in maximum-likelihood refinement of partial structures in conjunction with the *TNT* program (Section 25.1.5.8), maximum-entropy structure completion for missing or ambiguous parts of a structure, and accurate electron-density reconstruction based on high-resolution X-ray diffraction data. *BUSTER* is related to *SHARP* (Section 25.1.4.7). See Chapter 16.2 for more details.

Location: http://lagrange.mrc-lmb.cam.ac.uk/buster/Buster Home.phtml. Operating systems: IRIX and OSF1. Type: binary. Distribution: free academic.

25.1.6.2. *DM/DMMULTI*

DM (Cowtan, 1994) is a density-modification program package. *DM* applies real-space constraints based on known features of a protein electron-density map in order to improve the approximate phasing obtained from experimental sources. Various information can be applied, including the following diverse elements: solvent flattening, histogram mapping, multi-resolution modification, NCS averaging, skeletonization and Sayre's equation. *DM* is part of the *CCP*4 suite (Section 25.1.2.4). See Chapter 15.1 and Section 25.2.2 for more details.

Operating systems: UNIX, VAX/VMS and LINUX. Type: source code and binary. Distribution: free academic.

25.1.6.3. *FINDNCS*

FINDNCS (Lu, 1999) is a program that automatically determines NCS operations from heavy-atom sites to aid in applying averaging techniques in the MIR procedure. The program outputs the NCS operations (a rotation matrix and translation vector), r.m.s. deviations, polar angles and screw distance, matching sites and other useful information for users. The program can also generate files that can be used to display NCS operations using the program *O* (Section 25.1.7.7).

Location: http://gamma.mbb.ki.se/∼guoguang/findncs.html. Operating systems: UNIX, IRIX and OSF1. Type: binary. Language: Fortran77. Distribution: free academic.

25.1.6.4. *RAVE*

RAVE (Jones, 1992; Kleywegt & Jones, 1994) is a suite of programs for real-space electron-density averaging of crystallographic electron density between single and multiple domains, and between single and multiple crystal forms. It also contains tools for the detection of secondary-structure elements in macromolecular electron-density maps. See Chapter 17.1 for a detailed description.

Location: http://xray.bmc.uu.se/usf/menu.html#sof; ftp://xray.bmc.uu.se/. Operating systems: UNIX, SGI and DEC Alpha/OSF1. Type: binary. Distribution: free.

25.1.6.5. *SOLOMON*

SOLOMON (Abrahams & Leslie, 1996) is a program that modifies electron-density maps by averaging, solvent flattening and protein truncation. It can also remove overlapped parts of a mask between itself and its symmetry equivalents. *SOLOMON* is part of the *CCP*4 suite (Section 25.1.2.4).

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

25.1.6.6. *SQUASH*

The *SQUASH* program (Zhang & Main, 1990*a*,*b*) provides a tool for phase refinement and extension of macromolecular structures. The starting point is a set of native structure factors to some resolution, and estimated phases and figures of merit for some subset of the phases. The result is a set of improved phases and figures of merit for the whole data set. The program combines Sayre's equation with density modification by histogram matching, solvent flattening and noncrystallographic symmetry averaging. The real-space formulation enables any electron-density constraint to be applied easily, *e.g.* solvent flattening with (eventually) known regions of density. The least-squares solution of a large system of nonlinear equations is achieved by Newton–Raphson iteration that converts the system of nonlinear equations into linear ones. The system of linear equations is solved by the conjugate-gradient method using FFTs.

Location: http://www.msc.com/brochures/software/squash.html. Operating system: UNIX. Type: binary. Distribution: commercial.

25.1.7. Graphics and model building

25.1.7.1. *AMBER*

AMBER (Assisted Model Building with Energy Refinement; Cornell *et al.*, 1995) is a molecular-dynamics and energy-minimization program. *AMBER* refers to two things: a molecular-mechanical force field for the simulation of biomolecules (which is in general use in a variety of simulation programs) and a package of molecular-simulation programs which includes source code and demonstrations.

