



Table of Contents

Preface

Summary Article

Individual Contributors

Statistics as the information science

Statistical issues for databases, the internet, and experimental data

Mathematics in image processing, computer graphics, and computer vision

Future challenges in analysis

Getting inspiration from electrical engineering and computer graphics to develop interesting new mathematics

Research opportunities in nonlinear partial differential equations

Risk assessment for the solutions of partial differential equations

Discrete mathematics for information technology

Random matrix theory, quantum physics, and analytic number theory

Mathematics in materials science

Mathematical biology: analysis at multiple scales

Number Theory and its Connections to Geometry and Analysis

Revealing hidden values: inverse problems in science and industry

Complex stochastic models for perception and inference

Model theory and tame mathematics

Beyond flatland: the future of space

## Mathematics in Molecular Biology and Medicine

D. SUMNERS

Two major scientific revolutions began in the mid-twentieth century, one in computation/information processing and one in molecular biology. Spurred by wartime computational needs at Los Alamos, the electronic computer was born. A decade later at the Cavendish Laboratory in Cambridge, the landmark discovery of the DNA double helix was the key that opened the door and signaled the beginning of the revolution in molecular biology. In the fifty years since then, exciting scientific breakthroughs in both areas have been made at a mind-numbing rate. On the one hand we have relatively enormous and cheap computing power, and on the other we have the sequencing of the complete human genome. In terms of prospects for future progress, molecular biology is positioned at the beginning of the twenty-first century where physics was at the beginning of the twentieth century. Biology enjoys exquisite experimental ability and control, teasing out ever-deeper secrets. Simultaneously, computerized instrumentation and automation of experimental processes allow acquisition of biological data at an exponentially increasing rate. Biology now stands poised to climb the mountain of the understanding of life. Success of this climb depends on the increasing involvement of mathematics. Biology needs a great and immediate increase in the number of mathematicians inspired by biological problems, and a corresponding increase in the number of biologists appreciative of the potential for mathematical and computational techniques to spur scientific progress. Mother Nature does not give up her secrets easily; data does not come with instructions. Mathematics builds models which connect sparse data points with threads of logical argument, weaving these logical threads together to produce a fabric of understanding. Computation based on theory is essential in filling in gaps in understanding, extending the ability of the human mind to organize and comprehend the massive biological datasets now being produced. In order to drink from this firehose of biological data and convert some of it into knowledge, mathematics and computation (both old and new) are needed to build models and navigational tools.

The scientific heir apparent for the rapid increase in knowledge in molecular biology is medicine. The human body is an extremely complicated biological system. A goal now being formulated in research medicine is the complete human in silico. This dream is to generate interoperable computational models for human biology spanning all scales, from molecular to cellular to organs to organ systems. Access to such models would, for example, greatly enhance rational drug design, allowing computational testing of hypothetical drugs. It would greatly enhance understanding of "normal" function in human biological systems and the non-invasive diagnosis of disease states, using computational comparison of subject anatomy and function to template anatomy and function. Mathematics has maximal impact on molecular/cellular biology and medicine via the feedback loop of interaction: mathematical models are built in response to unsolved problems and experimental/clinical results; the theory is converted into algorithms for machine computation; computations and theory are used to make predictions and analyze experimental data; results of this analysis are used to refine the theory and computation.

As an example of impact of mathematics on biology, consider the spectacular recent advances in computational genomics. The process of sequencing the 3 billion base pair human genome is achieved through sequencing hundreds of thousands of smaller randomly overlapping contiguous genetic segments. In order to arrange these segments in the biologically correct linear order, one must compute the overlap probability among segments. After one has assembled the bits of the genome and has the correct linear sequence, one must then find the genes, the short, scattered regions of DNA which code for currently unknown proteins. This is a mathematical problem of pattern recognition and biological cryptology, using every mathematical trick in the book and some yet to be

[and time](#)

**Mathematics in molecular biology and medicine**

**The year 2000 in geometry and topology**

**Computations and numerical simulations**

**Numbers, insights and pictures: using mathematics and computing to understand mathematical models**

**List of Contributors with Affiliations**

invented. The mathematical nature of this task is the reason why every drug company in the world is hiring mathematicians and computer scientists, generating very stiff competition for mathematicians fluent in biology. After one has the DNA sequence for a gene, and hence the amino acid sequence for the encoded protein, one must then compute the three-dimensional structure (native fold) for the protein, and from this structure, (hopefully) deduce function. Current attempts at solving the well-known protein-folding problem rely on ab initio mathematical modeling and computation, and bootstrapping using annotated protein databases. After finding the 3D protein structure (or perhaps avoiding this step entirely), one must then find the function of the protein. For this, one uses 3D information, bootstrapping via database comparison of sequence and structure information with proteins of known sequence, structure and function, microarray chip technology which tells one that the gene which encodes the unknown protein is active or inactive in concert with genes that encode proteins of known structure and function, experiments based on geometric and topological assays for mechanism, etc. Determining structure and function of proteins is difficult, and this new science has been termed proteomics. Proteins rarely act alone; they act in concert with numerous other proteins. Understanding how proteins "talk" to each other and elucidating the cascade of regulation and function of proteins in the cell is the next hurdle. In every step of this climb of discovery, mathematics/computation has been and will be a major player.

One role of mathematics in medicine at the microscopic level is to "peer behind the curtain," using mathematics to compute the structure and function of life-sustaining enzymes that operate on DNA. The same enzymes that sustain life are also involved in life-threatening diseases, such as cancer; understanding structure and function opens the door to therapy. Microarray chip technology will soon permit disease treatment chemotherapy protocols to be tailor-made for the individual patient. The mathematics for microarray chip design and data analysis is now under intense development. At the macroscopic level, a fundamental problem in medicine is to understand and define "normal" anatomy and function for an organ. Due to high variability, comparison of anatomical and functional information across individuals and across groups of individuals requires sophisticated mathematical models. Mathematics can elucidate the function of vital organs, such as the heart and the brain. In the heart, detailed knowledge of heart geometry and muscle fiber orientation can be used to build sophisticated models, where the onset of fibrillation can be studied. Mathematicians are using these sophisticated models to design better defibrillators. In the brain, mathematics is useful in relating brain architecture (as revealed by high-resolution MRI scans) to brain function (as revealed by Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) scans). For example, one can use a modern realization of a 150-year-old theorem (the Riemann Mapping Theorem) to produce unique conformal flattenings of the surface of the brain, helping scientists to use canonical surface-based coordinate systems to compare brain function across individuals.

In summary, mathematics has been and will always be a major player in the team effort to understand the complexities of molecular biology and to harness this understanding to enhance human health.

Last Modified:  
Oct 17, 2001

[Previous page](#) | [Top of this page](#) | [Next page](#)

[Policies and Important Links](#)

| [Privacy](#)

| [FOIA](#)

| [Help](#)

| [Contact NSF](#)

| [Contact Web Master](#)

| [SiteMap](#)



The National Science Foundation, 4201 Wilson Boulevard, Arlington, Virginia 22230, USA  
Tel: (703) 292-5111, FIRS: (800) 877-8339 | TDD: (800) 281-8749

Last Updated:  
10/17/01  
[Text Only](#)