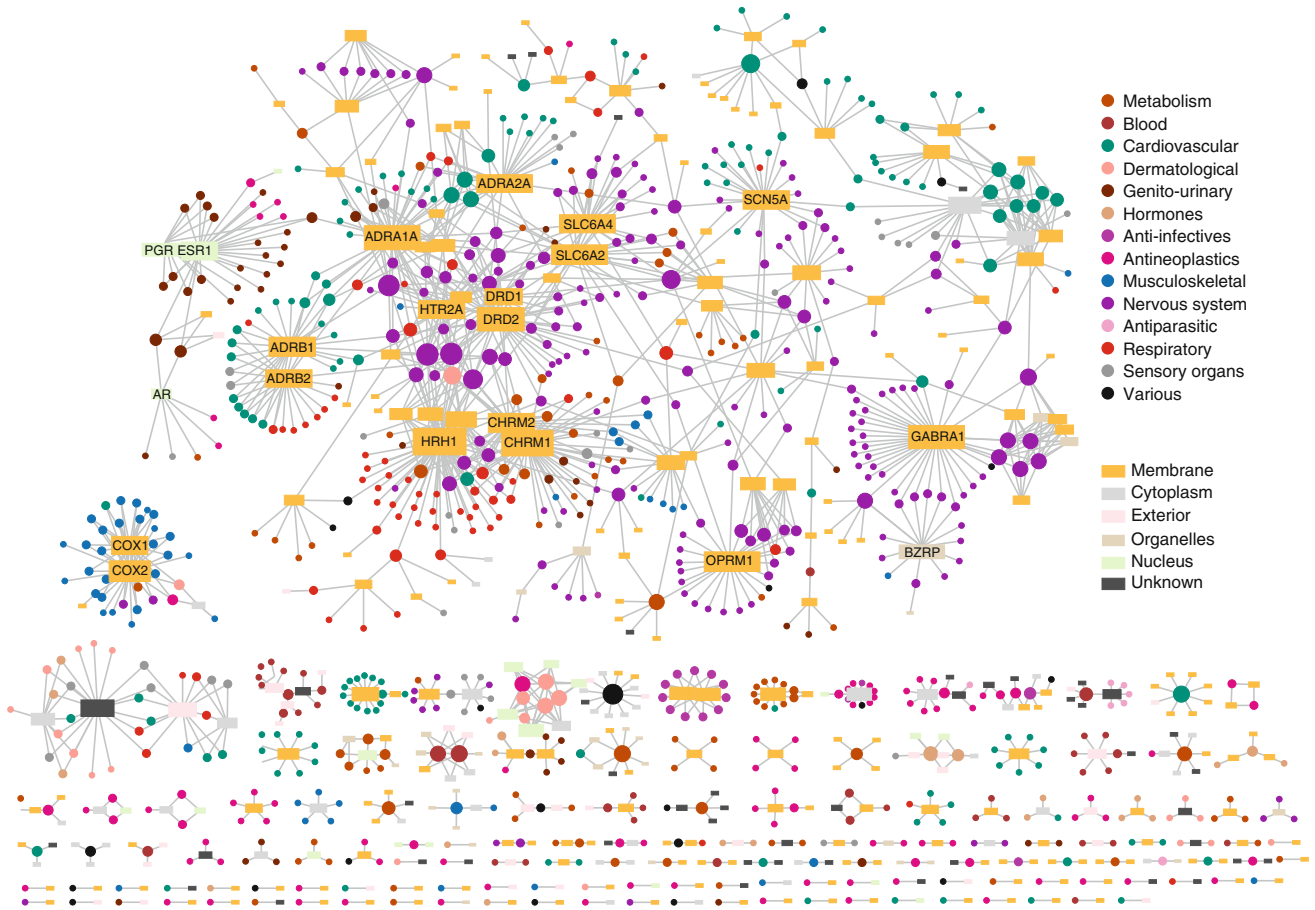

Fig. 5 Drug-target (DT) network taken from the work of Yildirim et al. [26]. The network (a bipartite graph) is constructed from associations of FDA-approved drugs with their known target proteins. A link is drawn between a drug node (filled circle colored according to its Anatomical Therapeutic Chemical Classification) and each of its known target proteins (filled rectangles colored according their cellular component obtained from the Gene Ontology database). The size of each drug (protein) node is proportional to the number targets (drugs) it interacts with. (*Reprinted with permission*)

