

Chapter 3

Crystal Pathologies

Todd O. Yeates

Abstract Truly ideal crystals are rarely realized in macromolecular crystallography. The conformational complexity of protein molecules and the promiscuity of their chance interactions often conspire to give crystals in which the molecules are present in alternative configurations. When the alternative configurations occur randomly throughout the crystal, one is faced by a case of static disorder (often indistinguishable from thermal motion), leading to limited resolution and potential challenges in modeling the underlying structural variations. Despite those challenges, the case of random disorder is arguably the simplest to understand and interpret. A variety of more complex categories of crystal disorder occur when alternative molecular configurations, orientations, or positions are not random, but correlated to each other in one way or another throughout the crystal specimen.

Keywords Twinning • Crystal disorder • Intensity statistics

3.1 Twinning

Twinning describes a broad set of situations where a crystal specimen is composed of multiple domains, which individually behave like ideal crystals, but which are oriented differently relative to each other (Fig. 3.1). The subject of twinning in macromolecular crystals has been well-reviewed [3, 8, 11, 24, 25, 27, 28]. The misorientation of distinct crystal domains in a twinned specimen is made possible (or even probable) by the ability of molecular spacings to be matched at the interfaces between differently oriented twin domains. A fairly non-specific type of twinning, referred to as non-merohedral twinning, gives rise in a diffraction

T.O. Yeates (✉)

Department of Chemistry and Biochemistry, University of California,
Los Angeles, CA 90095, USA
e-mail: yeates@mbi.ucla.edu

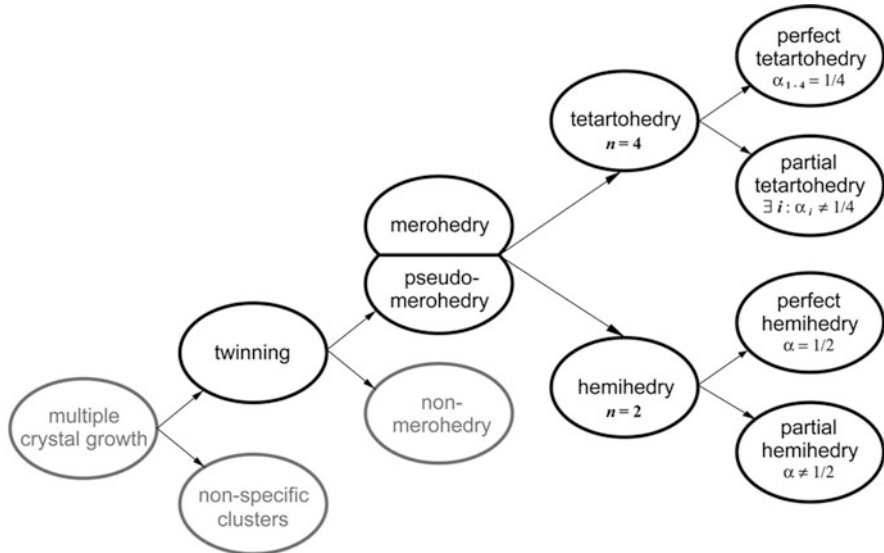


Fig. 3.1 Hierarchy of various types of twinning (Adapted from Yeates and Tsai [28])

experiment to a pattern composed of two (or more) independent, interpenetrating reciprocal lattices. This is usually easily recognized. Continual improvements in software have made it possible to deal effectively with diffraction patterns of this type, by integrating spots from distinct lattices separately, accounting for perfectly or closely overlapping reflections, etc.

Merohedral twinning is a more interesting, or at least more insidious, phenomenon. Here, the different twin domains of the specimen occur in (typically) two different orientations, related by an operation that is obeyed by the symmetry of the lattice (i.e. the holohedry) but which is not part of the crystal space group symmetry. This is possible whenever the lattice symmetry is higher than the space group symmetry. Figure 3.2 illustrates the case of space group P4; the underlying tetragonal lattice has extra rotational symmetry (422) not obeyed by the space group; the alternate twin domains are related by this extra operation. With merohedral twinning, the separate diffraction patterns arising from the multiple distinct twin domains are exactly superimposed, giving no visual indication that things are amiss.

