25.1.3.5. *SCALA*

The *SCALA* program (P. R. Evans, 1993, 1997) scales together multiple observations of reflections, and (optionally) merges multiple observations into an averaged intensity. Various scaling models are implemented. The scale factor is a function of the primary beam direction, either as a smooth function of φ (the rotation angle), or expressed as batch (image) number. In addition, the scale may be a function of the secondary beam direction derived from the spatial coordinates of the measured spot on the detector. In this case, the scaling is an interpolated three-dimensional function similar to that described by Kabsch (1988). The merging algorithm analyses the data for outliers and gives detailed analyses. It generates weighted means of the observations of the same reflection, after rejecting the outliers.

Location: *SCALA* is part of the *CCP*4 suite (Section 25.1.2.4). Operating systems: UNIX and VAX/VMS. Type: source code and binary. Distribution: free academic.

25.1.3.6. *STRATEGY*

STRATEGY (Ravelli *et al.*, 1997) is a program that aids in designing data-collection strategy. It is used to determine the optimal starting spindle angle, using a one-circle diffractometer with a 2D detector, in X-ray data collection from crystals of macromolecules. The input file of the program contains information, from a *DENZO* intensities x-file, about starting crystal orientation and cell parameters. The program simulates all the reflections that can occur during 360° rotation of the crystal, determines if reflections can be recorded on the detector, sorts them, provides pictures of the needed oscillation range as a function of the starting spindle angle for given degrees of completeness of the data set and produces redundancy tables for the shortest data collection possible for each desired completeness. However, neither mosaicity nor overlaps are taken into account. This program has been integrated into the *MOSFLM* package (Section 25.1.3.4).

Locations: http://www.crystal.chem.ruu.nl/distr/strategy.html; ftp://ftp.chem.uu.nl/. Operating system: UNIX. Type: binary. Distribution: free.

25.1.4. Phase determination and structure solution

25.1.4.1. *AMoRe*

AMoRe (Navaza, 1994) is a program package that carries out structure determination using molecular replacement. It reformats the data from the new crystal form, generates structure factors from the model, calculates rotation and translation functions, and applies rigid-body refinement to the solutions. *AMoRe* is part of the *CCP*4 suite (Section 25.1.2.4).

Location: http://www.dl.ac.uk/CCP/CCP4/dist/html/INDEX.html. Operating systems: UNIX, VAM/VMS and LINUX. Type: source code and binary. Distribution: free academic.

25.1.4.2. *GLRF*

GLRF (Tong & Rossmann, 1990) is a program that calculates the general locked rotation function. The self-rotation function determines noncrystallographic symmetry in a crystal. The cross-rotation function determines the orientation relationship of a structure in one unit cell to similar structures in another cell. Since the relationship between the assumed molecular symmetry axes is 'locked', the program can greatly enhance the signal peaks on the rotation function. Therefore, it is much more powerful for assemblies with high local symmetry, such as icosahedral viruses. *GLRF* is part of The Purdue University *XTAL* Program Library (*PUXTAL*: Section 25.1.2.9).

Location: http://www-structure.bio.purdue.edu/~kvz/#GLRF. Operating system: UNIX. Type: source code and binary. Distribution: free.

25.1.4.3. *HEAVY*

The *HEAVY* program package contains the *HEAVY* and *HASSP* programs. The package can carry out heavy-atom search, refinement and MIR/MAD phasing. Some of the major features of *HEAVY* include correlated phasing, Bayesian weighting and Bayesian difference refinement.

Location: http://www.iucr.org/sincris-top/logiciel/prg-heavy.html or e-mail Ncameron@lanl.gov. Operating systems: UNIX and VMS. Type: binary. Distribution: free academic.

25.1.4.4. *MADSYS*

MADSYS (Hendrickson, 1991) is a software package developed over the years in Dr Wayne Hendrickson's laboratory for determining experimental phases of macromolecular structures by multi-wavelength anomalous diffraction (MAD). The package consists of a set of programs that carry out MAD data handling, determination of anomalous-scatterer sites, refinement of MAD sites, MAD phases calculation and structure refinement.

