

### **Chasing the Function of Hypothetical Proteins**

Osnat Herzberg<sup>a</sup>, Gary Gilliland<sup>b</sup>, John Orban<sup>a</sup>, Andrew Howard<sup>c</sup>, John Moulton<sup>a</sup>, <sup>a</sup>*Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, MD, USA.* <sup>b</sup>*Center for Advanced Research in Biotechnology, National Institute for Standards and Technology, Rockville, MD, USA.* <sup>c</sup>*Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, Chicago, IL, USA.* E-mail: osnat@carb.nist.gov

The focus of the Structure2Function program at CARB during the first five years was on prokaryotic proteins of unknown function, mostly from *Haemophilus influenzae* and *E. coli*. The goal was to leverage the 3-D structural information to gain insight into the biochemical function. Approximately 75% of the new structures exhibit folds that have been seen before, and in many cases we were able to make new discoveries about function by utilizing standard assays and mining genome context. The remaining 25% structures defined new folds, a somewhat higher fraction of new folds compared with structures entering the PDB. In some of these cases, the function became known around the time the structure was determined, while in other cases we were able to make broad predictions about the function.

**Keywords:** structural genomics, hypothetical proteins, crystallography