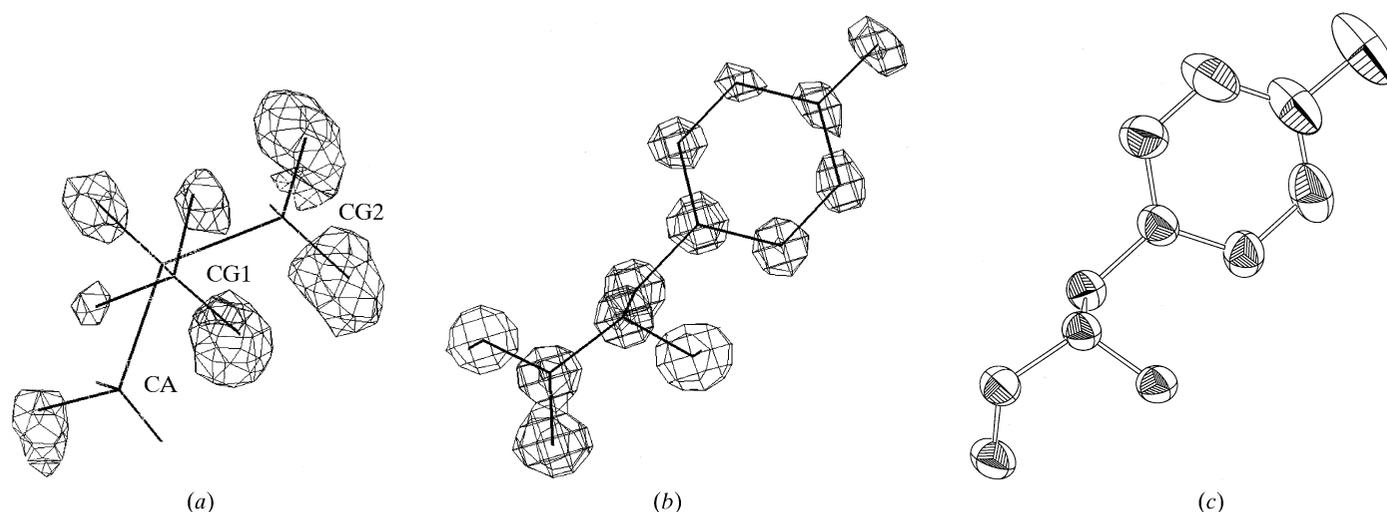


18.4. REFINEMENT AT ATOMIC RESOLUTION

**Figure 18.4.5.1**

(a), (b) Representative electron-density maps for the refinement of *Clostridium acidurici* ferredoxin at 0.94 Å resolution (Dauter, Wilson *et al.*, 1997). (a) The density for hydrogen atoms (at 3σ) omitted from the structure-factor calculation for Val42. (b) The $(2F_o - F_c)$ density for Tyr30, contoured at 3σ . (c) The thermal ellipsoids corresponding to (b), drawn at the 33% probability level using ORTEPII (Johnson, 1976). There is a clear correlation between the density in (b) and the ellipsoids in (c), showing increased displacement towards the end of the side chain, particularly in the plane of the phenyl ring.

prove difficult to refine the occupancy satisfactorily with statistical certainty.

18.4.4.4. Validation of extra parameters during the refinement process

The introduction of additional parameters into the model always results in a reduction in the least-squares or maximum-likelihood residual – in crystallographic terms, the R factor. However, the statistical significance of this reduction is not always clear, since this simultaneously reduces the observation-to-parameter ratio. It is therefore important to validate the significance of the introduction of further parameters into the model on a statistical basis.

Brünger (1992a) introduced the concept of statistical cross validation to evaluate the significance of introducing extra features into the atomic model. For this, a small and randomly distributed subset of the experimental observations is excluded from the refinement procedure, and the residual against this subset of reflections is termed R_{free} . It is generally sufficient to include about 1000 reflections in the R_{free} subset; further increase in this number provides little, if any, statistical advantage but diminishes the power of the minimization procedure. For atomic resolution structures, cross validation is important in establishing whether the introduction of an additional type of feature to the model (with its associated increase in parameters) is justified. There are two limitations to this. Firstly, if R_{free} shows zero or minimal decrease compared to that in the R factor, the significance remains unclear. Secondly, the introduction of individual features, for example the partial occupancy of five water molecules, can provide only a very small change in R_{free} , which will be impossible to substantiate. To recapitulate, at atomic resolution the prime use of cross validation is in establishing protocols with regard to extended sets of parameter types. The sets thus defined will depend on the quality of the data.

