

## 18. REFINEMENT

18.3.2.9. *Addition of tailored information sources*

When specific structural effects are observed which are not otherwise parameterized, new parameters might be desirable to encode this information. In the case of new statistical minima or variances of parameters already encoded by the refinement program, it is for most programs a simple matter to introduce new atom types, residues, or fragment names to the topology and parameter libraries. Examples include new parameterization for charged states or for *cis*- and *trans*-proline. If the reason for the new statistics is structural, the structure must be appropriately monitored during refinement to ensure that the conditions continue to hold, particularly in the case of simulated-annealing refinement steps.

18.3.3. **Strategy of application during building/refinement**

Refinement parameters necessarily and intentionally introduce bias into the refinement which may not disappear with later alterations of parameters. The importance of this fact is reflected by the observation that the parameters of refined structures can be recognized by statistical studies of the structures (Laskowski, Moss & Thornton, 1993). It is therefore important that the parameters initially reflect what can be confidently predicted about the structure. If unknown geometries may be expected, at metal or catalytic sites, for example, or if isomerization states need to be recognized from the refined structure, all relevant parameters must be initially eliminated from the refinement. Depending on the resolution of the structure and the detail required, the unbiased final refined structure may sufficiently demonstrate the unknown structural quantities. On the other hand, insufficient restraint may allow unreasonable geometries that do not allow recognition of the desired quantity. In this case, it may be necessary to test all possible restraint conditions and compare the results of the refinements.

18.3.3.1. *Confidence in restraints versus information from diffraction*

Primarily in cases of new structures, such as small-ligand- or metal-binding proteins, the refinement may indicate that the expected geometries and applied restraints seem incompatible with evidence from the electron density. Several sources for such discrepancies must be considered for an evaluation of the true geometries or the confidence level of such an evaluation. The

quality of the experimental information, such as data resolution and reduction parameters, must be considered. Physical phenomena possibly ignored by the refinement model might include anisotropies of motion and/or electron distribution, or disorder in the crystal. These might lead to systematic deviations in the refined structure that mimic alternate parameterizations. Finally, newly derived parameters should be examined to decide whether the fragments and chemical environments were inappropriate for the refinement problem, or whether errors in fragment structures artifactually distorted the parameterization.

18.3.4. **Future perspectives**

It seems obvious to seek the best (most accurate) possible parameterization and establish it as a standard (to enable statistical structure comparisons). This does not seem to be a realistic goal for several reasons. Firstly, the parameterization is less a determinant of accuracy than the quality of the data and the method of refinement. Secondly, the quality of existing parameterization and the potential for new environment-dependent parameters improves as more structures are solved and databases grow. Such new parameters can be derived from conformation-dependent statistics (*cis*- and *trans*-proline is an example described above), hydrogen-bonding geometries *etc.* Finally, protein structures are generally solved not to build a statistically optimized protein database, but to discover biophysical functional mechanisms.

The growth of structural databases will improve our understanding of structural properties (Wilson *et al.*, 1998); the highest-resolution protein structures will contribute most to the database, while low-resolution structures will profit most from improved predictive power. Structures that require restrained refinement both draw on the database for refinement parameters and integrity checks, and also contribute to it; a kind of boot-strapping procedure to re-refine deposited structures with iteratively improved parameters is conceivable (if convergent). The consequent removal of parameterization 'signatures' in, *e.g.*, bond and angle parameters seems unlikely to have practical consequences beyond identification of, *e.g.*, catalytically relevant outliers, but qualitative improvements in structure comparison might be revealing in unexpected ways. Such an effort will require adequate computational resources and the deposition of structure factors or, even better, diffraction images.