What's in a Structure?

While structural approaches are the rage in drug development, what makes a molecule a safe and efficient drug is best defined by functional properties within the context of a sick body. The belief in the adequacy of the structure–function hypothesis begs the question of what exactly we see in the structure of a molecule when we call it a drug.

Lukas K. Buehler
is the founder of SciScript Inc., in San Diego, California, USA. He also teaches at the University of California at San Diego's Extension Bioscience Program. He can be reached at lbuehler@ucsd.edu.

Structure determines function. This assertion is engrained deeply in the biologist's mind. In the biotechnology and pharmaceutical industries, analysis of structural information has become an inseparable part of the drug discovery process. Computer-assisted drug design (CADD), rational drug design and computational chemistry are just a few examples of the reliance on using structural information to solve functional problems. However, we find that the road to novel, non-toxic small-molecule therapeutics by rational approaches is impossible to travel without trial and error methods, though we don't know why.

This column will explore various aspects of the drug discovery process, comparing the promises with the limitations of rational design strategies. In this first installment, I will discuss the usefulness of the structural view as a reductionist model of function in biology and medicine.

During the past 50 years, X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy solution structures and computer modeling have produced astonishing insights into the structure–function relationship of biological macromolecules and their ligands. Watson and Crick implicitly referred to the copying mechanism of DNA replication in their groundbreaking 1953 *Nature* paper on the structure of the DNA double helix. Furthermore, I haven't met a scientist who is not intrigued by the explanatory power that comes from structural visualization. Yet, I believe that the structural view of function is deceptive if taken to mean that one can predict function from structure — in this case, a molecule functioning as a drug. The deception lies in the suggestion that structure is a surrogate of function, thereby confusing explanatory with predictive power.

Explaining What We Already Know

As humans, we inherently devise relationships between form and function in order to make sense of things. Consider the table as an example. Tables come in distinctly different yet closely related forms: coffee tables, desks, night stands, etc. Regardless of their specific function, each of these tables has a flat surface that can be used to put things on, such as food. But only the dinner table is used primarily for the purpose of perfunctory family dinners. So it seems that handling objects (tables) within the context of their proper surroundings (the family dining area) being used for their proper function (dining) makes form (the dinner table) understandable to humans. In biology we also need to realize that a structure explains a function we already know.

Indeed, biologists would not spend time determining the structure of a molecule were it not for its perceived usefulness. This is particularly true for the genome projects. Chief among the objections to the Human Genome Project in the late 1980s was the lack of usefulness (a string of letters without associated function), for fear of diverting funds for hypothesis-driven research. Today, this objection might seem surprising, as genome sequences have become treasure troves for functional genomics and bioinformatics. The structural organization of genomes already is shedding light on biologically important questions regarding evolution, development and disease.

Despite these insights, genome projects also have left us with the realization that the function of a large percentage of genes — in some cases as high as 40% — is unknown (1–2). This includes the genomes of even the most thoroughly studied organisms — the bacterium *Escherichia coli* and the fruit fly *Drosophila melanogaster*, for example. The functions of these genes remain unknown because their sequences have no relationship to the sequences of any known genes, demonstrating quite

clearly our current shortcomings in the ability to predict function from (primary) structure alone.

How Structure Informs about Function

Drug discovery is all about finding novel ligands to target defective receptors, enzymes or transport proteins. In the search for drugs, we have to focus our attention as much on the structure of a molecule as we do on its efficacy as a therapeutic agent. Molecules that function as drugs also are ligands of receptors. To call a drug a ligand asserts its binding property, which in turn can be explained by the chemical nature of both ligand and receptor surface structures. To understand the value of structural information in rational drug design, we obviously need to make a very careful distinction between the use of the terms "structure" (ligand) and "function" (drug). What we need to understand is how the meaning of each term is derived from the use of the other. Here we follow a simple logic: functional information conveys meaning to structures, not the other way around. All drugs are ligands, but not all ligands are drugs.

The structure–function relationship is a one-way street. When predicting function from structure, we do so with some guidance from similarities to known structures with known function. The pharmacophore concept — the mapping of common structural features of active analogs that bind to the same receptor — illustrates the importance of this finding. Faced with a completely novel structure, or amino acid sequence for that matter, we still lack the tools to confidently predict either function or structure of a protein. This is true even for the narrow application of predicting the ligand binding property of a novel molecule. When we look at the structure of a molecule and call it a ligand, we emphatically presume its receptor binding property (i.e., that it is a ligand, as determined by some experimental evidence, such as a high-throughput screening assay). We show a ligand's structure to explain how it binds.

Making Important Inferences

If not for predicting function, what exactly does structural information contribute to drug discovery that function alone cannot? For one, structure is a tool for synthesis in the hand of the chemist. For another, structure is an interpretation of a function in the hand of the designer — from structure to ligand, from ligand to drug. Most importantly, however, structure conveys a mechanism of the chemistry of a biological process, such as binding or catalysis. A mechanism is what Watson and Crick "saw" in their structure of the double helix: a molecular mechanism of inheritance.

We postulate that structure determines function, but we don't know exactly how. If structure is a representation of a mechanism at the molecular level, why should it be so difficult to predict function from it? The problem is rooted partly in a lack of understanding of the non-equilibrium dynamics, complexity and redundancy of biological systems. For example, it is a well-known observation that endogenous signaling molecules such as hormones and neurotransmitters can bind to more than one receptor. Not only are there multiple receptors for the same signaling molecule, there are multiple ligands for a single receptor, too. We know this because receptors can be distinguished pharmacologically (e.g., muscarinic receptors can be differentiated from nicotinic acetylcholine receptors). Considering that both endogenous signaling molecules and exogenous drugs (plant secondary metabolites, such as nicotine and muscarine) function as ligands, binding sites truly are multivalent entities. The general architecture and dynamics of proteins explain the many possible interactions, even with small ligand molecules.

This functional variability associated with proteins is not surprising, as it is the raw material of evolution, allowing for occasional changes within a fairly robust and stable "sequence space" that resists abrupt change due to a built-in redundancy and natural selection. From an information point of view, one could say that the code of life does not show the brittle behavior of computer language (software code) — where a single error causes an entire program to malfunction. Realizing that the immense sequence space encoded for by amino acid sequences is the foundation of robustness in biological systems (of course, there are mutations causing non-redundant defects — the very reason we develop drugs), it becomes easier to accept why it is anything but trivial to infer a biological activity of a novel small molecule, randomly pulled from a combinatorial chemistry library, from merely looking at its structure.

While reductionism is a powerful approach, with molecular structures being a visualization of the mechanism, the structure–function relationship of biological systems' components tells us little about the behavior of networks they belong to. Design is not strictly the reverse of the structure–function analysis of biological systems. I'll discuss this difference in a forthcoming column, using examples from bioinformatics and systems biology — as integrative sciences, both eventually will enable us to predict if certain molecules will behave as drugs or not.

References

1. P. Glaser et al., *Science* **294,** 849 (2001).
2. G.M. Rubin, *Nature* **409,** 820 (2001).**PG**