

## 8. Strategies for Macromolecular Refinement

SHELXL is designed to be easy to use and general for all space groups and uses a conventional structure-factor calculation rather than a FFT summation; the latter would be faster, but in practice involves some small approximations and is not very suitable for the treatment of dispersion or anisotropic thermal motion. The price to pay for the extra generality and precision is that SHELXL is much slower than programs written specifically for macromolecules, but this is to some extent compensated for by the better convergence properties, reducing the amount of manual intervention required (and also the *R*-factor).

Recent advances in cryogenic techniques, area detectors, and the use of synchrotron radiation enable macromolecular data to be collected to higher resolution than was previously possible. In practice this tends to complicate the refinement because it is possible to resolve finer details of the structure; it is often necessary to model alternative conformations, and in a few cases even anisotropic refinement is justified. Although SHELXL provides a number of other features not found in many macromolecular refinement programs, it is probably the flexible treatment of disorder and the facilities for restrained anisotropic refinement that are most likely to be of immediate interest to macromolecular crystallographers.

An auxiliary program SHELXPRO (Chapter 9) is provided as an interface to other macromolecular programs. SHELXPRO is able to generate an *.ins* file from a PDB format file, including the appropriate restraints etc. SHELXPRO can also generate map files for the program O and can display the refinement results in the form of Postscript plots, as well as including the updated coordinates in the *.ins* file for the next refinement. SHELXL produces PDB and CIF format files that can be read by SHELXPRO and used for archiving.

### 8.1 The radius of convergence

A crucial aspect of any macromolecular refinement program is the radius of convergence. A larger radius of convergence reduces the amount of time-consuming manual intervention using interactive graphics. Many claims that SHELXL gives *R*-factors one or two percent lower than other programs have been tracked down either to subtle differences in the model or to not getting trapped in local minima. The differences in the model include the treatment of diffuse solvent and hydrogen atoms, and the ability to refine common occupancies for disordered groups. The inclusion of dispersion terms and the use of a conventional rather than a FFT structure factor summation are also more precise; the approximations in the FFT summation may become significant for high resolution data and atoms with small displacement parameters. There are probably a number of contributing factors to the good convergence typically observed for SHELXL, e.g. the refinement against ALL data, the inclusion of important off-diagonal terms in the least-squares algebra, the ability to refine all parameters at once (i.e. coordinates and displacement parameters in the same cycle), and the restriction to unimodal restraint functions; multimodal restraint functions such as torsion angles or hydrogen bonds tend to increase the number of spurious local minima. It is much better to reserve the multimodal chemical information such as torsion angles for verifying the structure with an independent program such as PROCHECK (Laskowski, MacArthur, Moss & Thornton, 1993), and to use the unimodal information as restraints. The errors in the FFT

calculation of derivatives are larger than those in the structure factors (for the same grid intervals); this would also impede convergence.

## 8.2 Residues

Macromolecular structures are conventionally divided up into *residues*, for example individual amino-acids. In SHELXL residues may be referenced either individually, by '\_' followed by the appropriate residue number, or as all residues of a particular class, by '\_' followed by the class. For example 'DFIX 2.031 SG\_9 SG\_31' could be used to restrain a disulfide distance between two cystine residues, whereas 'FLAT\_PHE CB > CZ' would apply planarity restraints to all atoms between CB and CZ inclusive in all PHE (phenylalanine) residues. Plus and minus signs refer to the next and previous residue numbers respectively, so 'DFIX\_\* 1.329 C\_- N' applies a bond length restraint to all peptide bonds ('\_\*' after the command name applies it to all residues). This way of referring to atoms and residues is in no way restricted to proteins; it is equally suitable for oligonucleotides, polysaccharides, or to structures containing a mixture of all three. It enables the necessary restraints and other instructions to be input in a concise and relatively self-explanatory manner. These instructions are checked by the program for consistency and appropriate warnings are printed.

