Notes

s3.m2.o5 **ARP/wARP: Procedures for automated model building and refinement.** A. Perrakis[1], R. Morris[2] and V. Lamzin[2]. *European Molecular Biology Laboratory (EMBL) [1] Grenoble Outstation, [2] Hamburg Outstation*
Keywords: computing, refinement, modelling.

We present tools to automatically and without user intervention build and refine protein models, starting from reasonable estimates of crystallographic phases at resolution better than ~3.0 Å, molecular replacement solutions or partial models. Native diffraction data extending to a resolution better than 2.3 Å are necessary for completely automated building. Model improvement, better than that of conventional refinement techniques, was possible with data extending to resolution as low as 3.2 Å.

All methods are implemented in the ARP/wARP suite (http://www.embl-hamburg.de/ARP) and are based on an iterative procedure of real space model autobuilding and update, combined with reciprocal space refinement. A central concept is the use of *hybrid* models, which consist of free atoms and protein fragments. These describe the electron density contents in an optimal way, accounting for all features of the electron density (free atoms) while providing the geometrical information (protein fragments) which give restraints necessary for efficient reciprocal space refinement.

Autotracing is based on the search for protein-like patterns, the chemically identical main chain fragments (Cα-C-N-Cα *trans* peptides) that compose linear non-branching chains. The result of the autotracing is usually a set of several main chain fragments. Docking in the amino acid sequence and building of the side chains in appropriate conformations follows. Based on the sequence position of docked fragments gaps between them can be filled in. Real space torsional refinement of the model is the final space.

The combination of several autotracing cycles with refinement allows an almost complete protein model to be obtained in a fully automated way. To date, around fifty protein structures have been reported to be built and refined using these procedures, within a few CPU hours.

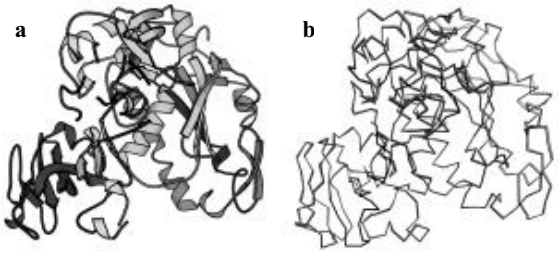

Figure 1. a: The model for *Leishmania* surface protein at the end of automated model building and refinement. Phases extended to 2.5 Å and native data to 2.0 Å. The model at the end of automated refinement comprised 462 residues (out of 472 in total) in 4 chains. The final Rfactor was 17.8 %. **b:** A Cα superposition of the autotraced (gray) and the final (black) models. The r.m.s. displacement from the final structure is only 0.28 Å.

Perrakis, A., Morris, R.J. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement, ***Nature Structural Biology* 6**, *458-463*