(e.g., multiple timescales, event sequences, and threshold effects)? True cellular concentrations tend to be quite low, how can they be appropriately modeled? In addition, the microenvironments of cells, organelles, organs, and organisms are highly heterogeneous; how is this accommodated into the models? Because of the large number of variables, are chance correlations a substantive issue? Can we find patterns in such large, noisy, and possibly sparse datasets, especially because it is quite likely that there will be many weak "signals"? Can we obtain data of sufficient precision to model such large and complex systems? Will we be able to understand the results? Regardless whether all of these factors need to be accounted for in every case, it is quite clear that the capability to develop and handle complex, computable, an extensible models of biological systems and subsystems is of paramount importance to future advances in biological and drug research. Given all of the potential impediments to developing effective models of biological systems or subsystems, considerable recent activity has, nevertheless, taken place [24, 34].

An important point is that computational and laboratory-based models are both by their very nature imperfect. While this is well appreciated with regard to the computational models of biological systems, it is not always the case when dealing with the experimental models. For example, as is well known from many studies, mice are not always suitable models for higher species [25], even though a mouse is a living system, and possesses some functions that are equivalent to those of a human. *Thus, one should not assume that because experiments on cells, organs, or whole animals are carried out in a laboratory that they are necessarily any more relevant than experiments based on mathematical models of these systems carried out on a computer.*

Where are the new drugs?

As noted in the Introduction [2], the number of approved NMEs has stayed relatively constant from 1988 to 2004 at

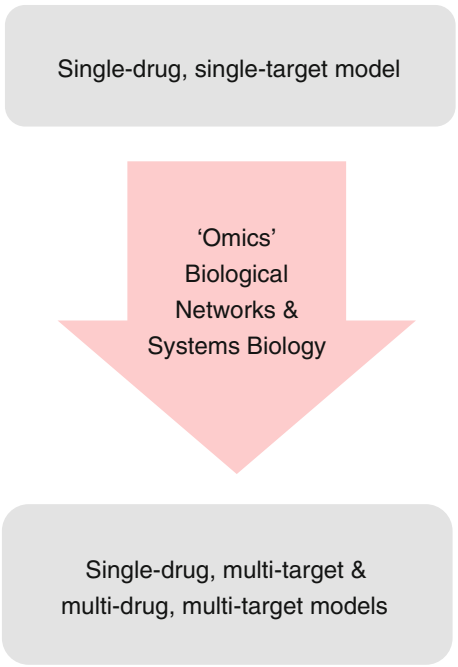

Fig. 6 Diagram depicting the emerging paradigm shift in drug research

about 20, which is half the number of non-NME and repurposed drugs developed over the same period. However, expenditures for research have skyrocketed from 10 to 40 million USD. What's the problem? While many analyses have been offered, it seems clear that the reductionist approach currently dominating drug research has played a role. What should be done about it? As depicted in Fig. 1, drug research needs to move to a more systems-biology oriented approach by putting more biology into pharmaceutical research. And this is, indeed, happening. For example, a plethora of network models are being developed that characterize the many relationships between drugs and their target proteins (and genes) [26–31]. These networks clearly show, as has been known for many years, that most drugs physically interact with more than one target. Figure 5 provides an example of one such network.

In addition, the growing power of today's computers has made it possible to develop a range of models of varying sophistication designed to simulate various biological processes [4, 32]. There is also a re-emergence of "classical" pharmacology coupled with the development of more reliable animal models and more effective biomarkers [33]. And clinical research is growing in importance [34].