## 8.3 Constraints and restraints for macromolecules

The lower data to parameter ratio for macromolecules makes the use of constraints and especially restraints essential. Rigid group constraints enable a structure to be refined with very few parameters, especially when the (thermal) displacement parameters are held fixed (BLOC 1). After a structure has been solved by molecular replacement using a rather approximate model for the whole protein or oligonucleotide, it may well be advisable to divide the structure up into relatively rigid domains (using a few AFIX 6 and AFIX 0 instructions) and to refine these as rigid groups, initially for a limited resolution shell (e.g. SHEL 8 3), then stepwise extending the resolution, e.g. using the STIR instruction. Restraints may still be required to define flexible hinges and prevent the units from flying apart. In view of the small number of parameters and the high correlations introduced by rigid group refinement, L.S. (full-matrix refinement) should be used for this stage (but CGLS will be necessary for the subsequent refinement). After this initial step which exploits the large convergence radius of rigid group refinement, in general the more flexible restraints will be used in preference to constraints for the rest of the refinement.

SHELXL provides distance, planarity and chiral volume restraints, but not torsion angle restraints or specific hydrogen bond restraints. For oligonucleotides, good distance restraints are available for the bases (Taylor & Kennard, 1982), but for the sugars and phosphates it is probably better to assume that chemically equivalent 1,2- and 1,3-distances are equal (using the SAME and SADI restraints) without the need to specify target values. In this way the effect of the pH on the protonation state of the phosphates and hence the P-O distances does not need to be predicted, but it is assumed the whole crystal is at the same pH. For proteins, since some amino-acid residues occur only a small number of times in a given protein, it is probably better to use 1,2- and 1,3-target distances based on the study of Engh and Huber (1991); these are employed in the restraints added by SHELXPRO to the *.ins* file.

The three bonds to a carbonyl carbon atom can be restrained to lie in the same plane by means of a *chiral volume restraint* (Hendrickson & Konnert, 1980) with a target volume of zero (e.g. 'CHIV\_GLU 0 C CD' to restrain the carbonyl and carboxyl carbons in all glutamate residues to have planar environments). The planarity restraint (FLAT) restrains the chiral volumes of a sufficient number of atomic tetrahedra to zero; in addition the r.m.s. deviation of the atoms from the best planes is calculated. Chiral volume restraints with non-zero targets are useful to prevent the inversion of  $\alpha$ -carbon atoms and the  $\beta$ -carbons of Ile and Thr, e.g. 'CHIV\_ILE 2.5 CA CB'. It is also useful to apply chiral volume restraints to non-chiral atoms such as CB of valine and CG of leucine in order to ensure conformity with conventional atom-labeling schemes (from the point of view of the atom names, these atoms could be considered to be chiral !).

*Anti-bumping restraints* are distance restraints that are only applied if the two atoms are closer to each other than the target distance. They can be generated automatically by SHELXL, taking all symmetry equivalent atoms into account. Since this step is relatively time-consuming, in the 1993 release it was performed only before the first refinement cycle, and the anti-bumping restraints were generated automatically only for the solvent (water) atoms (however they could be inserted by hand for any pairs of atoms). In practice this proved to be too limited, so in later releases the automatic generation of anti-bumping restraints was extended to all C, N, O and S atoms (with an option to include H...H interactions) and was performed each refinement cycle. Anti-bumping restraints are not generated automatically for (a) atoms connected by a chain of three bonds or less in the connectivity array (unless separated by more than a specified number of residues), (b) atoms with different non-zero PART numbers, and (c) pairs of atoms for which the sum of occupancies is less than 1.1. The target distances for the O...O and N...O distances are less than for the other atom pairs to allow for possible hydrogen bonds.

#### 8.4 Restrained anisotropic refinement

There is no doubt that macromolecules are better described in terms of anisotropic displacements, but the data to parameter ratio is very rarely adequate for a free anisotropic refinement. Such a refinement often results in 'non-positive definite' (NPD) displacement tensors, and at the best will give probability ellipsoids that do not conform to the expected dynamical behavior of the molecule. Clearly constraints or restraints must be applied to obtain a chemically sensible model. It is possible to divide a macromolecule up into relatively rigid domains, and to refine the 20 TLS parameters of rigid body motion for each domain (Driessen, Haneef, Harris, Howlin, Khan & Moss, 1989). This is a good model for the bases in oligonucleotides and for the four aromatic side-chains in proteins, but otherwise macromolecules are probably not sufficiently rigid for the application of TLS constraints, or they would have to be divided up into such small units that too many parameters would be required. As with the refinement of atomic positions, restraints offer a more flexible approach.