In the final analysis, validation of individual features depends on the electron density, and Fourier maps must be judiciously inspected. Nevertheless, this remains a somewhat subjective approach and is in practice intractable for extensive sets of parameters, such as the occupancies and ADPs of all solvent sites.

For the latter, automated procedures, which are being developed at present, are an absolute necessity, but they may not be optimal in the final stages of structure analysis, and visual inspection of the model and density is often needed.

The problems of limited data and reparameterization of the model remain. At high resolution, reparameterization means having the same number of atoms, but changing the number of parameters to increase their statistical significance, for example switching from an anisotropic to an isotropic atomic model or *vice versa*. In contrast, when reparameterization is applied at low resolution, this usually involves constraints, *i.e.*, a reduction in the number of independent atoms, but this is not an ideal procedure, as real chemical entities of the model are sacrificed. Reducing the number of independent atoms will inevitably result in disagreement between the experiment and model, which in turn will affect the precision of other parameters. It would be more appropriate to reduce the number of parameters without sacrificing the number of atoms, for example by describing the model in torsion-angle space. Water poses a particular problem, as at low as well as at high resolution the water molecules cannot all be described as discrete atoms. Algorithms are needed to describe them as a continuous model with only a few parameters. In the simplest model, the solvent can be described as a constant electron density.

18.4.5. Features in the refined model

All features of the refined model are more accurately defined if the data extend to higher resolution (Fig. 18.4.5.1). In this section, those features that are especially enhanced in an atomic resolution analysis are described. Introduction of an additional feature to the model should be assessed by the use of cross- or self-validation tools: only then can the significance of the parameters added to the model be substantiated.

18.4.5.1. Hydrogen atoms

Hydrogen atoms possess only a single electron and therefore have low electron density and are relatively poorly defined in X-ray studies. They play central roles in the function of proteins,

but at the traditional resolution limits of macromolecular structure analyses their positions can only be inferred rather than defined from the experimental data. Indeed, even at a resolution of 2.5 Å, hydrogen atoms should be included in the refined model, as their exclusion biases the position of the heavier atoms, but with their 'riding' positions fixed by those of the parent atoms.

As for small structures, independent refinement of hydrogen-atom positions is not always warranted, even by atomic resolution data, and hydrogen atoms are rather attached as riding rigidly on the positions of the parent atoms. Nevertheless, atomic resolution data allow the experimental confirmation of the positions of many of the hydrogen atoms in the electron-density maps, as they do account for one-sixth of the diffracting power of a carbon atom. Inspection of the maps can in principle allow the identification of (1) the presence or absence of hydrogen atoms on key residues, such as histidine, aspartate and glutamate or on ligands, and (2) the correct location of hydrogen atoms where more than one position is possible, such as in the hydroxyl groups of serine, threonine or tyrosine.

The correct placement of hydrogen atoms riding on their parent atoms involves computation of the appropriate position after each cycle of refinement. This is done automatically by programs such as *SHELXL*, *REFMAC* or *PHENIX.REFINE*. For rigid groups such as the NH amide, aromatic rings, $-\text{CH}_2-$ or $=\text{CH}-$, the position is accurately defined by the bonding scheme. For groups such as methyl CH_3 or OH, the position is not absolutely defined, and the software is required to make judgmental decisions. For example, *SHELXL* offers the opportunity to inspect the maximum density on a circular Fourier synthesis for optimal positioning. The bond length is fixed according to results from a small-molecule database. The location of hydrogen atoms on polar atoms can be assisted by software that analyses the local hydrogen-bonding networks; this involves maximization of the hydrogen-bonding potential of the relevant groups. Sheldrick advocates assigning ADP values to riding hydrogen atoms of 1.2 times that of the parent atom, or 1.5 times in the case of methyl and similar entities.

