Mathematical Molecular Bioscience and Biophysics

A Recurring Theme at the SIAM Conference on the Life Sciences

By Guo-Wei Wei

How effectively does a potential drug bind to its target biomolecule? Mathematics has the answer! The SIAM Conference on the Life Sciences (LS16), held in Boston from July 11-14, has, for the first time in its history, highlighted mathematical molecular bioscience and biophysics (MMBB) as a theme. MMBB concerns the development of mathematical theories, models, methods, schemes, and algorithms for elucidating molecular mechanisms and for solving open problems at the forefront of molecular biosciences and biophysics, such as those associated with drug design and discovery. All areas of mathematics—including differential equations, functional analysis, harmonic analysis, Lie group, Lie algebra, geometry, graph theory, and topology—are essential to MMBB and play a key role in addressing fundamental challenges in molecular biosciences and biophysics. One of these challenges is the emergent complexity in self-organizing biomolecular systems, such as HIV or Zika virus, molecular motors, Alzheimer's disease, and cancer cells. Mathematical approaches, such as multiscale modeling, invariant manifold, compressed sensing, and machine learning techniques, are becoming increasingly popular in molecular biosciences due to their ability to efficiently reduce the number of degrees of freedom while still maintaining an essential and adequate description of the biomolecules of interest [1, 2].

An important trend in contemporary life sciences is the fundamental transition of traditional disciplines, such as physiology, population biology, evolutionary biology, neuroscience, etc., from macroscopic and phenomenological subjects to molecular-based biosciences. Parallel to this development, the life sciences in the 21st century are transforming from qualitative and descriptive disciplines to quantitative and predictive ones, which are based on molecular mechanisms (the ultimate truth of biological sciences). This transformation has led to the burgeoning of MMBB, an emergent field in mathematics that generates mathematically-driven advances in molecular biosciences.

LS16 featured nearly 40 MMBB minisymposia organized by leading researchers in MMBB, covering various exciting advances including charge transport, ion channels, membrane modeling and computation, multiscale modeling of solvation, electrostatics computing and applications, topological and geometric methods for biomolecules, and macromolecular structures and interactions. The achievements in mathematical approaches for drug design and discovery are particularly worth noting.

Designing efficient drugs for curing diseases is especially important for life sciences in this century. Indeed, one of the ultimate goals of molecular bioscience and biophysics is to understand the molecular mechanism of human diseases and to develop efficient drugs—free of side effects—for disease treatment. The principal task of drug design and discovery is to predict whether a given molecule will bind to a biomolecule, such as a protein or DNA, and activate or inhibit its function, which in turn results in a therapeutic benefit to the patient. Typical drugs are comprised of small organic molecules, but biopolymer and protein-based drugs are becoming increasingly common. An ideal drug should be acceptable to the human metabolic system and bind firmly to the target, without affecting any other important "off-target" molecules or antitargets similar to the target molecule. Nevertheless, drug design and discovery involve an extremely complicated procedure that includes the following: disease identification, target hypothesis (the activation or inhibition of drug targets), screening of potential drugs that can effectively bind to the target, optimization of the structures of identified drugs, preclinical *in vitro* and *in vivo* tests, clinical trials to determine bioavailability and therapeutic potential, and optimization of a drug's efficacy, toxicity, and pharmacokinetic properties.

Computer-aided drug design and the design of protein containers for drug delivery have a proven record of success, not only because of improved understanding of the basic science—the molecular mechanism of drug and protein interactions—but also because of advances in mathematical modeling, geometric representations, topological characterization, graph theory analytics, computational methods, optimization procedures, machine learning algorithms, and the availability of massive parallel and graphics processing unit (GPU) computers. Indeed, mathematics plays an essential role in rational drug design, from the identification of drug binding hot spots, consensus scoring, geometric analysis, cluster analysis, and global optimization to drug efficacy, toxicity, and pharmacokinetic analysis.

Moreover, mathematical approaches—such as geometric analysis for high throughput drug screening, persistent homology for protein-drug binding detection, reduced manifold representation for discriminating false protein-protein and protein-drug interfaces, and machine and manifold learning techniques for protein-drug binding site analysis—greatly impact drug design and discovery. Specifically, these approaches lead to better homology modeling, geometric models, molecular docking algorithms, molecular dynamics, quantum calculation, *de novo* design, and statistical models for efficient drugs and functional proteins. Figure 1 illustrates the Pearson correlations between experimental protein-ligand binding affinities and various theoretical predictions. A mathematical approach called feature functional theory-binding predictor (FFT-BP) outperforms all the other eminent methods in molecular biophysics.

