The refinement scheme described above requires initial scaling factors $g_j$. With the now improved estimates for the partialities $R_j$, a new set of scaling factors can be obtained by the method outlined in Section 11.3.4. This alternating procedure of scaling and post refinement usually converges within three cycles.

The use of error functions for modelling partiality, as implicated by a Gaussian model for describing spot shape, was chosen here for reasons of conceptual simplicity and coherence. This choice is unlikely to alter significantly the results of post refinement that are based on other functions of similar form [see the discussion by Rossmann (1985)].

## 11.3.6. Space-group assignment

Identification of the correct space group is not always an easy task and should be postponed for as long as possible. Fortunately, all data processing as implemented in the program *XDS* can be carried out even in the absence of any knowledge of crystal symmetry and cell constants. In this case, a reduced cell is extracted from the observed diffraction pattern and processing of the data images continues to completion as if the crystal were triclinic. Clearly, the reflection indices then refer to the reduced cell and must be reindexed once the space group is known. For all space groups, the required reindexing transformation is linear and involves only whole numbers as shown in Part 9 of IT *A*. The following description and example are taken from Kabsch (1993).

Space-group assignment is carried out in two steps under control of the crystallographer once integrated intensities of all reflections are available. First, the Bravais lattices that are compatible with the observed reduced cell are identified. In the second step, any of the plausible space groups may be tested and rated according to symmetry *R* factors and systematic absences of integrated reflection intensities after reindexing. Additional acceptance criteria are obtained from refinement, now using a reduced set of independent parameters describing the conventional unit cell which should not lead to a significant increase of r.m.s. deviations between observed and calculated reflection positions and angles.

### 11.3.6.1. *Determination of the Bravais lattice*

The determination of possible Bravais lattices is based upon the concept of the reduced cell whose metric parameters characterize 44 lattice types as described in Part 9 of IT *A*. A primitive basis $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ of a given lattice is defined there as a reduced cell if it is right-handed and if the components of its metric tensor

$$A = \mathbf{b}_1 \cdot \mathbf{b}_1, \quad B = \mathbf{b}_2 \cdot \mathbf{b}_2, \quad C = \mathbf{b}_3 \cdot \mathbf{b}_3,$$
$$D = \mathbf{b}_2 \cdot \mathbf{b}_3, \quad E = \mathbf{b}_1 \cdot \mathbf{b}_3, \quad F = \mathbf{b}_1 \cdot \mathbf{b}_2$$

satisfy a number of conditions (inequalities). The main conditions state that the basis vectors are the shortest three linear independent lattice vectors with either all acute or all non-acute angles between them. As specified in IT *A*, each of the 44 lattice types is characterized by additional equality relations among the six components of the reduced-cell metric tensor. As an example, for lattice character 13 (Bravais type *oC*) the components of the metric tensor of the reduced cell must satisfy

$$A = B, \quad B \le C, \quad D = 0, \quad E = 0, \quad 0 \le -F \le A/2.$$

Any primitive triclinic cell describing a given lattice can be converted into a reduced cell. It is well known, however, that the reduced cell thus derived is sensitive to experimental error. Hence, the direct approach of first deriving the correct reduced cell and then finding the lattice type is unstable and may in certain cases even prevent the identification of the correct Bravais lattice.

A suitable solution of the problem has been found that avoids any decision about what the 'true' reduced cell is. The essential requirements of this procedure are: (*a*) a database of possible reduced cells and (*b*) a backward search strategy that finds the best-fitting cell in the database for each lattice type.

The database is derived from a seed cell which strictly satisfies the definitions for a reduced cell. All cells of the same volume as the seed cell whose basis vectors can be linearly expressed in terms of the seed vectors by indices $-1$, 0, or $+1$ are included in the database. Each unit cell in the database is considered as a potential reduced cell even though some of the defining conditions as given in Part 9 of IT *A* may be violated. These violations are treated as being due to experimental error.

