

## MS4-P8 Current and future potential of contact predictions for X-ray crystallography

Felix Simkovic<sup>1</sup>, Jens M.H. Thomas<sup>1</sup>, Olga Mayans<sup>1,2</sup>, Daniel J. Rigden<sup>1</sup>

1. Institute of Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, United Kingdom

2. Department of Biology, Universität Konstanz, 78457 Konstanz, Germany

email: hlfsimko@liv.ac.uk

Protein structures and protein-protein interactions are naturally constrained through contacting amino acids. Maintained under evolutionary pressure, a pair of contacting amino acids can remain invariant, accept only highly conservative changes over time, or accept more drastic mutations at one position when the other simultaneously alters in a coordinated fashion. Successful detection of these contacting residues, through the analysis of evolutionary covariance, is now reliably possible in suitable cases due to a new statistical treatment of the data that strongly reduces false positive predictions [1]. Such suitable cases are those where homologous sequences are sufficient in number and diversity - a limit expected to be surpassed by many more protein families in years to come [2]. Contact predictions provide great yet mostly unexplored opportunities to X-ray crystallography.

One benefit was recently demonstrated advancing ab initio Molecular Replacement (MR). The benefits of contact information in ab initio MR were successfully demonstrated using a software pipeline called AMPLE [3]. It was shown that a much larger target range is made tractable by guiding ab initio structure predictions with predicted residue contacts. Much larger - chain lengths up to 225 residues - and more  $\beta$ -rich protein targets were solved successfully using search model ensembles based on contact-guided ab initio structure predictions.

Additionally, further applications of contact data in the field of X-ray crystallography will be presented. Such applications can range from domain parsing to guide construct design [4] to the 3D-assembly of monomers into oligomeric search models for MR [5]. The use of contact information can also guide the interpretation of crystal structures by highlighting covarying and therefore functionally or structurally important interfaces or residues eg [6].

[1] Marks, D. S., Hopf, T. A., & Sander, C. (2012). *Nat. Biotechnol.* **30**, 1072–1080.

[2] Kamisetty, H., Ovchinnikov, S., & Baker, D. (2013). *P. Natl. Acad. Sci. USA* **110**, 15674–15679.

[3] Bibby, J., Keegan, R. M., Mayans, O., Winn, M. D. & Rigden, D. J. (2012). *Acta Cryst. D.* **68**, 1622–1631.

[4] Rigden, D.J. (2002). *Protein Eng.* **15**, 65–77.

[5] Ovchinnikov, S., Kamisetty, H. & Baker, D. (2014). *Life* **3**, e02030.

[6] Nicoloudis, J. M., Lau, S.-Y., Schärfe, C. P. I., Marks, D. S., Weihofen, W. A., & Gaudet, R. (2015). *Structure* **23**, 2087–2098.

**Keywords:** Molecular Replacement, AMPLE, Protein structure prediction, Evolutionary covariation

## MS4-P9 Combining unconventional MR with AMPLE *ab initio* models and experimental phasing

Adam J. Simpkin<sup>1</sup>, Olga Mayans<sup>2</sup>, Daniel J. Rigden<sup>1</sup>, Martin Savko<sup>3</sup>, William E. Shepard<sup>3</sup>, Ronan Keegan<sup>4</sup>

1. Institute of Integrative Biology, University of Liverpool, Liverpool, L69 7ZB, United Kingdom

2. Fachbereich Biologie, Universität Konstanz, D-78457 Konstanz, Germany

3. PROXIMA 2, Synchrotron SOLEIL, l'Orme des Merisiers, Saint Aubin, BP 48, 91192 Gif-sur-Yvette, France

4. Research Complex at Harwell, STFC Rutherford Appleton Laboratory, Didcot OX11 0FA, England

email: hlasimpk@liv.ac.uk

In recent years the combination of molecular replacement (MR) with single anomalous dispersion (SAD) via MRSAD, has been shown to solve structures that neither individual technique was able to [1, 2]. Whilst this success highlights the advantages of MRSAD, a partial model is required for MR which may not be available from conventional sources.

A potential way of generating these partial models is AMPLE [3]. AMPLE is a pipeline that uses a cluster-and-truncate approach on *ab initio* models to solve the structures of small proteins via unconventional MR. As *ab initio* modelling predicts the three dimensional structure of a protein independently of homolog structures, it can be used to make models for novel folds and structures. In the case of small proteins, the modelling of these features is often enough to solve them via MR, however in larger structures it may not. The modelling of features can nonetheless contain some useful information about the structure and in combination with experimental phasing information, could lead to a solution.

An attractive method of experimental phasing for crystallographers when a solution cannot be found by more traditional means is sulphur-SAD (S-SAD) as it requires no changes to crystal preparation. However, because S-SAD experiments are often demanding on data quality and because some proteins contain only a low sulphur content ratio, it is often hard to get phases out of S-SAD experiments. Indeed, as of 2015, only 80 de novo S-SAD structures have been deposited in the wwPDB [4]. Using AMPLE in combination with S-SAD therefore provides a potential route through which the success rate of S-SAD can be improved. We describe preliminary experiments exploring this approach, and enhancements to the underlying MR pipeline MrBUMP that facilitate the work.

[1] Busby, J. N., Lott, J. S., & Panjikar, S. (2016). *Acta Cryst. D.* **72(2)**, 182–191

[2] Skubák, P. & Pannu N. S., (2013). *Nat. Commun.* **4**, 2777

[3] Bibby, J., Keegan, R. M., Mayans, O., Winn, M. D., & Rigden, D. J. (2012). *Acta Cryst. D.* **68**, 1622–1631

[4] Weinert, T et al. (2015). *Nat. Methods.* **12**, 131–133

**Keywords:** Molecular replacement, Experimental phasing, AMPLE, MRSAD.