18.4.5.2. Anisotropic atomic displacement parameters

Refinement of an isotropic model involves four independent parameters per atom, three positional and one isotropic ADP. In contrast, an anisotropic model requires nine parameters per atom, with the anisotropic atomic displacement described by an ellipsoid represented by six parameters. At 1 Å resolution, the data certainly justify an anisotropic atomic model. Extension of the model from isotropic to anisotropic should generally result in a reduction in the *R* factor of the order of 5–6% and a comparable drop in R_{free} . As a consequence of the diminution of the observable-to-parameter ratio, the *R* factor at all resolutions will drop by a similar amount; however, R_{free} will not. Experience shows that at 2 Å or less there is no drop in R_{free} , and an anisotropic model is totally unsupported by the data. At intermediate resolutions, the result depends on the data quality and completeness. At lower resolution, to account for anisotropy of the atoms, the overall motion of molecules or domains can be refined using translation/libration/screw (TLS) parameters (Schomaker & Trueblood, 1968).

Until the end of the 1990s, anisotropic ADPs had only been handled by programs originally developed for small-molecule analysis, which use conventional algebraic computa-

tions of the calculated structure-factor amplitudes, *SHELXL* being a prime example. A limitation of this approach is the substantial computation time required. The use of fast-Fourier-transform algorithms for the structure-factor calculation leads to a significant saving in time (Murshudov *et al.*, 1999). Anisotropic modelling of the individual ADPs is essential if the thermal vibration is to be analysed in terms of coordinated motion of the whole molecule or of domains (Schomaker & Trueblood, 1968). Painter & Merritt (2006) have provided a means of analysing refined models to suggest appropriate TLS groupings.

18.4.5.3. Alternative conformations

Proteins are not rigid units with a single allowed conformation. *In vivo* they spontaneously fold from a linear sequence of amino acids to provide a three-dimensional phenotype that may exhibit substantial flexibility, which can play a central role in biological function, for example in the induced fit of an enzyme by a substrate or in allosteric conformational changes. Flexibility is reflected in the nature of the protein crystals, in particular the presence of regions of disordered solvent between neighbouring macromolecules in the lattice (see Section 18.4.5.6).

The structure tends to be highly ordered at the core of the protein, or more properly, at the core of the individual domains. Atoms in these regions in the most ordered protein crystals have ADP values comparable to those of small molecules, reflecting the fact that they are, in essence, closely packed by surrounding protein. In general, as one moves towards the surface of the protein, the situation becomes increasingly fluid. Side chains and even limited stretches of the main chain may show two (or multiple) conformations. These may be significant for the biological function of the protein.

The ability to model the alternative conformations is highly resolution dependent. At atomic resolution, the occupancy of two alternative but well defined conformations can be refined to an accuracy of about 5%, thus second conformations can be seen, provided that their occupancy is about 10% or higher. The limited number of proteins for which atomic resolution structures are available suggest that up to 20% of the 'ordered residues' show multiple conformations. This confers even further complexity on the description of the protein model. A constraint can be imposed on residues with multiple conformations: namely that the sum of all the alternatives must be 1.0. Protein regions (whether they are side- or main-chain regions) with alternative conformations and partial occupancy can form clusters in the unit cell with complementary occupancy. This often coincides with alternative sets of solvent sites, which should also be refined with complementary occupancies.

The atoms in two alternative conformations occupy independent and discrete sites in the lattice, about which each vibrates. However, if the spacing between two sites is small and the vibration of each is large, then it becomes impossible to differentiate a single site with high anisotropy from two separate sites. There is no absolute rule for such cases: programs such as *SHELXL* place an upper limit on the anisotropy and then suggest splitting the atom over two sites. Some regions can show even higher levels of disorder, with no electron density being visible for their constituent atoms. Such fully disordered regions do not contribute to the diffraction at high resolution, and the definition of their location will not be improved with atomic resolution data.

18.4. REFINEMENT AT ATOMIC RESOLUTION

18.4.5.4. Ordered solvent water

A protein crystal typically contains ~50% aqueous solvent. This is roughly divided into two separate zones. The first is a set of highly ordered sites close to the surface of the protein. The second, lying remote from the protein surface, is essentially composed of fluid water, with no order between different unit cells.

At room temperature, the solvent sites around the surface are assumed to be in dynamic equilibrium with the surrounding fluid, as for a protein in solution. Nevertheless, the observation of apparently ordered solvent sites on the surface indicates that these are occupied most of the time. The waters are organized in hydrogen-bonded networks, both to the protein and with one another. The most highly ordered water sites lie in the first solvent shell, where at least one contact is made directly to the protein. For the second and subsequent shells the degree of order diminishes: such shells form an intermediate grey level between the ordered protein and the totally disordered fluid. Indeed, the flexible residues on the surface form part of the continuum between a solid and liquid phase.

