Validation and Classification of Protein Structures

Manfred J. Sippl, Christian Weichenberger, Markus Wiederstein, Stefan Suhrer, *Center of Applied Molecular Engineering, University of Salzburg, Austria.* E-mail: sippl@came.sbg.ac.at

We discuss and summarize several new developments regarding the automated validation of protein structures, the decomposition of protein structures into domains, and the classification of protein domains.. As a specific example we present results on the assignment problem of oxygen and nitrogen atoms in the side chains of Glutamine (GLN) and Asparagine (ASN) in some detail. These atoms are difficult to distinguish in the interpretation of electron densities and it is known [1] that approximately 15% to 20% of the assignments in all known structures are incorrect.

We demonstrate how mean force potentials [2] derived from a set of high resolution PDB [3] protein structures can be used to recognize and correct erroneous N/O assignments. Since the potentials are derived from erroneous data sets this is an interesting and challenging problem for the development of potential functions. We show that within a few cycles of potential compilation and error correction the potentials converge to a stable functional form. The detected erroneous assignments fully agree with expert curated assignments [4]. The ASN/GLN flipper is available as a WEB service at http://services.came.sbg.ac.at/flipper.

[1] Hooft R.W.W et al, *Proteins*, 1996, **26**, 363. [2] Sippl, M.J., *Proteins*, 1993, **17**, 355. [3] Berman, H.M. et al, *Nucleic Acids Res., 2000,* **28**, 235. [4] Word, J.M. et al, *J. Mol. Biol.*, 1999, **285**, 1735.

Keywords: structure validation, domain assignment, structure comparison, potential of mean force