Location: http://convex.hhmi.columbia.edu/hendw/madsys/madsys.html. Operating system: UNIX. Type: binary. Distribution: free academic.

25.1.4.5. *MLPHARE*

MLPHARE is a program for maximum-likelihood heavy-atom refinement and phase calculation. This program refines heavy-atom parameters and error estimates, then uses these refined parameters to generate phase information. The maximum number of heavy atoms that may be refined is 130 over a maximum of 20 derivatives. The program was originally written for MIR, but may also be used for phasing from MAD data, where the different wavelengths are interpreted as different 'derivatives'. *MLPHARE* is part of the *CCP*4 suite (Section 25.1.2.4).

Location: http://www.dl.ac.uk/CCP/CCP4/dist/html/mlphare.html. Operating systems: UNIX, VAX/VMS and LINUX. Type: source code and binary. Distribution: free academic.

25.1.4.6. *Shake-and-Bake*

Shake-and-Bake (*SnB*) (Weeks & Miller, 1999) is a program that uses a dual-space direct-methods phasing algorithm based on the minimal principle to determine crystal structures of macromolecules. The program requires very high resolution data to 1.2 Å or better and $|E|$ values as input. *SnB* has been used to solve structures with more than 600 atoms in the asymmetric unit. Recently, *SnB* has also been used to determine the Se sites in large selenomethionyl-substituted proteins. See Chapter 16.1 for more details.

Location: http://www.hwi.buffalo.edu/SnB/. Operating systems: UNIX, VMS and LINUX. Type: source code. Language: Fortran77. Distribution: free.

25.1.4.7. *SHARP*

SHARP (Statistical Heavy-Atom Refinement and Phasing; de La Fortelle & Bricogne, 1997) operates on reduced, merged and scaled data from SIR(AS), MIR(AS) and MAD experiments, refines the heavy-atom model, helps detect minor or disordered sites using likelihood-based residual maps, and calculates phase probability distributions for all reflections in the data set. See Chapter 16.2 for more details.

Location: http://Lagrange.mrc-lmb.cam.ac.uk/sharp/SharpHome.phtml. Operating systems: IRIX and OSF1. Type: binary. Distribution: free academic.

25.1.5. Structure refinement

Several program packages that are used for structure refinement are described in Section 25.1.2. These include *CNS*, *X-PLOR*, *BIOMOL*, *PHASES* and *PROTEIN*. See Section 25.1.2 for further information.

25.1.5.1. *ARP/wARP*

The Automated Refinement Procedure, *ARP/wARP* (Lamzin & Wilson, 1993, 1997), is a program package for automated model building and refinement of protein structures. It combines, in an iterative manner, reciprocal-space structure-factor refinement with updating of the model in real space to construct and improve protein models. *ARP/wARP* can also be used for *ab initio* structure solution of metalloproteins at high resolution. *ARP/wARP* is distributed as part of the *CCP*4 suite (Section 25.1.2.4). See Section 25.2.5 for a detailed description.

Location: http://www.embl-hamburg.de/ARP/. Operating systems: UNIX, HPUX, IRIX and LINUX. Type: source code and binary. Language: Fortran77. Distribution: free academic.

25.1.5.2. *MULTAN88*

*MULTAN*88 (Main *et al.*, 1980) is a program that uses direct methods to determine crystal structures from single-crystal diffraction data. It can be used for very high resolution structure refinement and determination of heavy-atom positions.

Location: http://www.msc.com/. Operating systems: UNIX and VAX/VMS. Type: binary. Distribution: commercial.

25.1.5.3. *PROLSQ*

PROLSQ (Hendrickson & Konnert, 1979) is used for the restrained least-squares refinement of a protein structure. Prior to running *PROLSQ*, the program *PROTIN* must be run to analyse the protein geometry and produce an output file containing restraints information. *PROLSQ* cannot calculate structure factors. Use *SFALL* to calculate X-ray contributions to the matrix. *PROLSQ* is distributed as a unsupported program of the *CCP*4 suite (Section 25.1.2.4).

Location: http://www.dl.ac.uk/CCP/CCP4/dist/. Operating systems: UNIX, VAX/VMS and LINUX. Type: source code and binary. Distribution: free academic.