Location: http://www.amber.ucsf.edu/amber/amber.html. Operating systems: UNIX, SGI, SUN *etc*. Type: source code and binary. Languages: Fortran and C. Distribution: commercial.

25.1.7.2. *CHARMM*

CHARMM (Chemistry at HARvard Molecular Mechanics; Brooks *et al.*, 1983; MacKerell *et al.*, 1998) is a program for macromolecular simulations, including energy minimization, molecular dynamics and Monte Carlo simulations.

Location: http://yuri.harvard.edu/. Operating systems: UNIX, SGI, SUN *etc*. Type: source code. Language: C. Distribution: minor licence fee for academic users.

25.1.7.3. *Insight II*

Insight II is a 3D graphical environment for molecular modelling. *Insight* II creates, modifies, manipulates, displays and analyses molecular systems and related data, and provides the core requirements for all *Insight* II software modules. Its powerful user interface enables the seamless flow of data between a wide range of scientific applications. The *Insight* II environment integrates builder modules, development tools, force fields, simulation and visualization tools with tools specifically developed for applications in the life and materials sciences.

Location: http://www.msi.com/life/products/insight/index.html. Operating systems: SGI and IBM UNIX systems. Type: binary. Distribution: commercial.

25.1.7.4. *MidasPlus*

MidasPlus (formerly *Midas*) (Ferrin *et al.*, 1988) is an advanced molecular-modelling system developed by the Computer Graphics Laboratory (CGL) at the University of California, San Francisco. The system can be used for display and manipulation of macromolecules such as proteins and nucleic acids. Ancillary programs allow for such features as computation of molecular surfaces and electrostatic potentials and generation of publication-quality space-filling images with multiple light sources and shadows. To address the needs of the structure-based drug-design community, *MidasPlus* has been developed with an emphasis on the interactive

selection, manipulation and docking of drugs and receptors. Although quite powerful in this application, the system is also somewhat specialized in this respect: it requires three-dimensional atomic coordinate data for the structures being displayed and expects the primary structure to be based on linear chains of subunits such as amino acids or nucleic acids. Using *MidasPlus* for complex inorganic compounds or large polymers with many cross-links is another application.

Location: http://www.cgl.ucsf.edu/Outreach/midasplus/. Operating systems: SGI, DEC Alpha, NeXT, IBM RS6000 and LINUX. Type: source code. Distribution: minor licence fee for academic users.

25.1.7.5. *MODELLER*

MODELLER (Sali & Blundell, 1993) is most frequently used for homology or comparative modelling of protein three-dimensional structure. The user provides an alignment of a sequence to be modelled with known related structures and *MODELLER* will automatically calculate a full-atom model. More generally, *MODELLER* models protein 3D structure by satisfaction of spatial restraints. In principle, the restraints can be derived from a number of different sources. These include homologous structures (comparative modelling), NMR experiments (NMR refinement), rules of secondary-structure packing (combinatorial modelling), cross-linking experiments, fluorescence spectroscopy, image reconstruction in electron microscopy, site-directed mutagenesis, intuition, residue–residue and atom–atom potentials of mean force, *etc*. The output of *MODELLER* is a 3D structure of a protein that satisfies these restraints as well as possible. The optimization is carried out by the variable-target function procedure employing methods of conjugate gradients and molecular dynamics with simulated annealing. *MODELLER* can also do several other tasks, including multiple comparison of protein sequences and/or structures, clustering, and searching of sequence databases. *MODELLER* is also available as part of *QUANTA* (Section 25.1.7.8), *Insight* II (Section 25.1.7.3) and *Weblab GeneExplorer*.

Location: http://guitar.rockefeller.edu/modeller/modeller.html; ftp://guitar.rockefeller.edu/pub/modeller/. Operating system: UNIX. Type: source code and binary. Language: Fortran. Distribution: free academic.

