MS3-05 **Clustering procedures for the optimal selection of datasets from multiple crystals in macromolecular crystallography.** James Foadi,[ab] Pierre Aller,[b] Danny Axford,[b] Robin Owen,[b] Yilmaz Alguel,[a] Alex Cameron,[a] Wes Armour,[c] David Waterman,[d] and Gwyndaf Evans,[b] *[a]Imperial College London, UK, [b]Diamond Light Source Ltd, UK, [c]Oxford e-Research Centre, UK, [d]CCP4 – Research Complex at Harwell, UK*
E-mail: j.foadi@imperial.ac.uk

The availability of intense and collimated synchrotron sources and fast-reading detectors, with the increased likelihood of severe radiation damage, has enabled and forced the community of protein crystallographers to acquire large numbers of datasets of a same structures in a relatively short interval of time. A similar speeding up will have to be built in the processing of such multiple sets, or data collection efforts go wasted. A computer program, BLEND, has been implemented for this task. BLEND uses multivariate statistics, mainly in the form of cluster analysis, to bring together datasets with better merging likelihood. The program allows researchers to save time both in avoiding the combinatorial explosion implied in the analysis of multiple datasets and in the cumbersome and time-consuming amount of bookeeping that goes with it. This talk will highlight BLEND's workings and illustrate its ability to carry out optimal selection of groups of datasets with several real-data examples.

BLEND is currently installed at the Diamond Light Source synchrotron, where an increasing number of users employ it to sift through their datasets. BLEND has been used successfully in the solution of a novel membrane protein.

Keywords: multiple crystals; datasets; scaling; merging

MS4-01 **Molecular replacement and model-building using distant homology models as templates.** Thomas C. Terwilliger[a], Frank DiMaio[b], Randy J. Read[c], David Baker[b], Axel T. Brunger[d], Paul D. Adams[e], *[a]Los Alamos National Laboratory, Mailstop M888, Los Alamos, NM 87545, USA, [b]University of Washington, Department of Biochemistry, Seattle, WA, 98195, USA, [c]University of Cambridge, Department of Haematology, Cambridge Institute for Medical Research, Cambridge, CB2 0XY, UK, [d]Stanford University, Departments of Molecular and Cellular Physiology, Neurology and Neurological Science, Structural Biology, Photon Science, and Howard Hughes Medical Institute, 318 Campus Drive West, Stanford, CA 94305-5432, [e]Lawrence Berkeley National Laboratory, One Cyclotron Road, Bldg 64R0121, Berkeley, CA 94720, USA*
E-mail: terwilliger@lanl.gov

Molecular replacement is an immensely powerful method for determining macromolecular crystal structures. The method nevertheless remains limited by the requirement for a template structure that is similar to the structure to be determined. Normally a template with an rmsd of core main-chain atoms of about 1.5-2 Å is required. The applicability of molecular replacement would be greatly expanded if templates differing from the target structure by 2 to 3 Å rmsd could routinely be used. Here we describe three new approaches for carrying out molecular replacement with distant homology models.

The first approach [1] combines the power of Rosetta structure modeling with Phenix automated molecular replacement, model-building, density modification, and refinement to yield a new general-purpose and easy-to-use tool for crystallographic structure determination. Molecular replacement (MR) solutions are obtained with phenix.automr, rebuilt with Rosetta including electron density map information, and then are further rebuilt with phenix.autobuild. The combination of Rosetta rebuilding and phenix rebuilding is the key part of this method. This combination merges the benefits of structure-modeling, in which homology models can now be created that are more accurate than the templates used to create them, with crystallographic structure determination and refinement, in which models are built that are consistent with measured crystallographic structure factors.

The second approach takes advantage of the observation that many pairs of proteins have local structures that can be superimposed much more closely than can their complete structures. For example, a β-sheet in one protein may be locally similar to a sheet in another, but when the two structures are superimposed using all atoms, the sheets may be translated relative to each other. We have developed a method for applying local distortions to a structure (morphing it) using an electron density map to guide the distortions. This procedure allows local similarity to be maintained while global structure is changed. We show that morphing molecular replacement templates after placing them in their approximate locations in the crystallographic unit cell can greatly improve subsequent model-building.

The third approach is a combination of DEN (deformable elastic network) refinement and phenix.autobuild [2]. DEN refinement is also a general and easy-to-use tool for structure determination of difficult structures, especially for cases with low-resolution diffraction data and/or data with significant anisotropy. DEN-refinement allows local structure to be retained during refinement while secondary structural elements can undergo deformations and large regions such as domains can be shifted substantially. DEN refinement can be combined with morphing and autobuilding to yield a