Parallel Data Processing for High Throughput X-ray Structure Determination

Guenter Wolf, Herbert Axelrod, Henry van den Bedem, Hsiu-Ju Chiu, Mitchell D. Miller, Christopher L. Rife, Qingping Xu, Ashley M. Deacon, *Joint Center for Structural Genomics, Stanford University, Menlo Park, CA, USA*. E-mail: gwolf@slac.stanford.edu

The Structure Determination Core (SDC) of the Joint Center for the Structural Genomics (JCSG) has implemented a prototype system, Xsolve, which automates all of the processing steps needed to create an initial set of molecular coordinates from a dataset of diffraction images. The goal of Xsolve is to provide standardized, high quality data processing and automate the numerous time-consuming steps in the structure determination process. The current prototype produces a model that is over 95% complete in more than 80 % of the MAD cases with data to 2.5 Å or better.

Xsolve, a Java Message Service (JMS) based control system, can run on a Linux cluster different processing strategies in parallel, e.g. the data can be processed in several different space groups or MAD/SAD structure determination can be attempted using various wavelength combinations.

Xsolve supports a wide range of crystallography software programs, which can be used in parallel: data reduction with Mosflm, Denzo/HKL2000, XDS and Scala; heavy atom solution and phase determination with Solve, SHELXD/E and Sharp; phase improvement with Resolve and model building with Resolve and ARP/wARP.

The JCSG is supported by NIGMS/PSI (P50-GM 62411). SSRL operations are funded by DOE BES, and the SSRL Structural Molecular Biology program by DOE BER, NIH NCRR BTP and NIH NIGMS.

Keywords: automatic structure solution, MAD, software