All of this serves to further the emerging paradigm shift depicted in Fig. 6. The old single-drug, single-target model is giving way, albeit grudgingly, to new models that are based upon single-drug, multi-target and multi-drug, multi-target models. A growing number of omics technologies and the corresponding interest in biological networks and systems biology are driving this shift. Nevertheless, the single-drug, single-target paradigm, which is based upon the notion that a drug that attacks a single target with high specificity is more desirable than a more promiscuous "dirty drug" that attacks multiple targets, still dominates current research efforts [35]. This view, however, is changing, due to a heightened awareness of "off-target effects" [36], and what were at one time called dirty drugs are now being call multi-functional drugs. These drugs are seeing broader application in many therapeutic areas including psychopharmacology, which is experiencing an increasing use of such agents for multiple therapeutic applications [37]. From a broader perspective, the promising new field of *polypharmacology* is beginning to more systematically document the actual instances where single or multiple drugs interact with multiple targets [27, 29], and this is being further extended into the area of rational drug design [28].

What is the role of computational methods in the "new" biology?

Although theoretical and computational methods have a long and distinguished history in physics, this has not been the case for chemistry and biology. The situation is, however, changing, and a flood of papers in computational chemistry and biology has been published over the last decade. This upward trend in publication, which shows no sign of abatement, can be attributed, albeit not exclusively, to a number of factors: (1) the incredible increases in computer power now available even on the desktop, (2) the need to handle the massive amounts of data being generated by an increasing variety of chemical and biological experiments, (3) the need to discover and to elucidate relationships within the data, and ultimately, (4) the need to construct chemical/biological models that explain the data.

Although increases in computer power have enabled computations on a scale unimaginable only a relatively few years ago, this is not by itself sufficient. To paraphrase the mathematician John Casti, "New mathematics in addition to that traditionally used in physics and chemistry is needed if we are to fully understand biology [38]. Mathematicians, computer scientists, statisticians, engineers, and even sociologists have answered the call, providing a rich framework from which to study biological systems at all levels of the hierarchy. These methods include graphs, hypergraphs, and networks (including probabilistic ones), a wide variety of pattern recognition and machine learning methods, continuous, discrete, and stochastic biological simulations, and computational decision theory.

While the impact of these methods has been dramatic, many issues remain to be dealt with. Not the least among them are those associated with the size and complexity of biological systems and the consequent amount of detailed information that arises from these systems. To paraphrase Lofti Zadeh, the acknowledged father of fuzzy set theory and fuzzy logic, "As the amount of information grows, the level of detail at which it can be treated effectively must decrease."

An apt analogy from simulations of protein dynamics springs to mind. If the level of molecular detail is maintained, that is if all of the atoms including the hydrogen atoms are explicitly accounted for, then the most prevalent motions observed are those associated with the relatively fast and biologically uninteresting motions of the hydrogens. None of the more interesting large-scale conformational changes will be observed unless the simulations are carried out over extremely long time periods, which are very demanding of computational resources. The problem has been approached in some cases by simplifying the protein structural features using simple "ball and spring" models. Such models have been reasonably successful in characterizing the relatively slow conformational changes associated with the mechanisms of many biochemical processes.

In addition, to the problems created by the vast amount of detail information that biological systems can generate, an even greater problem may arise from the fact that information gathered in most biological experiments is to varying degrees imprecise and uncertain. Thus, the need arises for additional mathematical methods that are capable of handling such information. Soft computing methods provide a potential means for achieving this goal [39], which begs the question of what is soft computing? To again paraphrase Lofti Zadeh [40, 41],

> Soft computing is an emerging approach to computing, which parallels the remarkable ability of the human mind to reason and learn in an environment of uncertainty and imprecision.

Methods associated with soft computing include fuzzy set theory and fuzzy logic [42, 43], neural networks [44], probabilistic reasoning [45], genetic and evolutionary algorithms [46], belief networks [47], rough set theory [48], and granular computing [49] to name a few. And although some of these methods have experienced tremendous growth over the last decade or so, very few except, perhaps, neural networks and genetic/evolutionary algorithms have been applied to biological problems in a substantive manner. Thus, the time is ripe for bringing some of the other methodologies such as those associated with fuzzy mathematics and rough set theory to bear on biological problems.