The chief consequence of merohedral twinning is that the measured intensities are not really the true crystallographic intensities of individual reflections. Instead, each is a weighted sum of two twin-related but crystallographically independent reflections, $I(h_1)$ and $I(h_2)$, according to the value of the twin fraction, α

$$I_{\text{obs}, 1} = (1 - \alpha) I(h_1) + \alpha I(h_2)$$

$$I_{\text{obs}, 2} = \alpha I(h_1) + (1 - \alpha) I(h_2)$$

One challenge in recognizing and dealing with merohedral twinning is that the problem can manifest itself in different ways, depending on the twin fraction. When

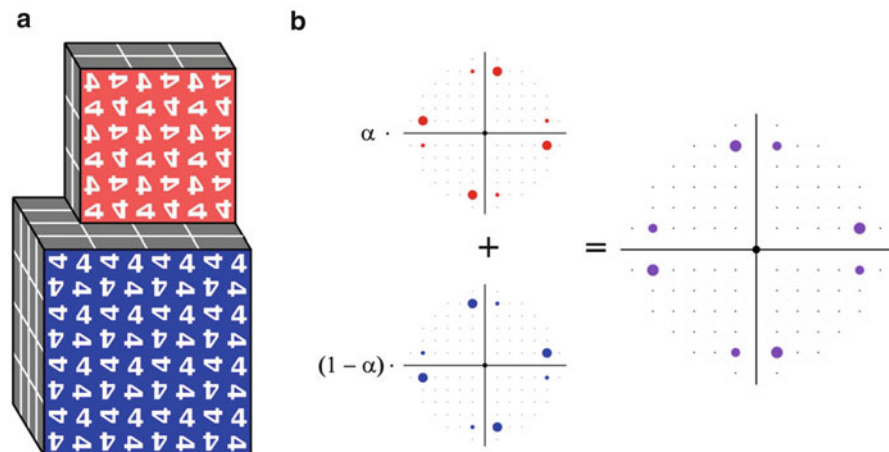


Fig. 3.2 A cartoon depicting partial merohedral twinning in space group P4. (a) Two twin domains growing together, related by a twofold twin-operation perpendicular to the fourfold symmetry axis. (b) The diffraction patterns of individual domains and their overlapping combination expected in a diffraction experiment (Adapted from Yeates and Fam [25])

α is equal to or very nearly equal to $1/2$ (a situation often referred to as ‘perfect twinning’), the outcome is erroneously high symmetry in the recorded X-ray data. For a successful structure determination, the crystallographer must come to realize that the true crystal space group symmetry is lower than it seems. When $\alpha < 1/2$ (‘partial twinning’), the observed symmetry is correct, but one must realize and deal with the fact that the observed intensities do not reflect correct crystallographic quantities. Understanding which of these two scenarios is at play is essential in arriving at a correct interpretation in the end.

3.1.1 Testing for Twinning

It is possible to delineate two distinct effects that twinning has on intensity data. Each effect gives rise to various statistical tests for twinning. These are now broadly implemented in macromolecular software packages. Here again, understanding the distinction between the different effects of twinning and their respective tests, and how they relate to the dichotomy between perfect and partial twinning, is critical for a proper analysis.

First, twinning causes twin-related reflection pairs, which should be crystallographically independent, to have intensities more similar to each other than expected by chance; in the extreme case of $\alpha = 1/2$, they are exactly equal. The magnitude of the effect depends on α , and statistical tests based on a comparison of twin-pairs (sometimes referred to as tests for partial twinning) typically return an estimate for α . A variety of useful comparison metrics have been developed over the years

[5, 15, 23]. One of these gives an easy to remember interpretation. If H is defined to be the difference between two twin-related observed intensities divided by their sum, then the mean value of $|H|$ over the data set should be equal to $(1-2\alpha)/2$ [23]. Rearrangement gives a quick estimate for α as $1/2 - \langle |H| \rangle$. Methods for treating errors in estimating α have been developed [4, 5, 9, 15].