The *rigid bond restraint* (DELU) assumes that the components of the anisotropic displacement parameters (ADPs) along bonded (1,2-) or 1,3-directions are zero within a given esd. This restraint should be applied with a low esd, i.e. as a 'hard' restraint. Didisheim & Schwarzenbach (1987) showed that for many non-planar groups of atoms, rigid bond restraints effectively impose TLS conditions of rigid body motion. Although rigid-bond restraints involving 1,2- and 1,3-distances reduce the effective number of free ADPs per atom

from 6 to less than 4 for typical organic structures, further restraints are often required for the successful anisotropic refinement of macromolecules.

The *similar ADP restraint* (SIMU) restrains the corresponding  $U_{ij}$ -components to be approximately equal for atoms which are spatially close (but not necessarily bonded because they may be in different components of a disordered group). The isotropic version of this restraint has been employed frequently in protein refinements. This restraint is consistent with the characteristic patterns of thermal ellipsoids in many organic molecules; on moving out along side-chains, the ellipsoids become more extended and also change direction gradually.

Neither of these restraints are suitable for isolated solvent (water) molecules. A linear restraint (ISOR) restrains the ADP's to be *approximately isotropic*, but without specifying the magnitude of the corresponding equivalent isotropic displacement parameter. Both SIMU and ISOR restraints are clearly only approximations to the truth, and so should be applied as 'soft' restraints with high esds. When all three restraints are applied, structures may be refined anisotropically with a much smaller data to parameter ratio, and still produce chemically sensible ADP's. Even when more data are available, these restraints are invaluable for handling disordered regions of the structure.

Constraints and restraints greatly increase the radius and rate of convergence of crystallographic refinements, so they should be employed in the early stages of refinement wherever feasible. The difference electron density syntheses calculated after such restrained refinements are often more revealing than those from free refinements. In large small-molecule structures with poor data to parameter ratios, the last few atoms can often not be located in a difference map until an anisotropic refinement has been performed with geometrical and ADP restraints. Atoms with low displacement parameters that are well determined by the X-ray data will be relatively little affected by the restraints, but the latter may well be essential for the successful refinement of poorly defined regions of the structure. Premature removal or softening the restraints (to improve the  $R$ -value !) is probably the most common cause of unstable macromolecular refinements with SHELXL.

## 8.5 The free $R$ -factor

The questions of whether the restraints can be removed in the final refinement, or what the best values are for the corresponding esds, can be resolved elegantly by the use of  $R_{\text{free}}$  (Brünger, 1992). To apply this test, the data are divided into a working set (about 95-90% of the reflections) and a reference set (about 5-10%). The reference set is only used for the purpose of calculating a conventional  $R$ -factor that is called  $R_{\text{free}}$ . It is very important that the structural model is not in any way based on the reference set of reflections, so these are left out of ALL refinement and Fourier map calculations. If the original model was in any way derived from the same data, then many refinement cycles are required to eliminate memory effects. This ensures that the  $R$ -factor for the reference set provides an objective guide as to whether the introduction of additional parameters or the weakening of restraints has actually improved the model, and not just reduced the  $R$ -factor for the data employed in the refinement (' $R$ -factor cosmetics'). The best way to set up the  $R_{\text{free}}$  test is to use SHELXPRO to flag reflections in the *.hkl* file for use in the reference set, and to set the second CGLS parameter to '-1'. If NCS or twinning is anticipated, it is advisable to use the 'thin shells' method of flagging the reflections for  $R_{\text{free}}$  in SHELXPRO.

$R_{\text{free}}$  is invaluable in deciding whether a restrained anisotropic refinement is significantly better than an isotropic refinement. Experience indicates that both the resolution and the quality of the data are important factors, but that restrained anisotropic refinement is unlikely to be justified for crystals that do not diffract to better than 1.5 Å. An ensemble distribution created by molecular dynamics is an alternative to the harmonic description of anisotropic motion (Gros, van Gunsteren & Hol, 1990; Clarage & Phillips, 1994), and may be more appropriate for structures with severe conformational disorder that do not diffract to high resolution.

Despite the overwhelming arguments for using  $R_{\text{free}}$  to monitor macromolecular refinements, it is only a single number, and is itself subject to statistical uncertainty because it is based on a limited number of reflections. Thus  $R_{\text{free}}$  may be insensitive to small structural changes, and small differences in  $R_{\text{free}}$  should not be taken as the last word; one should always consider whether the resulting geometrical and displacement parameters are *chemically reasonable*. The final refinement and maps should always be calculated with the **full** data, but without introducing additional parameters or changing the weights of the restraints.  $R_{\text{free}}$  is most useful for establishing refinement protocols; for a series of closely similar refinements (e.g. for mutants to similar resolution) the  $R_{\text{free}}$  tests only need to be applied to the first.