25.1.7.6. *MOLMOL*

MOLMOL is a molecular graphics program for display, analysis and manipulation of three-dimensional structures of biological macromolecules, with special emphasis on nuclear magnetic resonance (NMR) solution structures of proteins and nucleic acids. *MOLMOL* has a graphical user interface with menus, dialogue boxes and online help. The display possibilities include conventional presentations, as well as novel schematic drawings, with the option of displaying different presentations in one view. Covalent molecular structures can be modified by addition or removal of individual atoms and bonds. The three-dimensional structure can be manipulated by interactive rotation about individual dihedral angles. Special efforts were made to allow for appropriate display and analysis of sets of (typically 20–40) conformers that are conventionally used to represent the result of an NMR structure determination, using functions for superimposing sets of conformers, calculation of root-mean-square-distance (r.m.s.d.) values, identification of hydrogen bonds, checking and displaying violations of NMR constraints, and identification and listing of short distances between pairs of hydrogen atoms.

Location: http://www.mol.biol.ethz.ch/wuthrich/software/ molmol/. Operating systems: UNIX, IRIX, AIX, OSF1 and LINUX. Type: source code and binary. Distribution: free.

25.1.7.7. *O*

O (Jones *et al.*, 1991) is a general-purpose macromolecular-modelling package. The program is aimed at scientists with a need to model, build and display macromolecules. Unlike other molecular-modelling programs, such as *FRODO* (Section 25.1.7.10), *O* is a graphical display program built on top of a versatile database system. All molecular data are kept in this database, in a predefined data structure. However, any data can be stored in the database. Data produced by associated stand-alone programs can be stored very easily in the database and used by the program, for example for colouring of atoms. The powerful macro facility of *O* enables the user to customize the use of the program to satisfy his or her specific needs. The current version of *O* is mainly aimed at the field of protein crystallography, bringing into use several new tools which ease the building of models into electron density, allowing it to be done faster and more correctly. Notably, some new auto-build options greatly enhance the speed of building and rebuilding molecular models. See Chapter 17.1 for a detailed description.

Locations: http://kaktus/imsb.au.dk/∼mok/o/; ftp://xray.bmc.uu.se/. Operating systems: UNIX and ESV. Type: binary. Distribution: free academic.

25.1.7.8. *QUANTA*

QUANTA is an extensive library of crystallographic software programs that streamline and accelerate protein structure solution. *QUANTA* provides a powerful and comprehensive modelling environment for 2D and 3D modelling, simulation and analysis of macromolecules and small organic compounds.

Location: http://www.msi.com/life/products/quanta/index.html. Operating system: SGI. Type: binary. Distribution: commercial.

25.1.7.9. *SYBYL*

SYBYL is a comprehensive computational tool kit for molecular design and analysis, with a special focus on the creation of new chemical entities. *SYBYL* provides essential construction and analysis tools for both organic and inorganic molecular structures. It is especially useful in building the structures of ligands, substrates and inhibitors.

Location: http://www.tripos.com/software/sybyl.html. Operating system: UNIX. Type: binary. Distribution: commercial.

25.1.7.10. *Turbo FRODO*

FRODO (Jones, 1978) is a general-purpose molecular-modelling program which can be used to model *de novo* macromolecules, polypeptides and nucleic acids from experimental 3D data obtained from X-ray crystallography and NMR, and to display the resulting models using various representations including van der Waals and Connolly molecular dot surfaces, as well as spline surfaces. *Turbo FRODO* is designed for ligand fitting and protein stacking. The user can interactively mutate a protein or chemically modify it, and evaluate the resulting conformational changes. There are several versions of *FRODO* around the scientific community. For LINUX and HPUX use the *Turbo FRODO* X version.

Location: http://afmb.cnrs-mrs.fr/TURBO_FRODO/turbo.html. Operating systems: HPUX, IRIX and LINUX. Type: binary. Distribution: commercial.

25.1.8. Structure analysis and verification

25.1.8.1. *DSSP*

The *DSSP* program was designed by Kabsch & Sander (1983) to standardize secondary-structure assignment. The DSSP database is a