Two points of caution are called for in estimating the twin fraction by comparing potential twin-pairs. First, non-crystallographic symmetry (NCS) can cause the same effect as partial twinning – i.e. similarity between potentially twin-related reflections – so tests of this type, without further scrutiny, can lead to false conclusions of twinning. Second, tests of this type are of no utility in situations of perfect or near-perfect twinning; the equivalence between twin-related reflections would already be implicit from the apparent higher symmetry obtained during data reduction. And reducing data in a lower symmetry and then performing a test for partial twinning (i.e. comparing potentially twin-related reflections) can only lead to confusion; such a comparison would necessarily report near-equivalence of potential twin-pairs, which is consistent with perfect twinning, though no twinning may be present.

Various statistical measures are commonly employed to examine overall intensity distributions for evidence of perfect or very-high twinning [14, 18]. Perhaps the easiest to remember is $\langle I^2 \rangle / (\langle I \rangle^2)$, which should be 2.0 for untwinned (acentric) data, and 1.5 for perfectly twinned (acentric) data. A more recent approach was designed to try to circumvent the obfuscating effects of anisotropy, whose presence along with other phenomena such as pseudo-translational symmetry can shift distributions in a way that masks the presence of twinning. In this more modern variation, the overall intensity distribution is not evaluated over individual reflections, but instead for reflection pairs nearby in reciprocal space (but not related by a potential twin operation). The local difference, $L = (I_A - I_B)/(I_A + I_B)$ obeys a simple distribution and has a simple expected mean value: $\langle |L| \rangle = 1/2$ for untwinned (acentric) data and $3/8$ for perfectly twinned (acentric) data [10]. This local test is generally more robust than the traditional approaches that date back to Wilson [22] (Fig. 3.3).

3.1.2 NCS

Non-crystallographic symmetry can confound attempts to analyze diffraction data for twinning. This situation is worsened by the observation that NCS very often occurs as an underlying feature in cases of twinning, typically with an NCS operator nearly parallel to a twin-operator. Although dissecting the two effects can be problematic, a generally useful approach can be to examine the behavior of various tests as a function of resolution; the effects of NCS typically break at higher resolution, whereas the effects of twinning persist across all resolution ranges.

Tests that give resolution-dependent results can be illuminating. For example, if an initial test for partial twinning (by comparing twin-pairs) suggests that twinning may be present, but repeating the test using only higher resolution data shows

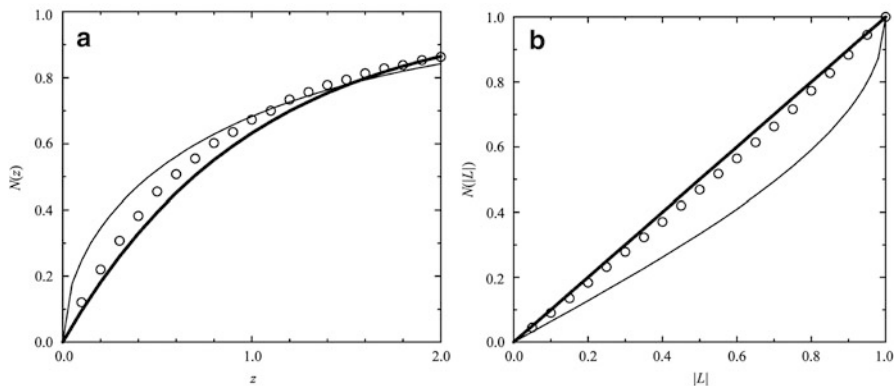


Fig. 3.3 Robustness of local intensity difference statistics in the presence of anisotropic scattering. Theoretical distributions for acentric data are shown by *bold curves*, while those for centric data are shown by the *thinner curves*. Distributions for observed acentric data are shown by *open circles*. The example (PDB code 1awu) illustrates a case where anisotropic scattering partially obscures the presence of twinning based on a traditional intensity distribution test (**a**), while the test of local differences, L , gives a clear indication of twinning, as seen in (**b**) (Adapted from Padilla and Yeates [10])

