MS44-O5 Use of clustering algorithms to combine partial solutions in reciprocal space

Claudia Millán[1], Massimo Sammito[1], Rafael J. Borges[1,2], Isabel Usón[1,3]

1. Dept. of Structural Biology, Institute of Molecular Biology of Barcelona (IBMB), Spanish National Research Council (CSIC), Barcelona, Spain
2. Dept. of Physics and Biophysics, Biosciences Institute (IBB), São Paulo State University (UNESP), Botucatu, São Paulo, Braz
3. Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

email: cmncri@ibmb.csic.es

Molecular replacement (MR) is the most frequent approach to solve the phase problem in macromolecular crystallography, but depends on the availability of a good enough starting model. Recently, MR is going ab initio in a number of approaches and as a consequence, its previous limits are being overcome through extremely sophisticated methods, targeting the area between purely ab initio and traditional MR, like AMPLE[1,2], MR_Rosetta[3], or the method proposed by Zhang[4]. ARCIMBOLDO[5] and BORGES[6] methods take the approach of using very small ideal models or libraries of local folds. Clustering methods had previous application in low resolution phasing approaches such as the Few Atoms Method[7]. More recently, Buehler et al. proposed the application of cluster analysis to MR phasing[8], pointing out that in cases in which partial or poor models are used, discrimination of the optimal position is not straightforward, and often, a number of weak peaks with partially correct orientations are found. Comparison of coordinates of partial solutions is not an easy task, as search models may fit the electron density in different ways. Instead, maps can be generated and its mean phase difference measured, producing a phase-combined map that may model differences and result in a better estimate of the atomic coordinates. We have applied phase combination to ARCIMBOLDO partial solutions and proof of principle has been established by studying optimal strategies starting with test cases. We have evaluated the reliability of possible figures of merit and how to best characterize differences (using mean phase differences or map correlation coefficients). These results, just submitted for publication, have been implemented to increase the radius of convergence and efficiency of the ARCIMBOLDO methods. The approach developed has been essential to solve a previously unknown structure with 430 aminoacids in the asymmetric unit and data to 1.5Å where classical MR methods had failed.

Bibby, J. *et al.* (2013). *Acta Cryst*. D69, 2194-2201
Bibby, J. *et al.* (2012). *Acta Cryst.* D68, 1622-1631
DiMaio, F. *et al.* (2011). *Nature*, *473*(7348), 540–543
Shrestha, R. *et al.* (2015). *Acta Cryst.* D71, 304-312
Millán, C. *et al.*(2015). *IUCrJ* 2, 95-105
Sammito, *et al*. (2013). *Nature Methods*, 10, 1099–1101
Lunin, V.Yu. *et al.* (1995). *Acta Cryst.*, D51, 896-903.
Buehler, A. *et al.* (2009). *Acta Cryst.* D65, 644-650

Keywords: Phasing, Clustering, Ab initio, Molecular Replacement

MS45. Chemical information as prior knowledge in crystallography

Chairs: Anthony Linden, Rob Nicholls

MS45-O1 Fitting ensemble models of disorder using a priori chemical structure information

Richard I. Cooper[1]

1. Department of Chemistry, University of Oxford, UK

email: richard.cooper@chem.ox.ac.uk

Finding and refining suitable crystallographic models of highly disordered regions of a structure is often a time consuming and thankless task: the improvement in fit and model phases gained by carefully fitting molecules to disordered regions of scattering density makes little difference to the interpretation of the 'important' parts of a crystal structure. Difficult cases can include: partially occupied solvent in voids, unidentified mixtures of disordered solvents, and solvent which is not constrained by the space group symmetry of the parent structure.

A more convenient option is to apply the bypass algorithm, implemented in PLATON as Squeeze.[1] The method adds a reciprocal space correction to the calculated structure factors based on the Fourier transform of the difference scattering density in the void regions of the crystal structure. Iterative improvement of phases using this scattering contribution can converge on a solution which gives a reliable correction to the data. Additional checks, such as the approximate number of scattering electrons in the solvent volume can help to rationalise the correction. The form of the scattering density used in the correction is not constrained, except that it must lie in the unoccupied regions (solvent accessible voids) of the structure. The results usually show an improvement in expected geometry and standard uncertainties of the remaining structure.[2]

An alternative approach is to make use of computing power to find a disordered model which gives the best fit to the data. Such a model can be constructed using an ensemble of partially occupied molecules placed in the solvent accessible regions of the structure and optimised using random sampling of different occupancies, conformations and positions. The sampling is optimised by using observed distributions of geometry and interaction from known structures.