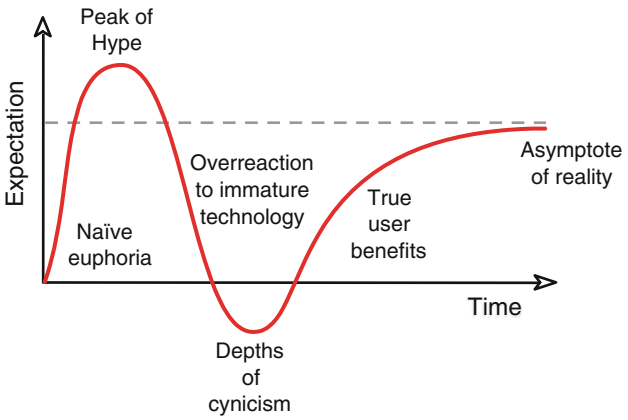

Fig. 7 The "Technology Cycle" after Bezdek [50]

Recapitulation and conclusions

The concept of biological reductionism has served us well over the years, but at a cost that is now becoming clearer. Elegant details that have been uncovered by cell, chemical, and structural biologists in the last several decades have spurred efforts to continue in this direction, but it is becoming apparent that more attention needs to be paid to the overall functional aspects of living systems. Although, it is easier, but by no means easy, to resolve living systems into their component parts, because they are complex and thus possess emergent properties, it is difficult to reconstruct them without some knowledge of what a reconstructed system is to accomplish. A bottom-up approach can, however, be taken to some extent if a theoretical framework exists for the system under study [3].

A related question is whether all of the laws governing biological systems are contained within those of physics and chemistry or whether laws specific to biological systems exist? Such laws were once associated with the concept of vitalism, namely, that the principles underlying biological systems go beyond those governing purely physico-chemical systems. This seems to contradict the notion that biological systems are natural systems and thus, should be governed by the laws that govern natural sciences.

For example, consider the Central Dogma of Molecular Biology (Fig. 4a), which delineates the fundamental sequential relationship: DNA → {RNA} → Protein. While the laws of physics and chemistry provide a basis for understanding the structure and interactions of all of these molecules, they do not by themselves describe the function(s) that these molecules carry out with respect to the Central Dogma, namely, the transfer of information from a specific sequence of bases in DNA to a specific sequence of amino acids in the corresponding protein. But the Central Dogma is even simpler than biological reality, which

requires a more complex set of relationships that couple genotypic and phenotypic factors to each other and to the "Environment" within which the system resides through a series of "Biological Sensors" (Fig. 4b).

It is clear that a more systems-oriented approach is needed that takes account of the hierarchical nature of biological systems and the many complex interactions amongst their components if we are to begin to get a handle on how drugs influence and are influenced by the biological systems into which they are administered. This can only be accomplished using the tools of computational biology (e.g., cellular simulations) in close conjunction with the burgeoning number of methods being used in modern, experimental biological and clinical research.

The need for computational methods and mathematical models of living systems is paramount because of the complexity of these systems, especially those of mammals and primates, since there is simply no way that the human mind can grasp the many features of these systems. Even the information in cartoons of biological pathways, although helpful, can be too complex to grasp completely. In addition, information is usually neglected on the time-dependent processes, threshold effects, and heterogeneous environmental factors in such cartoons, which may crucial to an understanding of the biological function(s) under study.

In this regard, biological systems are particularly difficult to study because there are so many small interactions that must be properly accounted for if meaningful results are to be obtained. A simple, almost trivial example of this is the prediction of the tertiary structure of proteins. As there are no dominant interactions responsible for folding, the many small interactions must be accounted for reasonably accurately since small errors in each term can accumulate leading to large errors in predictions. Such accuracy is difficult to obtain in practice, especially when one considers the many interactions among elements of a protein's structure that must be accurately accounted for, and the need to also the account for micro-environmental factors that arise from the biological milieu in which the protein resides.

Developing enhanced experimental and computational technologies is important, but no single method is capable of providing a definitive understanding of living systems. Millions of dollars have been invested in attempts to find what is equivalent to the "Silver Bullet" of drug research, namely, a single method that is so powerful that the results it produces will overcome any of the many obstacles that must be faced in the drug discovery and development process. Remember the hype that surrounded the introduction of combinatorial chemistry. There are many more examples of technologies that over the last two decades were over-hyped, leading to unrealistic expectations that could not be attained, the end result being a resistance to further use of the technology. This cycle continues to repeat itself today. Bezdek [50] has nicely conceptualized what he terms "The Technology Cycle," which is depicted in Fig. 7. Initially, new technology typically elicits a "Naïve euphoria," which shortly reaches the "Peak of hype." However, since essentially any new technology cannot live up to its hype there is an inevitable "Overreaction to immature technology" that plunges inexorably into the "Depths of cynicism." Slowly "True user benefits" begin to emerge as the "Asymptote of reality is approached.