strongly reduced evidence for twinning, one might suspect that the situation results from simple NCS falsely mimicking twinning at lower resolution. Alternatively, one might examine the overall intensity statistics (in a test for perfect twinning) and find weak or ambiguous evidence for twinning at low to moderate resolution, but much stronger shifts in the intensity distribution based only on higher resolution data. This is consistent with true twinning nearly coincident with a nearly crystallographic NCS operation. At low resolution, the twin operator would mix together reflections whose intensities are already nearly equal to each other because of the NCS. Therefore, the intensity distribution might be almost normal at low resolution, and show strong evidence for twinning only at resolutions where the nearly crystallographic nature of the NCS breaks down.

3.1.3 The End Game for Twinning

If twinning is properly recognized, chances for successful structure determination are often good, especially by molecular replacement. Assuming that the true crystal symmetry has been correctly assigned, modern programs provide robust routines for refining structures against twinned data. How is this possible, given that the true intensities were never measured? In one type of approach, it is sometimes possible to effectively correct the observed data and estimate what the observed intensities should have been in the absence of twinning; this is referred to as ‘detwinning’. In contrast, most approaches to structure determination and refinement take the reverse strategy, modeling the effects of twinning into the calculated intensities instead.

As a precautionary note, it should be understood that the averaging effects of twinning tend to produce lower R-values for purely statistical reasons not reflective of model quality. Therefore, obtaining a lower R-value in atomic refinement when twinning is invoked is, by itself, not evidence for the presence of twinning. More careful analyses of the type described above (and in more thorough reviews) are essential.

3.1.4 Other Variations on Twinning

The twinning situations noted above cover only the simpler types; there are numerous more complicated scenarios (Fig. 3.1). Pseudo-merohedral twinning can occur in space groups where twinning is not ordinarily expected, if a fortuitous unit cell geometry causes a low symmetry lattice to have nearly higher rotational symmetry. A rare situation known as reticular merohedral twinning can occur when only a subset of the reflections superimpose on each other; this can lead to strange diffraction patterns. Finally, twinning of higher order – i.e. with more than two distinct domain orientations – is possible. Several macromolecular cases of tetartohedral twinning ($n = 4$) have been reported in recent years [1, 6, 16, 29], and equations for handling such cases have been introduced [26].

3.2 Other Disorder Pathologies

Other kinds of disorder, distinct from twinning, have been reported in macromolecular crystals. One broad category includes cases where a single molecular configuration is maintained in a crystallographically ordered fashion in one layer (or row), but successive layers (or rows) might contain the molecule in an alternative configuration. When this occurs stochastically from layer to layer, the result has been described as an order-disorder (OD) phenomenon. Although cases are known where the distinction between alternate molecular configurations is a difference in orientation [12], most cases occur as a difference in relative position between molecules in different layers or rows [7, 17, 19–21, 30]. These cases are often described as lattice translocation disorders (LTD); their discovery dates to the case of methemoglobin in 1954, before the first crystal structures of proteins were determined [2].

LTD and other OD cases do not typically give the kinds of intensity distribution shifts seen in twinning; the short length scale of the stochastic variations between molecules causes structure factors to sum by interference in the usual way; complex F's add rather than intensities. LTD is therefore diagnosed in different ways. During the middle stages of structure determination, the presence of interpenetrating molecular density may provide a clue, echoing the presence of unmodeled molecular configurations that cannot exist simultaneously. Warning signs can often be seen before this stage. Intensity statistics can be 'hyper-centric' (shifted opposite from