The application of this method to some disordered structures is presented and shown to give a physically interpretable description of the scattering from a disordered solvent (Figure 1). Additional a priori real-space information, such as satisfying weak intermolecular interactions can also be used to guide solutions.

[1] P. van der Sluis and A. L. Spek (1990) *Acta Cryst* **A**46, 194-201.

[2] A. L. Spek (2015) *Acta Cryst* **C**71, 9–18

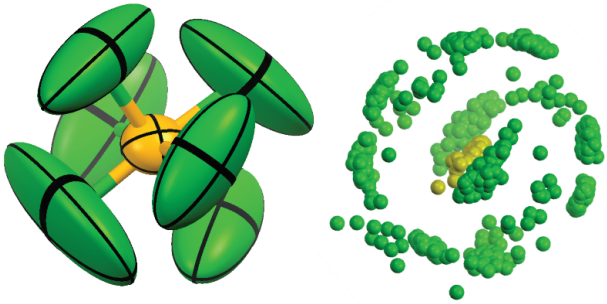

Figure 1. Conventional a.d.p. model of disordered PF_6 anion (left); ensemble of partially occupied molecular models (right) revealing more detailed orientational disorder.

Keywords: disorder, a priori chemical information, solvent

MS45-O2 Use of chemical restraints in Phenix

Nigel W. Moriarty[1]

1. Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

email: NWMoriarty@lbl.gov

The models obtained from crystallography can be improved by the use of prior chemical knowledge. Phenix (Adams, 2010) makes use of previously obtained experimental data and also predictive methods such as quantum chemistry (QC) to improve the accuracy of crystallographic models.

The electronic Ligand Builder & Optimization Workbench (eLBOW) (Moriarty, 2009) is used in Phenix to generate geometries and restraints for use in structure refinement. These can be determined by the using built-in semi-empirical QC methods, including RM1 and AM1, or via third party QC packages. It can also interface with the experimental geometries available via the Mogul program (Bruno, 2004). Libraries of restraints can be generated automatically, however, validation is essential to ensure accurate results. These range from simple checks to ensure topological correctness to detailed checks that highlight limitations of the generation method.

The Phenix structure refinement program, phenix.refine (Afonine, 2012), can make use of prior and predictive chemical methods in multiple ways to improve molecular models. The geometry of the protein can be improved using the Conformation Dependent Library (Moriarty, 2014) to adjust the main chain geometry based on the conformation of the backbone. Macromolecular geometry can also be improved by using more physically realistic potentials from molecular mechanics, such as Amber (Case, 2014). Finally, ligand geometries can be improved by the application of molecular modeling force fields (MMFF) or semi empirical QC (PM6) gradients from the AFITT program (Wlodek, 2006), or the use of QC methods from the DivCon library (Borbulevych, 2014).

References

Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, et al. 2010. *Acta Cryst. D* 66:213–21

Afonine, P.V., R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, et al. 2012. *Acta Cryst. D* 68:352–67

Borbulevych, O.Y., N.W. Moriarty, P.D. Adams & L.M. Westerhoff. 2014. *Comput. Cryst. Newsl.* 5:26–30

Bruno, I.J., J.C. Cole, M. Kessler, J. Luo, W.D.S. Motherwell, L.H. Purkis, et al. 2004. *J. Chem. Inf. Comput. Sci.* 44:2133–44

Case, D.A., V. Babin, J.T. Berryman, R.M. Betz, D.S. Cai, et al. 2014. *AMBER 14*. UCSF

Moriarty, N.W., R.W. Grosse-Kunstleve & P.D. Adams. 2009. *Acta Cryst. D* 65:1074–80

Moriarty, N.W., D.E. Tronrud, P.D. Adams & P.A. Karplus. 2014. *FEBS J.* 281:4061–71

Wlodek, S., A.G. Skillman, and A. Nicholls. 2006. *Acta Cryst. D* 62:741–49

Keywords: ligands, restraints, quantum chemistry, force fields, ligand libraries, validation