Ultimately, understanding of living systems will come from a holistic approach that involves both computational and experimental methods. We are now entering an exciting new phase in biological research, in general, and drug discovery, in particular, and we have many of the tools needed to gain a much more fundamental understanding of biological systems. However, the question remains: how long will it take to fully accomplish the paradigm shift to a more systems biology approach?

Acknowledgments Thanks are due to Susan Maggiora for her patient reading and editing of the original manuscript and to Andy Card for permission to use a picture of the Babbage difference engine he constructed totally from Legos™.

References

1. Wild D, Saqi AS (2004) Structural proteomics: inferring function from protein structure. Curr Proteom 1:59–65
2. Congressional Budget Office Study (2006) Research and development in the pharmaceutical industry. http://www.cbo.gov/ftpdocs/76xx/doc7615/10-02-DrugR-D.pdf. Accessed May 2, 2011
3. Brenner S (1999) Theoretical biology in the third millennium. Phil Trans R Soc Lond B 354:1963–1965
4. Powell LM, Lo M, Cole S, Trimmer J (2007) Application of predictive biosimulation to the study of atherosclerosis: development of the cardiovascular physiolab® platform and evaluation of CETP inhibitor therapy. In: Proceedings of the FOSBE, Stuttgart, Germany
5. Chawla PS, Kochar MS (2004) What's new in clinical pharmacology and therapeutics. Wis Med J 103:13–19
6. Ahn AC, Tewari M, Poon C-S, Phillips RS (2006) The limits of reductionisms in medicine: could systems biology offer an alternative? PLOS Med 3:709–713
7. Ahn AC, Tewari M, Poon C-S, Phillips RS (2006) The clinical applications of a systems approach. PLOS Med 3:956–960
8. Van Regenmortel MHV (2004) Reductionism and complexity in molecular biology. EMBO Report 5(11):1016–1020
9. Reductionism in Biology (2008) Stanford Encyclopedia of Philosophy, Metaphysics Research Lab, CSLI, Stanford University. http://plato.stanford.edu/entries/reduction-biology/. Accessed May 2, 2011
10. Polanyi M (1967) Life transcending physics and chemistry. Chem Eng News 45:54–69
11. http://acarol.woz.org/difference_engine.html. Accessed May 4, 2011