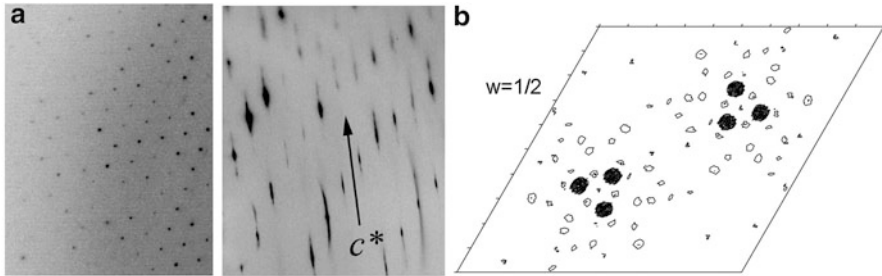


Fig. 3.4 A lattice translocation disorder in crystals of a bacterial microcompartment shell protein. (a) Prominent streaking is observed in certain directions, along c^* in this case. (b) Impossibly close packing peaks in a native Patterson map; the indicated molecular positions are not all simultaneously possible (Adapted from Tsai et al. [19])

the case of twinning) because of the modulating effects of translationally related molecules. However, hyper-centric intensity distributions are fairly common (e.g. whenever pseudo translational symmetry is present) even in crystals that do not suffer from disorder. Two features that appear to be common in LTD cases are systematic streaking of a subset of the reflections in a defined direction (Fig. 3.4a), and strong packing peaks in native Patterson maps at positions so close to the origin or to each other that they would imply impossibly close molecular packing if all the molecular positions were simultaneously occupied (Fig. 3.4b).

As with twinning, if a suitable model of the OD/LTD disorder can be developed, the structure can be determined correctly. As before, in some cases the observed intensities can be corrected by undoing the effects of having multiply shifted, partially occupied molecular positions. Alternatively, refinement can be performed in a way that incorporates the disorder into the model.

Beyond these kinds of disorders, others even more complex have been noted. Systematic off-Bragg peaks in crystals of a profilin-actin complex have been interpreted in terms of a complex modulated filament, whose period may not necessarily be commensurate with the lattice spacings in the crystal [13].

3.3 Concluding Remarks

Despite progress in identifying and dealing with disorder in macromolecular crystals, one thing that remains relatively clear is that the treatments employed are only approximations for what must be occurring in real crystals. The failures of final models to fully capture reality are especially evident in the cases of LTD treated so far. The models give reasonable approximations to the Bragg scattering (i.e. satisfactory R-values), but they do not account for the substantial scattering observed as streaking around Bragg peaks. To better understand and treat problems of disorder, renewed efforts are needed in the area of modeling diffraction from non-crystalline materials.