In the ordered region, the solvent structure can be modelled by discrete sites whose positional parameters and ADPs can be refined. For sites with low ADPs, the refinement is stable and their behaviour is well defined. As the ADPs increase, or more likely the associated occupancy in a particular site falls, the behaviour deteriorates, until finally the existence of the site becomes dubious. There is no hard cutoff for the reality of a weak solvent site. However, the number and significance of solvent sites are increased by atomic resolution data. Despite the fact that the waters contribute only weakly to the high-resolution terms, the improved accuracy of the rest of the structure and the reduction of noise due to the high resolution mean that their positions become better defined.

Indeed, the occupancy of some solvent sites can be refined if the resolution is sufficient, or at least their fractional occupancy can be estimated and kept fixed (Walsh *et al.*, 1998). This leads to the possibility of defining overlapping water networks with alternative hydrogen-bonding schemes. This can be a most time-consuming step in atomic resolution refinement, and a trade-off finally has to be made between the relevance of any improvement in the model and the time spent.

18.4.5.5. Automatic location of water sites

The protein itself has a clearly defined chemical structure, and the number of atoms to be positioned and how they are bonded to one another are known at the start of model building. The solvent region is in marked contrast to this, as the number of ordered water sites is not known *a priori*, and the distances between them are less well defined, their occupancy is uncertain, and there may be overlapping networks of partially occupied solvent sites. Those of low occupancy lie at the level of significance of the Fourier maps.

Selection of partially occupied solvent sites poses a most cumbersome problem in the modelling over and above that of the macromolecule itself, and can be highly subjective and very time consuming. Improved resolution of the data reveals additional weak or partially occupied solvent sites, which generally do not behave well during refinement. Water atoms modelled into relatively weak peaks in electron density tend to drift out of the density during refinement due to the inaccurate gradients that define their positions.

Given the huge number of water sites in question, automatic and at least semi-objective protocols are required. Several procedures have been developed for the automated identification of water sites during refinement [*inter alia* ARP (Lamzin & Wilson, 1997) and SHELXL (Sheldrick & Schneider, 1997)] and others allow selective inspection of such sites using graphics [such as O (Jones *et al.*, 1991) and COOT (Emsley & Cowtan, 2004)]. These depend on a combination of peak height in the density map and geometric considerations. However, these programs are currently optimized for structures at more typical resolutions, and future efforts could be made to adapt them for atomic resolution structures with overlapping water networks and other high levels of detail.

18.4.5.6. Bulk solvent and the low-resolution reflections

As stated in Section 18.4.5.3 and first reviewed by Matthews (1968) and more recently by Andersson & Hovmöller (1998), macromolecular crystals contain substantial regions of totally disordered, or bulk, aqueous solvent, in addition to those solvent molecules bound to the surface. The average electron density of the crystal volume occupied by protein is 1.35 g cm^{-3} (according to Matthews) or 1.22 g cm^{-3} (according to Andersson & Hovmöller), while that of water is 1.0 g cm^{-3} . This is because the atoms are more closely packed within the protein, as they are connected by covalent bonds, while in solvent regions they form sets of hydrogen-bonded networks.

To model both the solvent and protein regions of the crystal appropriately, it is necessary to have a satisfactory representation of the bulk solvent. The high *R* factors generally observed for most proteins for the low-resolution shells are in part symptomatic of the poor modelling of this feature or of systematic errors in the recording of the intensities of the low-angle reflections. For atomic resolution structures, the *R* factor can fall to values as low as 6–7% around 3–5 Å resolution. However, in lower-resolution shells it then rises steadily, often reaching values in the range of 20–40% below 10 Å. These observations indicate serious deficiencies in our current models or data.

