International Tables for Crystallography (2006). Vol. F, Section 25.2.1.1.3, p. 695.

25.2. Programs and program systems in wide use

BY W. FUREY, K. D. COWTAN, K. Y. J. ZHANG, P. MAIN, A. T. BRUNGER, P. D. ADAMS, W. L. DELANO, P. GROS, R. W. GROSSE-KUNSTLEVE, J.-S. JIANG, N. S. PANNU, R. J. READ, L. M. RICE, T. SIMONSON, D. E. TRONRUD, L. F. TEN EYCK, V. S. LAMZIN, A. PERRAKIS, K. S. WILSON, R. A. LASKOWSKI, M. W. MACARTHUR, J. M. THORNTON, P. J. KRAULIS, D. C. RICHARDSON, J. S. RICHARDSON, W. KABSCH AND G. M. SHELDRICK

25.2.1. *PHASES* (W. FUREY)

The program package *PHASES* came into being in the mid-to-late 1980s when it evolved largely from a series of independent computer programs written during the preceding decade for use in the Veterans Administration Medical Center's Biocrystallography Laboratory in Pittsburgh, PA, USA. The predecessor programs each carried out a particular task required in the processing, phasing and analysis of diffraction data from macromolecules, but these programs were usually computer, space-group and sometimes even protein specific. In addition, the programs were often poorly documented, if at all, and made use of incompatible data formats such that a battery of 'conversion' programs were required for transmitting information. While this was pretty much the situation in most laboratories at the time, it nevertheless unnecessarily complicated protein structure determination, particularly by graduate students and new postdoctoral workers. To overcome these problems, the original programs were rewritten (frequently combining several programs into one), generalized for all symmetries, modified to use a simple, standardized format and extensively documented. As methodologies developed, new programs and procedures were added, graphics programs were included, and the resulting package was optimized for use on interactive graphics workstations, which were then becoming the main computing resource in most laboratories. The first 'official' *PHASES* release was described at an American Crystallographic Association meeting (Furey & Swaminathan, 1990), although versions of the package had been in local use within the Pittsburgh laboratories for the preceding four years. There have been several major releases since the first as new features and strategies were incorporated, and the package as it existed in 1996 was extensively described in a *Methods in Enzymology* article (Furey & Swaminathan, 1997).

25.2.1.1. *Overall scope of the package*

The *PHASES* package was designed to deal with the major problem in macromolecular structure determination, *i.e.*, phasing the diffraction data. The package is not completely comprehensive, as it excludes programs for initial data reduction (production of a unique set of structure-factor amplitudes from the raw measurements), molecular-replacement calculations (rotation and translation functions), model building (interactive graphic fitting) and complete structure refinement (restrained least squares, conjugate gradient, molecular dynamics *etc.*). There are already excellent programs and packages available to accomplish these tasks; instead, the *PHASES* package focuses on the initial phasing of diffraction data from macromolecules by heavy-atom- and anomalous-scattering-based methods. Also included are programs and procedures for phase improvement by noncrystallographic symmetry averaging, solvent flattening, phase extension and partial structure phase combination. The programs and additional procedure scripts allow one to start with unique structure-factor amplitudes for native and/or derivative data sets and generate electron-density maps and skeletons that can be utilized in popular graphics programs for chain tracing and model building. The major methods incorporated in the package are listed below and will be described in more detail later.

25.2.1.1.1. *Isomorphous replacement, anomalous scattering and MAD phasing*

Heavy-atom-based phasing by the methods of isomorphous replacement (Green *et al.*, 1954) and/or anomalous scattering (Pepinsky & Okaya, 1956) are initiated by reading one or more 'scaled' files into the program *PHASIT*, along with estimates of the heavy-atom or anomalous-scatterer positional, occupancy and thermal parameters. Each input file can contain either isomorphous-replacement, derivative anomalous-scattering or native anomalous-scattering data. MAD (multiple-wavelength anomalous diffraction) data are treated as both isomorphous and anomalous-scattering data, in which case one simply inputs the scattering-factor differences (real part) appropriate for the wavelengths comprising the 'isomorphous' data sets and the actual scattering factors (imaginary part) appropriate for the 'anomalous' data sets. All possible combinations of isomorphous-replacement data, conventional anomalous-scattering data and MAD data are allowed and can be used simultaneously during phasing.

25.2.1.1.2. *Solvent flattening and negative-density truncation*

Solvent flattening and negative-density truncation are carried out following the strategy developed by Wang (1985); however, a reciprocal-space equivalent of the automated solvent-masking procedure is used (Furey & Swaminathan, 1997; Leslie, 1987). In addition, during solvent-mask construction all density near heavy-atom sites is automatically ignored, leading to more accurate masks. The complete process is fully automated and carries out three solvent-mask iterations with at least 16 solvent flattening and phase combination cycles. Optionally, an arbitrary number of additional cycles can be carried out for phase extension. A program is provided to interactively examine or edit the solvent mask or to create the mask by hand if desired.

25.2.1.1.3. *Noncrystallographic symmetry averaging*

Noncrystallographic (NC) symmetry averaging cases (Rossmann & Blow, 1963; Bricogne, 1974) are treated in direct space by operating on 'submaps', *i.e.* arbitrary regions in an electron-density map encompassing all of the molecules to be averaged that are unique by true crystal symmetry. Many of the averaging programs were derived from routines originally written by W. Hendrickson & J. Smith and have been described earlier (Bolin *et al.*, 1993), but they were substantially rewritten for incorporation into the *PHASES* package. Programs are supplied to: generate and examine the required submaps; refine the NC symmetry operators; interactively create averaging envelope masks; average density within the envelopes; convert the submaps to full-cell maps; invert the modified maps; and combine the phases with those from another source. An automated procedure is provided to carry out a specified number of averaging and phase combination cycles in addition to solvent flattening and negative-density truncation. This procedure allows for a gradual phase extension, if desired, extending by one reciprocal-lattice point in each direction for a given number of cycles.

Copyright © 2006 International Union of Crystallography

references