## 8.6 Disorder in macromolecules

To obtain a chemically sensible refinement of a disordered group, we will probably need to constrain or restrain a sum of occupation factors to be unity, to restrain equivalent interatomic distances to be equal to each other or to standard values (or alternatively apply rigid group constraints), and to restrain the displacement parameters of overlapping atoms. In the case of a tight unimodal distribution of conformations, restrained anisotropic refinement may provide as good a description as a detailed manual interpretation of the disorder in terms of two or more components, and is much simpler to perform. With high-resolution data it is advisable to make the atoms anisotropic BEFORE attempting to interpret borderline cases of side-chain disorder; it may well be found that no further interpretation is needed, and in any case the improved phases from the anisotropic refinement will enable higher quality difference maps to be examined.

Typical warning signs for disorder are large (and pronounced anisotropic) apparent thermal motion (in such cases the program may suggest that an atom should be split and estimate the coordinates for the two new atoms), residual features in the difference electron density and violations of the restraints on the geometrical and displacement parameters. This information is summarized by the program on a residue by residue basis, separately for main-chain, side-chain and solvent atoms. In the case of two or more discrete conformations, it is usually necessary to model the disorder at least one atom further back than the maps indicate, in order that the restraints on the interatomic distances are fulfilled. The different conformations should be assigned different PART numbers so that the connectivity array is set up correctly by the program; this enables the correct rigid bond restraints on the anisotropic displacement parameters and idealized hydrogen atoms to be generated automatically even for disordered regions (it is advisable to model the disorder before adding the hydrogens).

Several strategies are possible for modeling disorder with SHELXL, but for macromolecules the simplest is to include all components of the disorder in the same residues and use the

same atom names, the atoms belonging to different components being distinguished only by their different PART numbers. This procedure enables the standard restraints etc. to be used unchanged, because the same atom and residue names are used. No special action is needed to add the disordered hydrogen atoms, provided that the disorder is traced back one atom further than it is visible (so that the hydrogen atoms on the PART 0 atoms bonded to the disordered components are also correct). Note that this very simple and effective treatment of disorder was not available in the original 1993 release of SHELXL.

## 8.7 Automatic water divining

It is relatively common practice in the refinement of macromolecular structures to insert water molecules with partial occupancies at the positions of difference electron density map peaks in order to reduce the  $R$ -factor (another example of '  $R$ -factor cosmetics'). Usually when two different determinations of the same protein structure are compared, only the most tightly bound waters, which usually have full occupancies and smaller displacement parameters, are the same in each structure. The refinement of partial occupancy factors for the solvent atoms (in addition to their displacement parameters) is rarely justified by  $R_{\text{free}}$ , but sometimes the best  $R_{\text{free}}$  value is obtained for a model involving some water occupancies fixed at 1.0 and some at 0.5.

Regions of diffuse solvent may be modeled using *Babinet's principle* (Moews & Kretsinger, 1975); the same formula is employed in the program TNT, but the implementation is somewhat different. In SHELXL it is implemented as the SWAT instruction and usually produces a significant but not dramatic improvement in the agreement of the very low angle data. Anti-bumping restraints may be input by hand or generated automatically by the program, taking symmetry equivalents into account. After each refinement job, the displacement parameters of the water molecules should be examined, and waters with very high values (say  $U$  greater than  $0.8 \text{ \AA}^2$ , corresponding to a  $B$  of 63) eliminated. The  $F_o - F_c$  map is then analyzed automatically to find the highest peaks which involve no bad contacts and make at least one geometrically plausible hydrogen bond to an electronegative atom. These peaks are then included with full occupancies and oxygen scattering factors in the next refinement job. This procedure is repeated several times; in general  $R_{\text{free}}$  rapidly reaches its minimum value, although the conventional  $R$ -index continues to fall as further waters are added. It should be noted that the automatic generation of anti-bumping restraints is less effective when the water occupancies are allowed to have values other than 1.0 or 0.5. This approach provides an efficient way of building up a chemically reasonable (but not necessarily unique) network of waters that are prevented from diffusing into the protein, thus facilitating remodeling of disordered side-chains etc. The occupancies of specific waters may also be tied (using free variables) to the occupancies of particular components of disordered side-chains where this makes chemical sense. This procedure may be facilitated by using SHELXPRO to convert the *.res* output file from one refinement job to the *.ins* file for the next, or fully automated using the program SHELXWAT that calls SHELXL repeatedly. A similar but much more sophisticated approach (ARP) described by Lamzin & Wilson (1993) may also be used in conjunction with SHELXL.