References

1. Barends TRM, de Jong RM, van Straaten KE, Thunnissen AWH, Dijkstra BW (2005) *Escherichia coli* MltA: MAD phasing and refinement of a tetartohedrally twinned protein crystal structure. *Acta Crystallogr D* 61:613–621
2. Bragg WL, Howells ER (1954) X-ray diffraction by imidazole methaemoglobin. *Acta Crystallogr* 7:409–411
3. Dauter Z (2003) Twinned crystals and anomalous phasing. *Acta Crystallogr D* 59:2004–2016
4. Dumas P, Ennifar E, Walter P (1999) Detection and treatment of twinning: an improvement and new results. *Acta Crystallogr D* 55:1179–1187
5. Fisher RG, Sweet RM (1980) Treatment of diffraction data from protein crystals twinned by merohedry. *Acta Crystallogr A* 36:755–760
6. Gayathri P, Banerjee M, Vijayalakshmi A, Azeez S, Balaram H, Balaram P, Murthy MRN (2007) Structure of triphosphate isomerase (TIM) from *Methanocaldococcus jannaschii*. *Acta Crystallogr D* 63:206–220
7. Hare S, Cherepanov P, Wang J (2009) Application of general formulas for the correction of a lattice-translocation defect in crystals of a lentiviral integrase in complex with LEDGF. *Acta Crystallogr D* 65:966–973
8. Helliwell JR (2008) Macromolecular crystal twinning, lattice disorders and multiple crystals. *Crystallogr Rev* 14:189–250
9. Lunin VY, Lunina NL, Baumstark MW (2007) Estimates of the twinning fraction for macromolecular crystals using statistical models accounting for experimental errors. *Acta Crystallogr D* 63:1129–1138
10. Padilla JE, Yeates TO (2003) A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr D* 59:1124–1130
11. Parsons S (2003) Introduction to twinning. *Acta Crystallogr D* 59:1995–2003
12. Pletnev S, Morozova KS, Verkhusha VV, Dauter Z (2009) Rotational order-disorder structure of fluorescent protein FP480. *Acta Crystallogr D* 65:906–912
13. Porta J, Lovelace JJ, Schreurs AM, Kroon-Batenburg LM, Borgstahl GE (2011) Processing incommensurately modulated protein diffraction data with Eval15. *Acta Crystallogr D* 67:628–638
14. Rees DC (1980) The influence of twinning by merohedry on intensity statistics. *Acta Crystallogr A* 36:578–581
15. Rees DC (1982) A general theory of x-ray intensity statistics for twins by merohedry. *Acta Crystallogr A* 38:201–207
16. Rosendal KR, Sinning I, Wild K (2004) Crystallization of the crenarchaeal SRP core. *Acta Crystallogr D* 60:140–143
17. Roversi P, Blanc E, Johnson S, Lea SM (2012) Tetartohedral twinning could happen to you. *Acta Crystallogr D* 68:418–424
18. Stanley E (1972) The identification of twins from intensity statistics. *J Appl Crystallogr* 5:191–194
19. Tsai Y, Sawaya MR, Yeates TO (2009) Analysis of lattice-translocation disorder in the layered hexagonal structure of carboxysome shell protein CsoS1C. *Acta Crystallogr D* 65:980–988
20. Wang J, Kamtekar S, Berman AJ, Steitz TA (2005a) Correction of x-ray intensities from single crystals containing lattice-translocation defects. *Acta Crystallogr D* 61:67–74
21. Wang J, Rho S, Park HH, Eom SH (2005b) Correction of x-ray intensities from an HslV-HslU co-crystal containing lattice-translocation defects. *Acta Crystallogr D* 61:932–941
22. Wilson AJC (1949) The probability distribution of x-ray intensities. *Acta Crystallogr* 2:318–321
23. Yeates TO (1988) Simple statistics for intensity data from twinned specimens. *Acta Crystallogr A* 44:142–144
24. Yeates TO (1997) Detecting and overcoming crystal twinning. *Methods Enzymol* 276:344–358
25. Yeates TO, Fam BC (1999) Protein crystals and their evil twins. *Structure* 7:R25–R29

26. Yeates TO, Yu F (2008) Equations for determining tetartohedral twin fractions. *Acta Crystallogr D* 64:1158–1164
27. Yeates TO, Sawaya MR (2011) Structure determination in the presence of twinning by merohedry. In: Arnold E, Himmel DM, Rossmann MG (eds) *International tables for crystallography*. Vol. F. Oxford, Wiley-Blackwell, pp 548–551
28. Yeates TO, Tsai Y (2011) Detecting twinning by merohedry. In: Arnold E, Himmel DM, Rossmann MG (eds) *International tables for crystallography*, vol F. Wiley-Blackwell, Oxford, pp 311–316
29. Yu F, Song A, Xu C, Sun L, Li J, Tang L, Yu M, Yeates TO, Hu H, He J (2009) Determining the DUF55-domain structure of human thymocyte nuclear protein 1 from crystals partially twinned by tetartohedry. *Acta Crystallogr D* 65:212–219
30. Zhu X, Xu X, Wilson IA (2008) Structure determination of the 1918 H1N1 neuraminidase from a crystal with lattice-translocation defects. *Acta Crystallogr D* 64:843–850