The poorest approach is to ignore bulk solvent and assign zero electron density to those regions where there are no discrete atomic sites, as this leads to a severe discontinuum. An improved approach is to assign a constant value of the electron density to all points of the Fourier transform that are not covered by the discrete, ordered sites. This provides substantial reduction in the *R* factor for low-resolution shells of the order of 10% and requires the introduction of only one extra parameter to the least-squares minimization. An improvement of this simplistic model is the introduction of a second parameter, B_{sol} , described by

$$\text{scale} = k_0 \exp(-B_0 s^2/4) [1 - k_{\text{sol}} \exp(-B_{\text{sol}} s^2/4)], \quad (18.4.5.1)$$

where k_0 and B_0 are the scale factors for the protein, and k_{sol} and B_{sol} are the equivalent parameters for the bulk solvent (Tronrud, 1997). In effect, this provides a resolution-dependent smoothing of the interface contribution, rather than an overall term applied equally to all the data. The physical basis of this is discussed by Tronrud and implemented in several programs, for example SHELXL, REFMAC and PHENIX.REFINE (Fig. 18.4.5.2).

Another approach (Jiang & Brunger, 1994) to account for the effect of solvent is as follows: (1) calculate the mask covering the atoms in the crystal; (2) include a constant value for the electron density in the region not covered by the mask; (3) calculate structure factors from the solvent region; and (4) compute a total

18. REFINEMENT

structure by applying the appropriate scale to the solvent contribution:

$$F_{\text{total}} = F_{\text{prot}} + k_m \exp(-B_m s^2/4) F_{\text{mask}}, \quad (18.4.5.2)$$

where k_m and B_m are the scale and temperature factor for the solvent contribution, and F_{total} , F_{prot} and F_{mask} are the complex structure factors corresponding to the combined contribution from the protein and solvent region. The scale and B values are usually refined iteratively to find the best match to the observed structure factors, and are implemented in *CNS*, *REFMAC* and *PHENIX.REFINE*.

Nevertheless, there remain severe problems in the modelling of the interface. The border between the two regions is not abrupt, as there is a smooth and continuous change from the region with fully occupied, discrete sites to one which is truly fluid, but this passes through a volume with an increasing level of dynamic disorder and associated partial occupancy. Modelling of this region poses major problems, as described above, and the definition of disordered sites with low occupancy remains difficult even at atomic resolution. At which stage the occupancy and associated ADP can be defined with confidence is not yet an objective decision. In addition, refinement and modelling at this level of detail is very time consuming in terms of human intervention.

18.4.5.7. Metal ions and other ligands in the solvent

In general, proteins are crystallized from aqueous solutions which contain various additives, such as anions or cations (especially metals), organic solvents, including those used as cryoprotectants, and other ligands. Some of these may bind in specific or indeed non-specific sites in the ordered solvent shell, in addition to any functional binding sites of the protein. To identify such entities at limited resolution is often impossible, as the range of expected ADPs is large and there is very poor discrimination in the appearance of such sites and of water in the electron density. Atomic resolution assists in resolving ambiguities, as the interatomic distances, ADPs and occupancies are all better defined.

For metal ions, two additional criteria can be invoked. Firstly, the coordination geometry, with well defined bond lengths and angles, provides an indication of the identity of the ion, as different metals have different preferred ligand environments (Harding, 1999, 2006). The bond-valence approach is also applicable (Müller *et al.*, 2003). In addition, the value of the refined ADP and/or occupancy is helpful. Secondly, the anomalous signal in the data should reveal the presence of metal and some other non-water sites in the solvent through computation of the anomalous difference synthesis (Dauter & Dauter, 1999). This emphasizes the need to retain the anomalous signal during the collection and reduction of native data. While these approaches can be applied at lower resolution, they both become much more powerful at atomic resolution.

The presence of bound organic ligands has become especially relevant since the advent of cryogenic freezing. Compounds such as ethylene glycol and glycerol possess a number of functional hydrogen-bonding groups that can attach to sites on the protein in a defined way. Indeed, these may often bind in the active sites of enzymes such as glycosyl hydrolases, where they mimic the hydroxyl groups of the sugar substrate. It is most important to identify such moieties properly, particularly if substrate studies are to be planned successfully.