## 8.8 Refinement of structures at modest resolution

Although the unique features of SHELXL are primarily useful for refinement against very high resolution data, tests indicated that only small changes would be required to the original release to extend its applicability to medium resolution data (say 2.5Å). The most important of these changes were improved diagnostics and more sophisticated anti-bumping restraints (see above), and the addition of non-crystallographic symmetry (NCS) restraints. The use of NCS restraints considerably improves the effective data to parameter ratio, and the resulting Fourier maps often look as though they were calculated with higher resolution data than were actually used (because the phases are more accurate). Two types of NCS restraint may be generated automatically with the help of the NCSY instruction. The first type uses the connectivity table to define equivalent 1,4-distances, which are then restrained to be equal. The second restrains the isotropic  $U$ -values of equivalent atoms to be equal. It is not normally necessary to restrain equivalent 1,2- and 1,3-distances to be equal because the DFIX and DANG restraints will have this effect anyway; but SAME may be used to add such restraints in the absence of DFIX and DANG. The use of *restraints* rather than applying NCS as an exact constraint (e.g. in the structure factor calculation) is more flexible (but slower) and does not require the specification of transformation matrices and real-space masks. Experience indicates that NCS restraints should be used wherever possible; it is not difficult to relax them later (e.g. for specific side-chains involved in interactions with other non-NCS related molecules) should this prove to be necessary.

## 8.9 A typical SHELXL refinement using high resolution data

An example of a typical SHELXL refinement against high resolution data, for an inhibited form of serine protease, is summarized in Table 5.2. Data were collected at 120 K on a synchrotron to 0.96 Å resolution with an overall mean  $I/\sigma$  of 15.2 and a  $R_{\text{merge}}$  (based on intensities) of 3.7%. Molecular dynamics refinement using X-PLOR (Brünger, Kuriyan & Karplus, 1987) from initial  $R$ -values of 42.5% produced the results shown as Job 1. A reference set consisting of 10% of the reflections and a working set of the remaining 90% were used throughout the X-PLOR and SHELXL refinement. The final X-PLOR and initial SHELXL refinements were performed with the resolution range restricted to 1.1 to 8 Å (48495 working set reflections) to save computer time. 10 conjugate gradient cycles were performed in each of the SHELXL refinement jobs; where new atoms were introduced they were always refined isotropically for 2 cycles before making them anisotropic. The CPU times (150 MHz Silicon Graphics R4400 processor) varied from 6.1 hours for job 2 to 21.7 for job 13. The weighting scheme was fixed at 'WGHT 0.2' until jobs 12 and 13, where the two-parameter scheme with values suggested by the program was employed. The restraints (DEFS 0.015 0.2 0.01 0.025) were made tighter than usual to make the refinements more comparable with X-PLOR; the mean distance deviation was 0.009 Å for X-PLOR and 0.014 Å for the final SHELXL job.

Introduction of the diffuse solvent parameter in job 3 (which started from the same parameters as job 2) was not significant, although usually it reduces  $R_{\text{free}}$  by about 0.5%; probably this was a consequence of leaving out the low angle data at this stage. Making all atoms anisotropic resulted in almost a 3% drop in  $R_{\text{free}}$ , but from experience of similar structures we believe that the drop would have been larger if all the data had been used at this stage. This helps to explain the further drop in  $R_{\text{free}}$  on using all the reflection data (job 8), and the fact that the

difference between  $R_1$  and  $R_{\text{free}}$  was about 3% for jobs 4 to 7 and about 2% for the remaining jobs. Particularly noteworthy is the drop in the  $R$ -factors on introducing hydrogens (no extra parameters); a parallel job using exactly the same model but excluding hydrogens showed that 1.25% of the drop in  $R_{\text{free}}$  was contributed by the hydrogens. On the other hand the drop in job 12 is caused almost entirely by the improvements to the model; the same job with the original weights gave an  $R_{\text{free}}$  of 10.90%. After using  $R_{\text{free}}$  to monitor the refinement as discussed here, a final refinement was performed against all 80102 unique reflections without any further changes to the model; this converged to  $R_1 = 8.77\%$ , essentially identical to the final  $R_1$  for the working set.

