

## **Protein Engineering By Means of Statistical Analysis**

Yuval Nov, Department of Statistics, Haifa University

The talk reports advances in two statistics-based methods for protein engineering. Both methods are aimed to save laboratory time, resources, and labor.

The first method's goal is to combine optimally beneficial mutations of a protein. This method is based on a stochastic sequence-activity model, devised specifically for proteins, whose parameters are estimated from experimental data, and which allows prediction of highly-active protein variants. Two empirical applications of this method are discussed.

The second method is concerned with new criteria for determining the sample size in a protein engineering technique called saturation mutagenesis. By using a criterion of this type, one may reduce significantly the sample size while compromising minimally the quality of the best variant discovered. The method is also used to compare the efficiency of four randomization schemes, which underlie the saturation mutagenesis protocol.