Poster Presentation

MS96.P06

ValiFrag: Evaluating fragment quality during automated protein model building

J. Pereira[1], T. Wiegels[1], V. Lamzin[1]

[1]*European Molecular Biology Laboratory, Hamburg Outstation, Hamburg, Germany*

X-ray diffraction data from flexible macromolecules and their complexes can rarely be measured to a resolution better than 3 Å. Due to a loss of detectable atomic features, the determination of low-resolution structures is beyond the current operational range of crystallographic software and requires a large amount of manual intervention. ARP/wARP [1] v7.4 generates structures that are up to 80% complete at 3.0 Å, but the completeness drops sharply as the resolution gets worse. Reduction of the model completeness is accompanied with an increase in the number of fragments built, which become shorter. Such fragments are applicable for further model building if they are correct. Though, if they are wrong they may cause the formation of incorrectly built regions in the final model. Thus, there is a need to improve fragment quality before automated model completion is applied. We exploit the vast amount of structural information deposited in the Protein Data Bank (PDB) [2], to make use of it for structural validation of built fragments. Precisely, we evaluate the conformation of each fragment. If the conformation is present in several different protein models in the PDB, it is likely to be modelled correctly in the built model and is accepted. If, on the contrary, it cannot be found in any PDB model, it is probably incorrect. Here we present the software implementation of this validation, called ValiFrag, which checks the validity of automatically built protein chain fragments by evaluating their occurrence in the PDB. Protein models from the PDB were broken into dipeptides and conformational parameters for each of these were then stored in a database. For each automatically built fragment, ValiFrag computes the probability of it to be correct according to the conformation of all possible dipeptides. It can, therefore, assess which fragments are likely to be structurally incorrect and should possibly be modified, or even removed, to improve the final model.

[1] *G. Langer, S.X. Cohen, V.S. Lamzin, A. Perrakis, A., Nature Protocols, 2008, 3(7), 1171-1179,* ***[2]*** *H.M. Berman, J. Westbrook, Z. Feng, G.Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Research, 2000, 28(1), 235–242*

Keywords: automated model building, fragment validation, software implementation