

*SIMBAD: A Sequence-independent molecular replacement pipeline*

Adam James Simpkin<sup>1</sup>, Andrey Lebedev<sup>2</sup>, Ville Uski<sup>2</sup>, Charles Ballard<sup>2</sup>, William Shepard<sup>3</sup>, Daniel Rigden<sup>1</sup>, Ronan Keegan<sup>2</sup>

<sup>1</sup>University Of Liverpool, Liverpool, United Kingdom, <sup>2</sup>STFC, Rutherford Appleton Laboratory, Harwell, Oxford, United Kingdom,

<sup>3</sup>SOLEIL synchrotron, Gif-sur-yvette, France

E-mail: hlasimpk@liv.ac.uk

In protein crystallisation it is not uncommon for contaminant proteins to crystallise instead of the protein of interest. Such cases can remain unsolved for many years, because there is no way to determine whether the crystallised structure is a contaminant or if the model used for molecular replacement is not close enough to the target structure.

Here we present SIMBAD, a highly streamlined pipeline to solve crystal structures independently of sequence which can provide a route to solve for contaminant and novel structures alike. Indeed, a case in which a novel structure remained unsolved for 14 years was recently determined [1] using a similar approach. Furthermore, SIMBAD has solved cases of mistaken identity (swapped crystallisation trays) and structures of proteins that were unidentified and unsequenced at the time of crystallisation.

SIMBAD utilises a multi-core cluster to solve the crystal structure using a series of different approaches. First, it runs a lattice parameter search against every structure in the PDB. Structures crystallising with closely related lattices parameters are then trialled by conventional MR and refinement using MOLREP and REFMAC. If this approach fails then a search against common contaminants is carried out, again trialling the top ranked structures with MOLREP/REFMAC. Finally, if necessary, a novel method that ranks structures from a non-redundant domain database provided by MoRDa applies a modified version of the AMORE rotation function to score potential templates. The top-scoring structures are trialled in MOLREP/REFMAC once again.

SIMBAD is currently being adapted for use at synchrotron beamlines as a means of screening for crystals containing contaminant structures and to help to derive the phases of novel structures.

Keegan, R. et al. (2016). Acta Cryst, D72, 933-943.

**Keywords:** [Molecular replacement](#), [Sequence independent](#), [Contaminants](#)