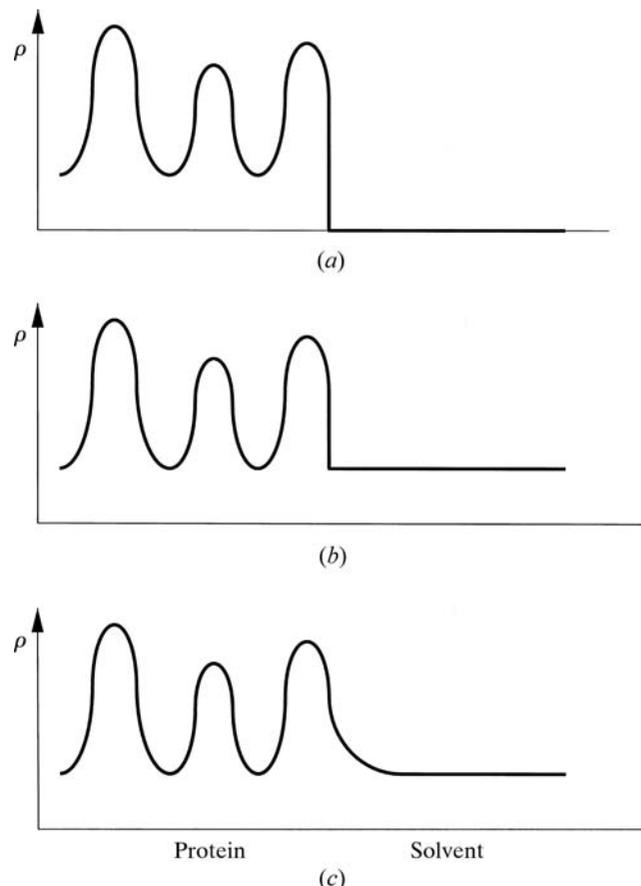


Figure 18.4.5.2

Schematic representation of the bulk-solvent models described in the text. (a) No bulk-solvent correction, *i.e.* solvent density set to zero. (b) Constant level of solvent outside the macromolecule and ordered water envelope. Here, sharp edge effects remain. (c) The model as in (b), but smoothed at the edge of a macromolecule, equivalent to the application of a B value to the solvent model.

18.4.5.8. Deformation density

X-ray structures are generally modelled using the spherical-atom approximation for the scattering, which ignores the deviation from sphericity of the outer bonding and lone-pair electrons. Extensive studies over a long period have confirmed that the so-called deformation density, representing deviation from this spherical model, can be determined experimentally using data to very high resolution, usually from 0.8 to 0.5 Å. An excellent review of this field has been provided by Coppens (1997). The observed deviations can be compared with those expected from the available theories of chemical bonding and the densities derived therefrom.

The application of atomic resolution analysis to proteins allowed the observation of the deformation density in macromolecules (Lamzin *et al.*, 1999). Data for two proteins were analysed: crambin (molecular weight 6 kDa) at 0.67 Å resolution and a subtilisin (molecular weight 30 kDa) at 0.9 Å resolution. Significant and interpretable deformation density could not be observed for the individual residues. However, on averaging the density over 40 peptide units for crambin and more than 250 for the subtilisin, the deformation density within the peptide unit was clearly visible and could be related to the expected bonding features in these units. This shows the real power of atomic resolution crystallography, which can reveal features containing no more than $0.2 \text{ e } \text{Å}^{-3}$.

Deformation density studies are now being applied to many polypeptides (Jelsch *et al.*, 1998; Koritsanszky *et al.*, 2002; Pichon-

Pesme *et al.*, 2004; Afonine *et al.*, 2007; Zarychta *et al.*, 2007). It has been observed that details of deformation densities are most clearly revealed when only the highest resolution (so-called high order) terms are included in the refinement and the Fourier maps (Coppens, 1997). This is reported to result from the deconvolution of the effects of the anisotropic ADPs, which can to some extent take up the fine features corresponding to the deformation density (bonding electrons and lone pairs). After proper modelling of the deformation density features, the overall model is refined against all the data.

18.4.6. Quality assessment of the model

The refinement of proteins at resolutions lower than atomic depends upon the use of restraints on the geometry and ADPs. The almost exclusively used library of target geometric restraints for refinement and validation of protein structures (Chapter 18.3) is derived from structures of amino acids and peptides in the Cambridge Crystallographic Data Centre's small-molecule crystal structure database (Allen *et al.*, 1979). Stereochemical parameters, such as conformational angles φ , ψ , should ideally not be restrained, as they allow independent validation of the model. As stated in Section 18.4.4.2, these restraints are required even at atomic resolution to maintain the chemical integrity of flexible regions, although their impact will be limited on ordered regions.

