18.1.8.3. *Choice of optimization method*

First-order methods are generally the most economical for macromolecular problems. The most general approach is to treat the problem as a non-linear optimization problem from the beginning. This strategy is used by *TNT* (Tronrud *et al.*, 1987; Tronrud, 1997) and by *X-PLOR* (Kuriyan *et al.*, 1989), although by very different methods.

TNT uses a preconditioned conjugate gradient procedure (Tronrud, 1992), where the preconditioning function is the second derivatives of the objective function with respect to each parameter. In other words, at each step the parameters are normalized by the curvature with respect to that parameter, and a normal conjugate gradient step is taken. This has the effect that stiff parameters, which have steep derivatives, are scaled down, while soft parameters (such as *B* factors), which have small derivatives, are scaled up. This greatly increases both the rate and radius of convergence of the method.

X-PLOR (and its intellectual descendent, *CNS*) (Chapter 18.2 and Section 25.2.3) uses a simulated annealing procedure that derives sampling points by molecular dynamics. Simulated annealing is a process by which the objective function is sampled at a new point in parameter space. If the value of the objective function at the new point is less than that at the current point, the new point becomes the current point. If the value of the objective function is greater at the new point than at the current point, the Boltzmann probability $\exp(-\Delta E/kT)$ of the difference in function values ΔE is compared to a random number. If it is less than the random number, the new point is accepted as the current point; otherwise it is rejected. This process continues until a sufficiently deep minimum is found that the sampling process never leaves that region of parameter space. At this point the 'temperature' in the Boltzmann factor is reduced, which lowers the probability that the current point will move out of the region. This produces a finer search of the local region. The cooling process is continued until the solution has been restricted to a sufficiently small region. There are many variations of the strategy that affect the rate of convergence and the completeness of sampling. The primary virtue of simulated annealing is that it does not become trapped in shallow local minima. Simulated annealing can be either a zero-order method or a first-order method, depending on the strategy used to generate new sampling points. *X-PLOR* treats the fit to the diffraction data as an additional energy term, and the gradient of that 'energy' is treated as a force. This makes it a first-order method.

The first widely available macromolecular refinement program, *PROLSQ* (Konnert, 1976), uses an approximation to the second-order problem in which the matrix is greatly simplified. The parameters for each atom are treated as a small block on the diagonal of the matrix, and the off-diagonal blocks for pairs of atoms related by geometric constraints are also filled in. The sparse set of linear equations is then solved by an adaptation of the method of conjugate gradients.

The most comprehensive refinement program available for macromolecules is the same as the most comprehensive program available for small molecules – *SHELXL*98 (Sheldrick, 1993; see also Section 25.2.10). The primary adaptations to macromolecular problems have been the addition of conjugate gradients as an optimization method for cases in which the full matrix will not fit in the available memory and facilities to process the polymeric models required for macromolecules.

18.1.8.4. *Singularity in refinement*

Unless there are more linearly independent observations than there are parameters to fit them, the system of normal equations has no solution. The inverse of the matrix does not exist. Second-order methods fail in these circumstances by doing the matrix equivalent of dividing by zero. However, the objective function is still defined and has a defined gradient at all points. First-order methods will find a point at which the gradient is close to zero, and zero-order methods will still find a minimum value for the objective function. The difficulty is that the points so found are not unique. If one computes the eigenvalues and eigenvectors of the matrix of normal equations, one will find in this case that there are some eigenvalues that are very small or zero. The eigenvectors corresponding to these eigenvalues define sets of directions in which the parameters can be moved without affecting the value of the objective function. This region of the parameter space simply cannot be determined by the available data. The only recourses are to modify the model so that it has fewer parameters, add additional restraints to the problem, or collect more data. The real hazard with this situation is that the commonly used refinement methods do not detect the problem. Careful use of cross validation and keeping careful count of the parameters are the only remedy.

18.1.9. Evaluation of the model

Macromolecular model refinement is a cyclic process. No presently known refinement algorithm can remove all the errors of chain tracing, conformation, or misinterpretation of electron density. Other methods must be interspersed with refinement to help remove model errors. These errors are detected by basic sanity checks and the use of common sense about the model. This topic is discussed comprehensively in Part 21 and in Kleywegt (2000).