### SHELXL refinement of a serine protease (188 residues)

Job	Action taken	NP	NH	NW/NW <sub>1/2</sub> /NX	N <sub>par</sub>	R <sub>1</sub>	R <sub>free</sub>
1	Final X-PLOR, 1.1-8Å	1337	0	176 / 0 / 19	6129	19.47	21.14
2	Same atoms, SHELXL	1337	0	176 / 0 / 19	6129	17.15	18.96
3	SWAT added	1337	0	176 / 0 / 19	6130	17.07	18.95
4	All atoms anisotropic	1337	0	176 / 0 / 19	13790	12.96	16.10
5	Disorder, added solvent	1376	0	207 / 0 / 34	14565	11.46	14.20
6	More disorder and solvent	1422	0	214 / 2 / 39	14831	11.35	14.22
7	Disorder, half occ. waters	1447	0	213 / 20 / 37	15478	11.13	14.10
8	Resolution: 0.96Å-Inf.	1447	0	218 / 28 / 37	15595	10.75	12.95
9	Riding Hydrogens added	1451	1088	220 / 38 / 40	15769	9.58	11.56
10	Minor adjustments	1477	1052	222 / 48 / 40	16114	9.15	11.19
11	Minor adjustments	1491	1042	211 / 64 / 48	16173	9.29	11.31
12	Weighting changed	1491	1029	222 / 84 / 38	16357	8.74	10.85
13	Further refinement	1499	1025	212 / 96 / 38	16353	8.76	10.79

NP = Number of protein atoms (including partially occupied atoms), NH = Number of hydrogens (all fully occupied), NW = Number of fully occupied waters, NW<sub>1/2</sub> = Number of half occupied waters, NX = Number of other atoms (inhibitor, formate, glycerol, some of them partially occupied), and N<sub>par</sub> = Number of least-squares parameters.

## 8.10 Summary of useful SHELXL keywords for macromolecular refinement

The more important keywords for macromolecular refinement are summarized in the following table (\* indicates significant changes from SHELXL-93):

---

DEFS	Set global restraint esd defaults.
DFIX	Restrain 1,2-distance to target (which may be a free variable).
DANG*	Restrain 1,3-distance to target (which may be a free variable).
SADI	Restrain distances to be equal without specifying target.
SAME	Generate SADI automatically for 1,2- and 1,3-distances using connectivity.
CHIV	Restrain chiral volume to target (default zero; may be a free variable).
FLAT*	Planarity restraint.
DELU	Generate rigid bond $U_{ij}$ restraints automatically using connectivity.
SIMU	Generate similar $U$ (or $U_{ij}$ ) restraints automatically using distances.
ISOR	'Approximately isotropic' restraints.
BUMP*	Generate anti-bumping restraints automatically (incl. symm. equivalents).
NCSY*	Generate non-crystallographic symmetry restraints.
FVAR	Starting values for overall scale factor and free variables.
SUMP	Restrain linear combination of free variables.
PART	Atoms with different non-zero PART numbers not connected by program.
AFIX	Riding H, rigid groups and other constraints on individual atoms.
HFIX	Generate hydrogens and suitable AFIX instructions for their refinement.
MERG	'MERG 4' averages equiv. reflns., incl. Friedel opp., sets all $f$ " to 0.
SHEL	Set maximum and minimum resolution (data ignored outside range).
STIR*	Stepwise increase of resolution.
SWAT	Refine diffuse solvent parameter (Babinet's principle).
WGHT	Weighting scheme, probably best left at default 'WGHT 0.1' throughout.
CGLS	No. of cycles conjugate gradient least-squares, select $R_{free}$ reflections.
BLOC,	L.S. Blocked-matrix least-squares (for esds).
RTAB, MPLA, HTAB*	Tables of bonds, angles, torsions, planes, H-bonds etc.
WPDB*, ACTA, LIST*	Output PDB and CIF files for archiving and data transfer..

---

I am grateful to Thomas R. Schneider for providing the serine protease example and helping with this chapter.