The backward search strategy starts with the hypothesis that the lattice type is already known and identifies the best-fitting cell in the database of possible reduced cells. Contrary to a forward directed search, it is now always possible to decide which conditions have to be satisfied by the components of the metric tensor of the reduced cell. The total amount by which all these equality and inequality conditions are violated is used as a quality index. This measure is defined below for lattice type 13 *oC* testing a potential reduced cell $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ from the database for agreement. Positive values of the quality index $p_{13}$ indicate that some conditions are not satisfied.

$$p_{13}(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = |A - B| + \max(0, B - C) + |D| + |E| + \max(0, F) + \max(0, -F - A/2).$$

All potential reduced cells in the database are tested and the smallest value for $p_{13}$ is assigned to lattice type 13. This test is carried out for all 44 possible lattice types using quality indices derived in a similar way from the defining conditions as listed in Part 9 of IT *A*. For each of the 44 lattice types thus tested, the procedure described here returns the quality index, the conventional cell parameters and a transformation matrix relating original indices with respect to the seed cell to the new indices with respect to the conventional cell. These index-transformation matrices are derived from those given in Table 9.3.1 in IT *A*.

The results obtained by this method are shown in Table 11.3.6.1 for the example of a 1.5° oscillation data film containing 1313 strong diffraction spots which were located automatically. The space group of the crystal is $C222_1$ and the cell constants are $a = 72.9, b = 100.1, c = 92.6$ Å. The entry for the correct Bravais lattice *oC* with derived cell constants close to the true ones has a low value for its quality index and thus appears as a possible explanation of the observed diffraction pattern.

### 11.3.6.2. *Finding possible space groups*

Inspection of the table rating the likelihood of each of the 44 lattice types usually reveals a rather limited set of possible space groups. Furthermore, the absence of parity-changing symmetry operators required for protein crystals restricts the number of possible space groups to 65 instead of 230. Any space group can be tested by repeating only the final steps of data processing. These steps include a comparison of symmetry-related reflection intensities, as well as a refinement of the parameters controlling the diffraction pattern after reindexing the reflections by the appropriate transformation. Low r.m.s. deviations between the observed and refined spot positions, as well as small *R* factors for symmetry-related reflection intensities, indicate that the constraints imposed by the tentatively chosen space group are satisfied. The space group with highest symmetry compatible with the data is almost certainly correct if the data set is sufficiently complete and redundant, which requires that each symmetry element relates a sufficient number of reflections to one another.

For the example of a 1.5° oscillation data film given above, space-group determination consists of the following steps. Inspection of Table 11.3.6.1 indicates that lattice characters 10, 13, 14 and 34, besides the triclinic characters 31 and 44, are approximately