25.1.5.4. *REFMAC*

REFMAC (Murshudov *et al.*, 1997, 1999) is a macromolecular refinement program which has been integrated into the *CCP*4 suite (Section 25.1.2.4). *REFMAC* can carry out rigid-body, restrained or unrestrained refinement against X-ray data, or idealization of a macromolecular structure. It minimizes the coordinate parameters to satisfy either a maximum-likelihood or least-squares residual. There are options to use different minimization methods. If the user wishes to invoke geometric restraints, the program *PROTIN*, which analyses the protein geometry and produces an output file containing restraints information, must be run prior to running *REFMAC*. *REFMAC* also produces an MTZ output file containing weighted coefficients for *SIGMAA*-weighted mFo-DFcalc and 2mFo-DFcalc maps, where 'missing data' have been restored.

Location: http://www.dl.ac.uk/CCP/CCP4/dist/html/refmac.html. Operating systems: UNIX, SGI, SUN, DEC and LINUX. Type: source code and binary. Distribution: free.

25.1.5.5. *RSRef*

RSRef (Chapman, 1995) is a package of programs that enables an atomic model to be optimized by fitting to an electron-density map. *RSRef* uses an electron-density function that is resolution dependent, so that it accurately models a medium-resolution map. When combined with *TNT*'s (Section 25.1.5.8) *Geometry*, full stereochemical refinement is possible. *RSRef* can be used to quickly prerefine a protein structure during or after model building, or to completely refine structures with high noncrystallographic symmetry that have good electron density.

Location: http://www.sb.fsu.edu/∼rsref/. Operating systems: SGI and EVS. Type: source code and binary. Distribution: minor licence fee for academic users.

25.1.5.6. *SHELX97*

SHELX (Sheldrick & Schneider, 1997) is a set of programs for crystal structure determination from single-crystal diffraction data. Originally *SHELX* was intended only for small molecules. However, improvements in computing performance and data-collection methods have led to increased use of *SHELX* for macromolecules, especially the location of heavy atoms from isomorphous and anomalous-difference data, and the refinement of proteins against high-resolution data (2.5 Å or better). See Section 25.2.10 for a detailed description.

Location: http://shelx.uni-ac.gwdg.de/SHELX/. Operating systems: UNIX, VMS, DOS and Windows. Type: binary. Language: Fortran77. Distribution: free academic.

25.1.5.7. *SIR97*

*SIR*97 (Altomare *et al.*, 1999) is an integrated program package for the determination and refinement of small-molecule structures from single-crystal diffraction data. It is also useful in solving the heavy-atom positions in protein structure determination.

Location: http://www.ba.cnr.it/IRMEC/Sir_Waremain.html. Operating systems: UNIX, VMS, MacOS and Windows. Type: binary. Distribution: free academic.

25.1.5.8. *TNT*

TNT (Tronrud *et al.*, 1987; Tronrud, 1997) is a general-purpose program package for the structure refinement of macromolecules using single-crystal X-ray diffraction data. It is normally used to optimize a model to X-ray diffraction data while maintaining proper stereochemistry using least-squares function-minimization techniques. It can restrain a model to bond lengths, bond angles, dihedral angles, pseudo-rotation angles, planarity and non-bonded 'close' contacts (including symmetry-related contacts). A principal advantage of the *TNT* package is its great flexibility, making it ideal for restraining structures that contain cofactors, inhibitors, or nucleic acids. The package is composed of separate programs, each performing clearly defined tasks. To use the package with other forms of data you simply write programs that produce the value and first derivative of the functional term you wish to minimize. See Section 25.2.4 for a detailed description.

Location: http://www.uoxray.uoregon.edu/tnt/welcome.html. Operating systems: UNIX, VAX/VMS, DEC Alpha, EVS, AIX, SUN and SGI. Type: source code and binary. Distribution: free academic.

25.1.6. Phase improvement and density-map modification

25.1.6.1. *BUSTER*

BUSTER (Bricogne, 1997*a*,*b*) is a program for recovering missing phase information by Bayesian inference. *BUSTER* has applications