clear distinction between alternative oxidation states requires an accuracy only provided by atomic resolution.

Almost all atomic resolution analyses require data recorded from cryogenically frozen crystals. This does pose some problems of biological relevance, as proteins *in vivo* have adapted to operate at ambient cellular temperatures. The structure that is required is that of the protein and surrounding solvent at the corresponding temperature. The trade-off is that cryogenic structures may be better defined, but that they are only so because of the increased order of the protein and solvent at low temperature. This has to be weighed against the lack of fine detail in a medium-resolution analysis at room temperature.

A question often raised with regard to the worth of atomic resolution data concerns the effort required in refining a protein at such resolution. To define all details, such as alternative conformations, hydrogen-atom positions and solvent, is certainly time consuming, especially if an anisotropic model is adopted. However, the advantages outweigh the disadvantages, as even if a full anisotropic model is not refined to exhaustion, nevertheless all density maps will be clearer if the resolution is better, resulting in an improved definition of the features of interest.

18.4.8. Practical strategies

Pioneering work was carried out by Teeter and colleagues on crambin, using data recorded on this small and highly stable protein using a conventional diffractometer (Teeter *et al.*, 1993). Some of the earliest atomic resolution structures using data from a synchrotron source with an imaging plate detector included rubredoxin at 1.0 Å (Dauter *et al.*, 1992), ribonuclease Sa at 1.1 Å (Sevcik *et al.*, 1996) and triclinic lysozyme at 0.9 Å resolution (Walsh *et al.*, 1998), at room temperature for the rubredoxin and ribonuclease, and 100 K for the lysozyme. The strategy used involved the application of conventional restrained least-squares or maximum-likelihood techniques in the early stages of refinement, followed by switching over to *SHELXL* to introduce a full anisotropic model and riding hydrogen atoms.

Subsequent structures have almost exclusively used cryogenically vitrified crystals to minimize the effects of radiation damage on high-intensity SR beamlines. A review of the field (Dauter, 2003) reported over 100 structures at better than 1.2 Å spacing, and in January 2011 there were around 1450 such coordinate sets in the PDB.

Initially, only *SHELXL* was able to refine macromolecular structures with anisotropic ADPs, riding hydrogen atoms and appropriate stereochemical restraints. Subsequently, such options have been implemented in programs such as *REFMAC* and *PHENIX.REFINE*. The use of fast Fourier algorithms gives these programs some speed advantage over *SHELXL*, but this is now of less importance given the power of contemporary computers. The key advantage of *SHELXL* is the ability to perform full- (or at least block-) matrix inversion, and hence extract error estimates for each individual parameter.

It is not straightforward to provide hard and fast rules for the optimum strategy. The power of refinement at atomic resolution lies in the sheer number of X-ray observations which drive the minimization to a global minimum. The dangers of bias in the final model and of false minima can largely be ignored. As a guideline the following steps would seem to be a common-sense approach when data to 1.2 Å or better are available.

(1) Refine an initial isotropic model at modest resolution with one of the conventional programs to ensure the correct global minimum has been reached. This step was probably more appropriate in the days when computing resources were more limiting, and could probably be omitted today.
(2) Introduce an anisotropic model with riding hydrogen atoms using one of the maximum-likelihood programs such as *REFMAC*, *PHENIX.REFINE* or *SHELXL* (with the conjugate-gradient option), and refine against the experimental data, omitting around 1000–2000 reflections for validation by $R_{\rm free}$. It is *not* necessary to omit 5% of the data, particularly at this resolution.
(3) Complete the water model using one of the autobuild options within or associated with (*e.g. ARP/wARP* or *COOT*) the refinement program. This is an area amenable to considerable automation in the future to allow partially occupied sites and overlapping networks to be defined without user intervention.
(4) Introduce alternative conformations for the protein. Much of this is also amenable to automation, especially for the side chains, where programs such as *COOT* can already suggest likely alternate rotamers.
(5) Inspect appropriately weighted electron-density maps for features to be added, deleted or moved. This includes ligands, metal atoms, ions, cryprotectants and other additives. Current developments in programs such as *COOT* are increasingly automating this step. Software libraries that provide accurate descriptions of the ligand chemistry and geometry are essential, and are the subject of intense development. It is vital that structures are deposited in the PDB with the correct geometrical restraints. To some extent, steps (3)–(5) go in parallel, and indeed their order is not vital.
(6) Consider the *SHELXL* full-matrix option for the final refinement cycles. This will provide estimates of the accuracy of the individual parameters, not available from other programs.
(7) As a final step, run some cycles with all reflections including those previously used for $R_{\rm free}$, but without changing the model parameters. At all stages, it will remain necessary to impose restraints to avoid poor behaviour of any flexible regions. As the resolution goes beyond 1 Å, the effect of such restraints on good parts of the structure becomes minimal.

In summary, at atomic resolution models require special attention to detail, reflecting the wealth of information with regard to features such as alternative conformations of both side and main chains, and extensive ordered water structure often with partial occupancy and overlapping networks. With the present software, this can be a lengthy operation, and indeed this can persuade some scientists to work at lower, less informative, resolution to speed up and simplify the analysis. Many of the presently tedious manual tasks in terms of model construction are, however, amenable to extensive automation. As the number of atomic resolution structures increases, we can expect the refinement and modelling process to be greatly simplified with new software algorithms.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Adams, P. D., Lunin, V. Y. & Urzhumtsev, A. (2007). *On macromolecular refinement at subatomic resolution with interatomic scatterers*. *Acta Cryst.* D**63**, 1194–1197.

Agarwal, R. C. (1978). *A new least-squares refinement technique based on the fast Fourier transform algorithm*. *Acta Cryst.* A**34**, 791–809.

Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O.,