BIOPOLYMER STRUCTURES: WHERE DO THEY COME FROM? WHERE ARE THEY GOING? EVOLUTIONARY PERSPECTIVES ON BIOPOLYMER STRUCTURE AND FUNCTION

RICHARD A. GOLDSTEIN
Chemistry Department, Biophysics Research Division, University of Michigan
Ann Arbor, Michigan, USA - 48109 - 1055
richardg@umich.edu

ERICH BORNBERG-BAUER
Abt. Theoretische Bioinformatik, Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 280, D - 69 120 Heidelberg
and
Institut für Mathematik, Universität Wien, A - 1090 Wien
Bornberg@dkfz-heidelberg.de

Biological polymers such as DNA, RNA, and proteins are the result of eons of evolution. In order to understand these macromolecules, it is necessary to understand the evolutionary pressures that determined their form and function. Conversely, the properties of these macromolecules encode this heritage, and can provide insight into the process of evolution.

The effort to understand, reconstruct, and simulate the evolution of biopolymers, including their structure and function, has become a key task involving such diverse fields as evolutionary biology, molecular biology, structural biology, and biophysical chemistry. The rapid growth of sequence and structure databases, combined with the continuing geometric increases in computer speed, has made this field fertile ground for biocomputing and bioinformatics. In this session a number of diverse contributions to this effort are presented.

There has been much interest in using simple models to investigate basic principles that govern the relationship between biopolymer sequence and structure. In their paper, Abkevich, Gutin, and Shakhnovich develop a model where proteins are confined to cubic lattices. They apply optimization procedures to find evolved sequences that fold faster than initial random sequences, and show that the ability of proteins to fold may actually have arisen as a side-effect of other physiochemical factors.

In their contribution, Renner and Bornberg-Bauer use the popular "HP-type" lattice model to explore the fitness landscapes of lattice proteins. By generating large ensembles of random structures, they can characterize the relationship of sequence to structure using statistical methods. They find structure landscapes to be rugged, and this ruggedness to be sensitive to alphabet

and shape space size. This suggests that simple HP-models may be too crude for realistic analyses of evolutionary issues.

Many applications, such as identifying and aligning homologous proteins, require the use of some sort of "similarity matrix" which measures how much different amino acid types resemble each other. In their paper, Wei, Altman, and Chang develop a novel approach to the generation of these matrices, looking at similarities in the *environments* in which different amino acids are more likely to be found. In addition to enabling homology searches with a sensitivity comparable to that obtained using matrices based on relative substitution rates, this approach helps us to understand the relationship between amino acid property and local environment.

The exceedingly growing abundance of sequence data poses the need for new and efficient methods to scan data bases quickly and efficiently. Matsuda, Taniguchi, and Hashimoto introduce a method to tackle the notorious sequence-to-structure alignment problem. They present a simple algorithm to represent structures by mapping dihedral angles to an alphabet of 20 letters. Using Smith-Waterman local alignment algorithms and appropriate scoring matrices, their method appears to be more sensitive than standard procedures for performing motif searches.

The evolutionary process may leave intriguing residual clues that can be uncovered by increasingly sophisticated statistical methods. Korotkov and Phoenix present a method based on information theory to look for latent periodicities in the DNA sequences, and suggest that such periodicities may be quite widespread. In addition to suggesting duplication events during evolution, these features may also have certain physiologically-important roles.

Acknowledgments:

We are thankful to all researchers who submitted work to this track. We are indebted to the reviewers who found the time to provide useful comments and criticisms.