Table 11.3.6.1. *Rating of lattice types implied by a given reduced cell*

| Lattice type | Quality index | Conventional cell constants (Å, °) | | | | | | Reindexing transformation |
|---|---|---|---|---|---|---|---|---|
| | | *a* | *b* | *c* | $\alpha$ | $\beta$ | $\gamma$ | |
| 1 *cF* | 999.0 | 119.3 | 137.3 | 119.1 | 121.1 | 77.5 | 122.6 | $11\bar{1}0/1\bar{1}10/\bar{1}\bar{1}\bar{1}0$ |
| 2 *hR* | 770.1 | 74.6 | 111.7 | 137.4 | 103.6 | 89.1 | 108.8 | $1100/\bar{1}0\bar{1}0/\bar{1}110$ |
| 3 *cP* | 769.7 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |
| 5 *cI* | 936.0 | 111.7 | 74.6 | 112.6 | 70.1 | 53.6 | 71.2 | 1010/1100/0110 |
| 4 *hR* | 769.5 | 101.1 | 111.9 | 119.1 | 116.5 | 89.1 | 116.4 | $1\bar{1}00/\bar{1}010/\bar{1}\bar{1}\bar{1}0$ |
| 6 *tI* | 999.0 | 112.6 | 111.7 | 74.6 | 71.2 | 70.1 | 53.6 | 0110/1010/1100 |
| 7 *tI* | 999.0 | 111.7 | 74.6 | 112.6 | 70.1 | 53.6 | 71.2 | 1010/1100/0110 |
| 8 *oI* | 999.0 | 74.6 | 111.7 | 112.6 | 53.6 | 70.1 | 71.2 | $\bar{1}\bar{1}00/\bar{1}0\bar{1}0/0\bar{1}\bar{1}0$ |
| 9 *hR* | 772.7 | 62.1 | 74.6 | 296.6 | 90.5 | 105.8 | 125.5 | $1000/\bar{1}\bar{1}00/\bar{1}1\bar{3}0$ |
| 10 *mC* | 24.0 | 101.1 | 74.6 | 92.9 | 90.1 | 90.0 | 91.3 | $1\bar{1}00/1100/0010$ |
| 11 *tP* | 174.8 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |
| 12 *hP* | 122.8 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |
| 13 *oC* | **23.8** | 74.6 | 101.1 | 92.9 | 90.0 | 90.1 | 88.7 | $1100/\bar{1}100/0010$ |
| 15 *tI* | 672.7 | 62.1 | 63.5 | 200.2 | 77.0 | 77.6 | 107.2 | 1000/0100/1120 |
| 16 *oF* | 999.0 | 74.6 | 101.1 | 200.2 | 90.5 | 111.8 | 88.7 | $\bar{1}\bar{1}00/1\bar{1}00/1120$ |
| 14 *mC* | 23.4 | 74.6 | 101.1 | 92.9 | 90.0 | 90.1 | 88.7 | $1100/\bar{1}100/0010$ |
| 17 *mC* | 999.0 | 101.1 | 74.6 | 111.7 | 71.2 | 116.4 | 88.7 | $1\bar{1}00/\bar{1}\bar{1}00/\bar{1}0\bar{1}0$ |
| 18 *tI* | 999.0 | 112.6 | 119.1 | 62.1 | 68.7 | 99.5 | 115.4 | $01\bar{1}0/1110/1000$ |
| 19 *oI* | 999.0 | 62.1 | 112.6 | 119.1 | 64.6 | 68.7 | 80.5 | $\bar{1}000/01\bar{1}0/\bar{1}\bar{1}\bar{1}0$ |
| 20 *mC* | 746.3 | 112.6 | 112.6 | 62.1 | 99.5 | 99.6 | 111.3 | $0\bar{1}\bar{1}0/0\bar{1}10/\bar{1}000$ |
| 21 *tP* | 748.0 | 63.5 | 92.9 | 62.1 | 90.1 | 107.2 | 90.0 | 0100/0010/1000 |
| 22 *hP* | 999.0 | 63.5 | 92.9 | 62.1 | 90.1 | 107.2 | 90.0 | 0100/0010/1000 |
| 23 *oC* | 747.8 | 112.6 | 112.6 | 62.1 | 80.5 | 99.6 | 68.7 | $0110/0\bar{1}10/1000$ |
| 24 *hR* | 999.0 | 154.8 | 112.6 | 62.1 | 80.5 | 80.9 | 84.3 | $1210/0\bar{1}10/1000$ |
| 25 *mC* | 746.1 | 112.6 | 112.6 | 62.1 | 80.5 | 99.6 | 68.7 | $0110/0\bar{1}10/1000$ |
| 26 *oF* | 624.9 | 62.1 | 123.9 | 195.9 | 86.4 | 108.4 | 101.5 | $1000/\bar{1}\bar{2}00/\bar{1}0\bar{2}0$ |