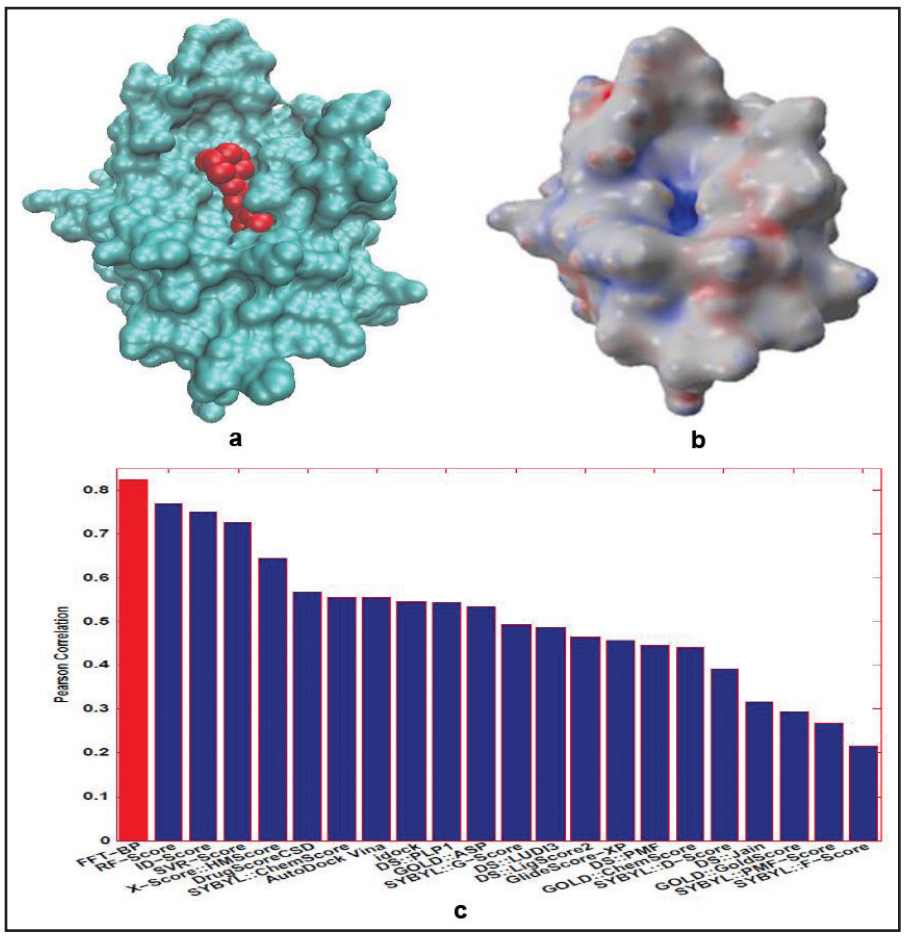

Figure 1. *Illustration of mathematical approaches to drug design and discovery.* **1a.** *Geometric representation of protein-ligand binding. The protein is green and the ligand is in red.* **1b.** *The binding site (blue) predicted by the product of minimal curvature and electrostatic potential obtained from the differential geometry-based Poisson-Boltzmann equation.* **1c.** *Comparison of Pearson correlations of various predictions and experimental binding affinity data for the PDBBind 2007 core set of 195 complexes. The winner, feature functional theory -binding predictor (FFT-BP), is based on machine learning and involves geometry, topology, graph theory, partial differential equations, and advanced numerical algorithms.*

The cutting edge of FFT-BP for drug design and discovery is a manifestation of the ever-increasing impact of mathematics on molecular biology and biophysics. There is enormous potential in this area for integrative interdisciplinary research in which mathematicians and experimentalists develop solutions to challenging problems in tandem. Driven by the advances in quantitative and predictive life sciences, MMBB will provide unprecedented opportunities to mathematicians for generations to come.

References

[1] Demerdash, O.N.A., Daily, M.D., & Mitchell, J.C. (2009). Structure-based predictive models for allosteric hot spots. *PLOS Computational Biology*, *5*, e1000531.

[2] Wei, G.W. (2010). Differential geometry based multiscale models. *Bulletin of Mathematical Biology, 72*, 1562-1622.

Guo-Wei Wei is a professor of mathematics at Michigan State University.