Automatic crystallographic structure determination with ARP/wARP

V.S. Lamzin[1], A. Perrakis[2]

[1] European Molecular Biology Laboratory (EMBL), Hamburg Outstation, Hamburg c/o DESY, Notkestraße 85, D-22603 Hamburg, Germany

[2] Netherlands Cancer Institute (NKI), Department of Molecular Carcinogenesis, Plesmanlaan 121 1066 CX, Amsterdam, The Netherlands

Development of underlying methodology for crystal structure determination of biological macromolecules has now become one of the central goals in the process of high-throughput structure determination. The Automated Refinement Procedure (ARP/wARP suite, Perrakis *et al*., 1999) for building and refinement of protein structures is in position to promote progress in automating the steps of deriving an essentially complete structural model, provided that the X-ray diffraction data extend to sufficient resolution. The time required for building a protein structure can be shortened from several man-days or even man-months to a few CPU hours.

ARP/wARP aims at delivering essentially complete models starting from electron-density maps alone or from partial models. ARP/wARP challenges the conventional view of refinement of a macromolecular crystal structure by automation of real space manipulation of the model, thus mimicking a user intervention *in silica* and re-defining the classical conception of a macromolecular model. The refinement of a *hybrid* model, a combination of automatically recognised and built polypeptide fragments with free atoms, constitutes the basic ARP/wARP concept. The polypeptides provide additional information in a form of restraints while the free atoms describe remaining, not yet accounted for, prominent features in the electron density. As refinement proceeds, the hybrid model converges towards a conventional protein structure: the polypeptide chains transform to an (essentially) complete protein chain while free atoms to solvent structure. ARP/wARP is tightly linked to the CCP4 suite and there is a productive cooperation with G. Murshudov, the author of the REFMAC refinement program.

The recent developments encompassed the development of the CCP4-based graphical user interface, global assessment of the X-ray data quality, development of new methods which exploit stereo-chemical properties of secondary structural elements and the use of approaches from the graph theory and constrained integer programming for chain tracing. The developments made ARP/wARP applicable for automatic model building in cases where native X-ray data (but not necessarily the phases) extend to around 2.5 Å resolution. The new version 6.0 incorporating these features was released in July 2002 and it proved to be faster and more powerful.

Perrakis, A., Morris, R.M. & Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458-463.