18.1.9.1. *Examination of outliers in the model*

Refinement-program output listings will normally provide some information on atoms that are showing non-standard bond lengths, bond angles or *B* factors. In addition, there is other software which can help identify non-standard or unusual geometry, such as *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT IF* (Vriend, 1990). These are very useful in identifying questionable regions of structure but should not be completely relied on to identify errors or how the molecular models may be improved. Overall, the constraints in the model must be satisfied exactly, and the restraints should have a statistically reasonable distribution of deviations from the ideal values.

18.1.9.2. *Examination of model electron density*

Refinement of the model to improve the agreement between the observed and calculated diffraction data and the associated calculated phases should result in improved electron-density and ΔF maps. Unexplained features in the electron-density map or difference map are a clear indication that the model is not yet complete or accurate. Careful examination of the Fourier maps is essential. Interactive graphics programs such as *XtalView* (McRee, 1993) and *O* contain a number of analysis tools to aid in the identification of errors in the models.

There are several different types of Fourier maps that can be useful in the correction of the models. This topic is discussed extensively in Chapter 15.2. Usual maps include F_o maps, ΔF maps and $(nF_o - mF_c)$ maps. The Fourier coefficients used to compute the maps should be weighted by estimates of the degree of bias as described in Chapter 15.2. While ΔF maps are very useful in highlighting areas in the maps that reflect the greatest difference between the F_o's and F_c's in Fourier space, they do not show the electron density of the unit cell. Positive and negative regions of a ΔF map may be the result of positional errors of an atom or group of atoms, *B*-factor errors, completely misplaced atoms or missing atoms. F_o maps show the electron density but are biased by the current model. A $(2F_o - F_c)$ map is a combination of an F_o map

and a ΔF map which results in a map better showing the changes due to errors. Some investigators prefer using further amplified ΔF contributions by using a $(3F_o - 2F_c)$ map or higher-order terms.

The contribution of the disordered solvent continuum has been discussed previously. Macromolecular crystals also contain significant quantities of discrete or partially discrete solvent molecules (*i.e.* water). Care needs to be taken in adding solvent to a model. Errors in models generate peaks in Fourier maps that can be interpreted as solvent peaks. Hence, adding solvent peaks too early in the refinement process may, in fact, lead to model errors. Automatic water-adding programs are becoming more common; examples include *SHELXL*98 and *ARP*/*wARP* (Lamzin & Wilson, 1997). These programs check if the waters are with in reasonable bonding distances of hydrogen-bonding atoms. There is a distribution of solvent molecules ranging from ones with low *B* factors at unit occupancy to ones with very large *B* factors. Various criteria are used to decide on a cutoff in the discrete solvent contribution. A rule of thumb for ambient-temperature data sets is frequently about one solvent molecule per residue in a protein molecule. As more data are being collected at cryogenic temperatures, this ratio is tending to go up. Noise is being fitted if too many peaks in a ΔF map are being assigned as solvent molecules. This can also contribute to reducing *R* factors on incorrect models. Solvent sites may not be fully occupied. Because of the large *B* factors and limited range of the diffraction data, the *B* factors and occupancy are highly correlated. Refinement of occupancy does not usually contribute either to improving a model or to reduction of *R* factors in structures with up to 2.0 Å resolution data. Beyond 1.5 Å data, it may be possible to refine solvent water occupancies and *B* factors. At even higher resolution, some programs, such as *SHELXL*98, provide anisotropic refinement methods which may further improve the solvent model while reducing *R* factors including $R_{\rm free}$.

18.1.9.3. *R and* $R_{\rm free}$

Cross validation is a powerful tool for avoiding over-interpretation of the data by a too elaborate model. The introduction of cross validation to crystallography (Brünger, 1992) has been responsible for significant improvement in the quality of structure determinations. A subset of the reflections, chosen randomly, is segregated and not used in the refinement. If the model is correct and the only errors are statistical, these reflections should have an *R* factor close to that of the reflections used in the refinement. Changes to the model should affect both *R* and $R_{\rm free}$ similarly. Kleywegt & Jones (1997) have pointed out that it is necessary to treat the selection of free reflections very carefully in the presence of noncrystallographic symmetry.

18.1.10. Conclusion

It is always important to bear in mind that macromolecular crystal structures are models intended to explain a particular set of observations. Statistical measures can determine how well the model explains the observations, but cannot say whether the model is true or not. The distinction between precision and accuracy must always be kept in mind. The objective should not be simply to obtain the best fit of a model to the data, but, in addition, to find all of the ways in which a model does *not* fit the data and correct them. Until the day when all crystals diffract to atomic resolution, the primary objective of refinement of the models will be to determine just how well the structures are or are not determined.

references