MS4-P5 The end point of model refinement in macromolecules; what are the coordinate errors?

John R. Helliwell[1], Manickam Gurusaran[2], Kanagaraj Sekar[2]

1. School of Chemistry, University of Manchester, Brunswick Street, Manchester, M13 9PL, UK
2. Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, Karnataka, 560012, India

email: john.helliwell@manchester.ac.uk

The end point of model refinement obviously requires a knowledge of the individual atomic position errors. The provision of such 'metadata' for the individual atomic coordinates has not been routinely done. As G Murshudov and E J Dodson state "without giving some estimate of the reliability of the model parameters the refinement procedure cannot be complete" [1]. The only refinement program which has an option to do this is the full matrix inversion option in SHELXL of G Sheldrick [2], usually applied at atomic resolution. We have adopted the Diffraction Precision Index (DPI) average for coordinate errors in macromolecules [3] to characterize the positional errors of individual atoms involved in non-covalent interactions [4] based on the sqrt (individual atomic B factor / the average B factor). It is used where restraints of a protein model refinement are not applied eg ion pairs or for metal ligands or protonation state determinations when a bond distance restraint is removed. As a web accessible tool we provide an ion pairs knowledgebase (http://cluster.physics.iisc.ernet.in/sbps). Now we also provide a webtool to calculate the error on any protein PDB file atom [described in detail with examples in ref 5]. A modified PDB file can be saved (http://cluster.physics.iisc.ernet.in/dpi/). In the preparation of structures for PDB deposition the distance errors for solvent molecules to their binding partners can now be used in evaluating clashes [5], where atomic position restraints have not been used. In cases where non-crystallographic symmetry has been used in a model refinement an appropriate user guidance message is provided by our webtool.

REFERENCES

1. G Murshudov and E J Dodson Simplified error estimation a la Cruickshank in macromolecular crystallography CCP4 Newsletter - January 1997.

2. Sheldrick, G.M. (1995) SHELXL-93 Institut fuer Anorg.Chemie, Goettingen, Germany.

3. D. W. J. Cruickshank (1999). Acta. Cryst. D 55, 583–601.

4. M. Gurusaran, M. Shankar, R. Nagarajan, J. R. Helliwell and K. Sekar (2014). IUCrJ 1, 74-81.

5. K. S. Dinesh Kumar, M. Gurusaran, S.N. Satheesh, P. Radha, S. Pavithra, K.P.S. Thulaa Tharshan, J R Helliwell and K. Sekar Online_DPI: A web server to Calculate the Diffraction Precision Index for a Protein Structure (2015) J Appl Cryst in press.

Keywords: Macromolecular atomic coordinate errors; Ion pairs; Metal ligand; Solvent clashes; DPI.

MS4-P6 A nuclease cut three ways: phasing from distant homologues, an ideal α-helix and Zn-SAD

Huw T. Jenkins[1], Alfred A. Antson[1]

1. York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5DD, UK

email: huw.jenkins@york.ac.uk

An in-situ proteolytic fragment of the large terminase from bacteriophage G20C produced crystals that diffracted to 1.45 Å resolution. The structure was initially solved by molecular replacement using an ensemble of the five available structures of nuclease domains from homologues with between 13 and 20% sequence identity. The MR solution was autobuilt to an almost complete poly-alanine trace with SHELXE and completed with ARP/wARP.

The structure contained five α-helices including a 13-residue long unbent α-helix so an attempt was made to solve the structure starting from a single ideal α-helix. Phaser successfully placed a 14-residue ideal α-helix and density modification by ACORN using phases calculated from this α-helix combined with artificially extended data to 1.0 Å enabled the structure to be automatically built using ARP/wARP. Unexpectedly, the placed α-helix corresponded not to the straightest α-helix of the nuclease domain but to the third longest α-helix of 11 amino acids.

It was clear from the maps that the active site of the nuclease domain contained a metal ion with reasonable anomalous signal at the wavelength used for data collection (0.9763 Å). Subsequent XRF revealed the presence of zinc (from the crystallisation conditions) and structure solution by Zn-SAD using SHELX from a dataset collected at 1.2824 Å was extremely straightforward.

These three approaches will be discussed together with some recent experiences with using low identity models for molecular replacement and also fragment-based MR approaches.

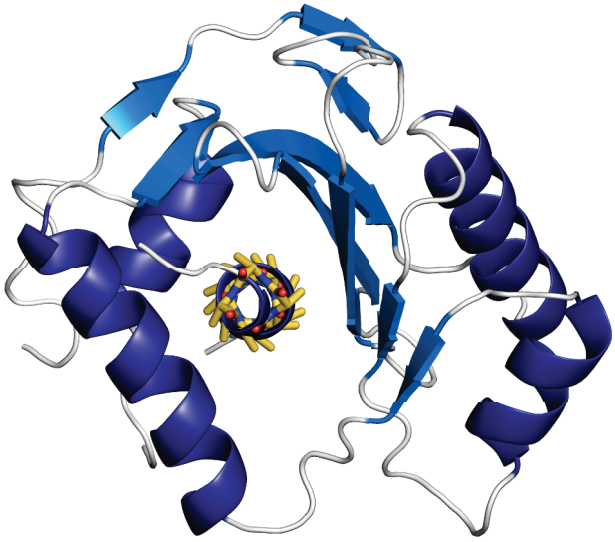

Figure 1. Structure of proteolytic fragment of G20C large terminase with 14-residue ideal α-helix used to solve the structure shown as yellow sticks.

Keywords: phasing, MR, fragments, SAD