8.5.4. Plausibility of results

A study of residuals to detect a pattern of discrepancies will reveal the presence of systematic error, or model inadequacy, only if different subsets of the data are affected differently. Some sources of bias, however, have noticeable effects throughout the data set, and missing parameters may mimic others that have been included, thus introducing bias without any apparent lack of fit. To cite an obvious example, determination of unit-cell dimensions requires an accurate value of the wavelength of the radiation being used. If this value is incorrect, inferred values of the cell constants may be reproduced repeatedly with great precision, but all will be subject to a systematic bias that has no effect on the quality of the fit. Even though the structure of the residuals in such a case reveals little about possible systematic error, it is still possible to detect it by critical examination of the estimated parameters.

Even before any data have been collected in preparation for the determination of a crystal structure, a great deal is known about certain details. It is known that the crystal is composed of atoms of certain elements that are present in certain proportions. It is known that pairs of atoms of various elements cannot be less than a certain distance apart, and, further, that adjacent atoms tend to be separated by distances that fall within a rather narrow range. It is known that thermal vibration amplitudes are likely to be larger in directions normal to interatomic vectors than parallel to them, although, particularly in the case of hydrogen bonds, there may be an apparent amplitude parallel to a vector because of atomic disorder. Even when there is a particularly unusual feature in a structure, most of the structure will conform to commonly observed patterns. Thus, if a refined crystal structure overall has reasonable features, such as interatomic distances that are appropriate to oxidation state and coordination number and displacement ellipsoids that make sense, one or two unusual features may be accepted with confidence. On the other hand, if there is wide variation in the lengths of chemically similar bonds, or if the eigenvectors of the thermal motion tensors point in odd directions relative to the interatomic vectors, there must be a presumption that systematic errors have been compensated for by biased estimates of parameters.

A particular problem arises when there is a question of the presence or absence of symmetry, such as a choice between two space groups, one of which possesses a centre of symmetry or a mirror plane, or a case where a symmetric molecule occupies a position whose environment has a less-symmetric point group. If symmetry constraints are relaxed, the model can always be refined to a lower sum of squared residuals. (For a discussion of numerical problems that occur in the vicinity of a symmetric configuration, see Section 8.1.3.) The problem comes from the fact that the removal of the symmetry element almost always introduces too many additional parameters. Statistical tests are then quite likely to indicate that the lower-symmetry model gives a significantly better fit, but consideration of internal consistency and chemical or physical plausibility is likely to suggest that much systematic error has been absorbed by the additional parameters. The proper procedure is to devise a model with noncrystallographic constraints (see Section 8.3.2) that expresses what is, for chemical or physical reasons, known or probable. To do so may require great patience and perseverence.

references