Owing to the excess of accurate X-ray observations over parameters at atomic resolution, extensive validation of individual structures should be less challenging than for those at lower resolution. It is hard to achieve an *R* factor around 10% with an incorrect model. However, considerable attention needs to be given to detail and great care taken to avoid over-interpretation, especially of the flexible regions.

An analysis of eight structures determined at atomic resolution some years ago (EU 3-D Validation Network, 1998) indicated that they follow the expected rules of chemistry more closely than those of lower-resolution analyses in the PDB, confirming that atomic resolution indeed provides more precise coordinates. A subsequent analysis of ten structures at ultra-high resolution, 0.8 Å or better (Jaskolski *et al.*, 2007) confirmed these conclusions but identified a few possible adjustments to some targets. Following this analysis, Karplus *et al.* (2008) proposed that protein stereochemistry is context dependent, *e.g.* it differs in detail between α -helices and β -strands, and that this should be reflected in future target libraries. Thus, the availability of atomic resolution structures will provide a more objective basis for the construction of such libraries.

18.4.7. Relation to biological chemistry

A question arises as to what biological issues are addressed by analysis of macromolecular structures at atomic resolution. For any protein, the overall structure of its fold, and hence its homology with other proteins, can already be provided by analyses at low to medium resolution. However, proteins are the active entities of cells and carry out recognition of other macromolecules, ligand binding and catalytic roles that depend upon subtle details of chemistry, for which accurate positioning of the atoms is required. Even at atomic resolution, the accuracy of structural definition is less than what would ideally be required for the changes observed during a chemical reaction. At lower resolutions, structure–function relations require yet further extrapolation of the experimental data.

To understand the function of many macromolecules, such as enzymes, it is not sufficient to determine the structure of a single state. Alongside the native structure, those of various complexes will also be required. The differences between the states provide additional information on the functionality. For an understanding of the chemistry involved, atomic resolution has tremendous advantages in terms of accuracy, as reliable judgments can be based on the experimental data alone.

Advantages of atomic resolution include the following:

- (1) The positions of all atoms that possess defined conformations are more accurately defined. This means that all bond lengths and angles in the structure have lower standard uncertainties (EU 3-D Validation Network, 1998). For regions of the molecule where the conformation is representative of the norm, this is of purely quantitative significance, but where the stereochemistry deviates from the expected value this accuracy takes on a special significance, which poses questions to the theoretical chemist. Such deviations from standard geometry often play an important role in biological function.
- (2) The better the ADP definition, notably its anisotropy, the greater the insight into the static or thermal flexibility of individual regions of the molecule. Macromolecules are crucially dependent upon flexibility for properties such as induced fit in substrate or ligand recognition, allosteric responses or responses to the biological environment. More detailed definition of the position and mobility of flexible regions may be assisted by atomic resolution analysis.
- (3) A few amino-acid side chains play an active role in catalysis. Those that do include histidine, aspartic and glutamic acids and serine, all through protonation–deprotonation events, and hydrogen atoms are crucial to their function. Hydrogen atoms are usually treated as riding on their parent atoms and should be included in the model, even at medium resolution. Unfortunately, those hydrogen atoms that are of interest can only rarely be treated as rigidly bonded at a predictable position. However, atomic resolution allows many hydrogen atoms to be clearly identified in the refined electron density. In addition, the presence or absence of hydrogen may be inferred by accurate estimation of the bond lengths between atoms, *e.g.* within the carboxylate groups.
- (4) The relative orientation of reacting moieties is crucial to enzyme catalysis. If chemical hypotheses of mechanism are to be subjected to appropriate Popperian scrutiny (Popper, 1959), then precise definition of atomic coordinates in native and complex structures is necessary.
- (5) Enzyme catalysis provides a reduction of the activation energy of the reaction, which can be achieved by distortion of the conformation of the substrate bound to the enzyme (the so-called Michaelis complex) towards the transition state or by the stabilization of the latter by the enzyme. For both, the study of complexes of inhibitors or substrate analogues at a sufficient resolution to clarify the fine detail of the structures is required.
- (6) Adaptation of the enzyme to the substrate is postulated by the induced-fit theory of catalysis. The level of adjustment can be very small, and energy calculations again require that this be precisely defined.
- (7) In metalloproteins, the ligand field, and hence geometry and bond lengths, around the metal ion are essential indicators of any variation in valence electrons between different states. For example, bond lengths between oxidized and reduced states of metal ions vary by the order of 0.1 Å or less, and