12. Otter T (2004) Toward a new theoretical framework for biology. GECCO 2004 Workshop on self-organization in evolutionary algorithms (SOE), pp 1–9
13. Evans GA (2000) Designer science and the "omic" revolution. Nat Biotech 18:127
14. Huang S (2000) The practical problems of post-genomic biology. Nat Biotech 18:471–472
15. Weinstein JN (2002) 'Omic' and hypothesis-driven research in the molecular pharmacology of cancer. Curr Opin Pharmacol 2:361–365
16. Rashevsky N (1960) Mathematical biophysics: physico-chemical foundations of biology, vol 2 (Revised Ed.). Dover Publications Inc, New York, p 308
17. Wikipedia (2011) Occam's razor. http://en.wikipedia.org/wiki/Occam's_razor. Accessed May 18, 2011
18. Wikipedia (2011) System identification. http://en.wikipedia.org/wiki/System_identification. Accessed May 12, 2011
19. Edelman GM, Galley GA (2001) Degeneracy and complexity in biological systems. Proc Natl Acad Sci USA 98:13763–13768
20. LePage DF, Conlon RA (2006) Animal models for disease: Knockout, knockin, and conditional mutant mice. In: Wang QK (ed) Cardiovascular disease, vol 2. Molecular medicine. Humana Press Inc, Totawa, New Jersey, Chapter 4
21. Wikipedia (2011) Small interfering RNA. http://en.wikipedia.org/w/index.php?title=Small_interfering_RNA. Accessed May 4, 2011
22. Hood L (2003) Quote in article by Henry CM on systems biology. Chem Eng News 81:45–55
23. Araujo RP, Liotta LA, Petricoin EF (2007) Proteins, drug targets and the mechanisms they control: the simple truth about complex networks. Nat Rev Drug Disc 6:871–880
24. Kitano H (2002) Computational systems biology. Nature 420: 206–210
25. Lee J-S, Chu I-S, Mikaelyan A, Calvisi DF, Heo J, Reddy JK, Thorgeirsson SS (2004) Application of comparative functional genomics to identify best-fit mouse models to study human cancer. Nat Genet 36:1306–1311
26. Yildirim MA, Goh K-I, Cusick ME, Barabasi A-L, Vidal M (2007) Drug-target network. Nat Biotech 25:1119–1126
27. Paolini GV, Shapland RHB, van Hoorn WP, Mason JS, Hopkins AL (2006) Global mapping of pharmacological space. Nat Biotech 24:805–815
28. Hopkins AL (2007) Network pharmacology. Nat Biotech 25: 1110–1111
29. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK (2007) Relating protein pharmacology by ligand chemistry. Nat Biotech 25:197–206
30. Janga SC, Tzakos A (2009) Structure and organization of drug-target networks: insights from genomic approaches to drug discovery. Mol Biosys 5:1536–1548
31. Pujol A, Mosca R, Farrés J, Aloy P (2010) Unveiling the role of network and systems biology in drug discovery. Trends Pharmacol Sci 31:115–123
32. Phillips A, Cardelli L (2007) Efficient, correct simulation of biological processes in the stochastic pi-calculus. In: Calder M, Gilmore S (eds) Computational methods in systems biology. Springer, New York, pp 184–189
33. Wendler A, Wehling M (2010) The translatability of animal models for clinical development: biomarkers and disease models. Curr Opin Pharmacol 10:601–606
34. Varmus H (1997) Testimony on the value of clinical research. Before the Senate Committee on Labor and Human Resources, Subcommittee on Public Health and safety. http://www.hhs.gov/asl/testify/t971009a.html Accessed May 9, 2011
35. Owens J (2006) 'Dirty drugs' secrets uncovered. Nat Rev Drug Disc 5:1
36. Mac Donald ML, Lamerdin J, Owens S, Keon BH, Bilter GK, Shang Z, Huang Z, Yu H, Dias J, Minami T, Michnick SW, Westwick JK (2006) Identifying off-target effects and hidden phenotypes of drugs in human cells. Nat Chem Biol 2:329–337
37. Stahl SM (2009) Multi-functional drugs: a novel concept for psychopharmacology. Trends Pyschopharm 14:71–73
38. Casti J (1982) Topological methods for social and behavioral systems. Int J Gen Syst 8:187–210
39. Bandyopadhyay S, Maulik U, Wang JTL (eds) (2007) Analysis of biological data: a soft computing approach. World Scientific, Singapore
40. Zadeh LA (1994) Soft computing and fuzzy logic. IEEE Softw 11:48–56
41. Zadeh LA (2001) Applied soft computing. Appl Soft Comput 1:1–2
42. Klir GJ, Yuan B (1995) Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, Upper Saddle River
43. Xu D, Keller JM, Popescu M, Bondugula R (2008) Applications of fuzzy logic in bioinformatics. Imperial College Press, London
44. Haykin S (1994) Neural networks: a comprehensive foundation. IEEE Press, Macmillan College Publishing Company, New York
45. Pearl J (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers, San Francisco, California
46. Fogel DB, Bäck T, Michalewicz Z (2000) Evolutionary computation 1: basic algorithms and operators. Taylor & Francis Group, New York
47. Almond RG (1995) Graphical belief modeling. Chapman & Hall, London
48. Pawlak Z (1991) Rough sets: theoretical aspects of reasoning with data. Kluwer Academic Publishers, Dordrecht
49. Pedrycz W, Skowron A, Kreinovich V (eds) (2008) Handbook of granular computing. Wiley, New York
50. Bezdek J (1993) Fuzzy models: what are they and why? IEEE Trans Fuzzy Syst 1:1–5