| 27 *mC* | 499.7 | 123.9 | 62.1 | 112.6 | 80.5 | 119.7 | 78.5 | $\bar{1}\bar{2}00/\bar{1}000/01\bar{1}0$ |
| 28 *mC* | 325.0 | 62.1 | 195.9 | 63.5 | 95.4 | 107.2 | 71.6 | $\bar{1}000/\bar{1}0\bar{2}0/0\bar{1}00$ |
| 29 *mC* | 99.8 | 62.1 | 123.9 | 92.9 | 90.0 | 90.1 | 78.5 | 1000/1200/0010 |
| 30 *mC* | 336.4 | 63.5 | 196.5 | 62.1 | 95.4 | 107.2 | 71.1 | $0\bar{1}00/0\bar{1}20/\bar{1}000$ |
| 31 *aP* | 0.2 | 62.1 | 63.5 | 92.9 | 90.0 | 89.9 | 72.8 | $1000/0\bar{1}00/00\bar{1}0$ |
| 32 *oP* | 152.0 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |
| 40 *oC* | 413.0 | 63.5 | 196.4 | 62.1 | 84.5 | 107.2 | 108.9 | $0\bar{1}00/0120/\bar{1}000$ |
| 35 *mP* | 151.8 | 63.5 | 62.1 | 92.9 | 90.1 | 90.0 | 107.2 | $0\bar{1}00/\bar{1}000/00\bar{1}0$ |
| 36 *oC* | 400.3 | 62.1 | 195.9 | 63.5 | 84.6 | 107.2 | 108.4 | $1000/\bar{1}0\bar{2}0/0100$ |
| 33 *mP* | 151.2 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |
| 38 *oC* | 100.1 | 62.1 | 123.9 | 92.9 | 90.0 | 90.1 | 101.5 | $\bar{1}000/1200/00\bar{1}0$ |
| 34 *mP* | 1.0 | 62.1 | 92.9 | 63.5 | 90.0 | 107.2 | 90.1 | $\bar{1}000/00\bar{1}0/0\bar{1}00$ |
| 42 *oI* | 661.3 | 62.1 | 63.5 | 200.2 | 103.0 | 102.4 | 107.2 | $\bar{1}000/0\bar{1}00/1120$ |
| 41 *mC* | 412.2 | 196.4 | 63.5 | 62.1 | 107.2 | 95.5 | 71.1 | $0\bar{1}\bar{2}0/0\bar{1}00/\bar{1}000$ |
| 37 *mC* | 400.1 | 195.9 | 62.1 | 63.5 | 107.2 | 95.4 | 71.6 | 1020/1000/0100 |
| 39 *mC* | 99.9 | 123.9 | 62.1 | 92.9 | 90.1 | 90.0 | 78.5 | $\bar{1}\bar{2}00/\bar{1}000/00\bar{1}0$ |
| 43 *mI* | 999.0 | 74.6 | 200.2 | 63.5 | 103.0 | 127.3 | 68.2 | $1100/1120/0\bar{1}00$ |
| 44 *aP* | 0.0 | 62.1 | 63.5 | 92.9 | 90.0 | 90.1 | 107.2 | 1000/0100/0010 |

compatible with the observed diffraction pattern. The highest lattice symmetry is orthorhombic (character 13, Bravais type *oC*), which limits the possible space groups for protein crystals to either $C222_1$ or $C222$. Processing of all films in the data set was completed in space group $P1$ using the cell constants shown for lattice character 44. To test whether the crystal has space-group symmetry $C222$ and conventional cell constants $a = 74.6, b = 101.1, c = 92.9$ Å, the final steps of data processing were repeated after reindexing the reflections by the transformation $h' = 1 \cdot h + 1 \cdot k + 0 \cdot l + 0$, $k' = -1 \cdot h + 1 \cdot k + 0 \cdot l + 0$, $l' = 0 \cdot h + 0 \cdot k + 1 \cdot l + 0$ as specified for lattice character 13. Note that the transformation also provides a simple tool for correcting the indices if all reflections are misindexed by a constant. The results clearly show that the crystal has space-group symmetry $C222_1$. The presence of the $2_1$ axis was deduced from the rather weak intensities observed for reflections of type $00l' = \text{odd}$.



references