Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A noncrystallographic screw axis parallel to a twin axis can corrupt intensity statistics

Vladimir Y. Lunin

Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russian Federation

Correspondence e-mail: lunin@impb.psn.ru

Received 9 January 2012
Accepted 15 March 2012

A number of methods to detect twinning are based upon the assumption that the statistical properties of diffracted intensities are different for twinned and untwinned specimens. This may not be true for a large portion of the reflections in a twinned specimen if a noncrystallographic screw axis parallel to the twinning axis is present. In this case, up to half of all reflections can obey Wilson statistics, which are typical of untwinned crystals. The distribution corresponding to a whole set of observed intensities is biased towards the Wilson distribution in this case.

1. Introduction

Twinning is a well recognized phenomenon in macromolecular crystallography (see, for example, Yeates, 1997; Yeates & Fam, 1999; Dauter, 2003; Parsons, 2003; Lebedev *et al.*, 2006; Zwart *et al.*, 2008). A hemihedrally twinned specimen is composed of two crystalline domains related by a twinning operation (180° rotation) which belongs to the point group of the lattice but not to the point group of the crystal. The observed (twinned) diffraction intensities are in this case given by linear combinations of the true untwinned intensities of reflections weighted by the relative volumes (twinning fractions) of the two domains. Several methods that have been suggested for the detection of twinning are based upon differences in the statistical properties of twinned and untwinned intensities (Rees, 1980; Yeates, 1988; Lunin *et al.*, 2007).

Generally, the normalized untwinned intensities z for acentric reflections may be considered as independent random variables that obey Wilson statistics (Wilson, 1949) with an exponential probability density function,

$$P_{\rm W}(z) = \exp(-z). \quad (1)$$

The twinned intensity is a weighted sum of two Wilson-type variables,

$$z_{\rm twin} = (1-\alpha)z_1 + \alpha z_2, \quad (2)$$

and has a different distribution from the Wilson distribution (Stanley, 1955, 1972; Rees, 1982). In particular, in the case of perfect twinning (α = 1/2) the probability density function is

$$P_{\rm ptwin}(z) = 4z\exp(-2z). \quad (3)$$

The corresponding cumulative distribution functions (c.d.f.s) $N_{\rm W}(z)$ and $N_{\rm ptwin}(z)$ reveal different shapes of their graphs. The function $N_{\rm W}(z)$ has a positive first derivative over z and a negative curvature at z = 0, while the first derivative of $N_{\rm ptwin}(z)$ equals 0 and the curvature is positive at z = 0, resulting in a characteristic 'sigmoidal' shape of the graph (Stanley, 1955; Rees, 1980; Dauter, 2003). Comparison of the c.d.f. $N_{\rm obs}(z)$ calculated from experimentally observed intensities with the two standard distributions $N_{\rm W}(z)$ and $N_{\rm ptwin}(z)$ is the usual tool that is used to make a decision on the presence or absence of twinning. A number of more advanced tools are beyond the scope of this paper.

Statistical procedures usually give reasonable results, but anomalies have been detected in several difficult cases. Even for a specimen twinned with a high twinning fraction the observed $N_{\rm obs}(z)$ distribution is sometimes more close to the Wilson distribution $N_{\rm W}(z)$ than to $N_{\rm ptwin}(z)$ (Abrescia & Subirana, 2002; Padilla & Yeates, 2003) or even almost coincides with it (Dauter *et al.*, 2005). Such anomalies are often detected in cases where noncrystallographic symmetry is present (Lebedev *et al.*, 2006; Zwart *et al.*, 2008). In particular, it is well recognized that the presence of pseudo-translational symmetry can change the intensity distributions significantly. This phenomenon is explained by a large fraction of very weak reflections resulting from pseudo-centring. This can mask the effect of the decrease in the number of small intensities expected for twinned specimens. Similarly, deviation of the observed $N_{\rm obs}(z)$ distribution from the theoretically expected $N_{\rm twin}(z)$ can be detected for cases of rotational pseudo-symmetry (Lee *et al.*, 2003; Lebedev *et al.*, 2006). To some extent, the difficulties caused by pseudo-translation may be overcome by the use of local L statistics (Padilla & Yeates, 2003).

In this paper, we discuss one further mechanism by which noncrystallographic symmetry can mask twinning. We show that for a large number of reflections twinned intensities can obey Wilson statistics if the studied structure has a noncrystallographic screw axis parallel to the twinning axis. As a result, the distribution of the whole set of intensities may be significantly biased towards the Wilson distribution.

2. Anomalies arising from noncrystallographic screw symmetry

2.1. Test example

Let us first consider an idealized case where

(i) the space group is $P1$;

(ii) the unit cell is monoclinic (with unique axis **b**), *i.e.* the point group of the lattice is $2/m$, and axis **b** is a potential twinning axis;

(iii) the unit cell contains two molecules related by a noncrystallographic screw axis which is a 180° rotation around axis **b** and a shift of **t** = **b**/4 along this axis.

The general formula that relates the structure factors calculated from two copies of a molecule separately is

$$\mathbf{F}_2(\mathbf{s}) = \mathbf{F}_1(\mathbf{G}^T\mathbf{s})\exp[2\pi i(\mathbf{s},\mathbf{t})]. \qquad (4)$$

Here, $\mathbf{F}_1(\mathbf{s}) = F_1(\mathbf{s})\exp[i\varphi_1(\mathbf{s})]$ and $\mathbf{F}_2(\mathbf{s})$ are the complex structure factors calculated for reflection **s** from the coordinates of the first and second molecules separately; the coordinates of two molecules are related by

$$\mathbf{r}_{2,j} = \mathbf{G}\mathbf{r}_{1,j} + \mathbf{t}, \qquad (5)$$

and $\mathbf{G}^T$ represents the transposed rotation matrix **G**. In the considered example, this equation takes the form

$$\mathbf{F}_2(hkl) = \mathbf{F}_1(\bar{h}k\bar{l})\exp\left(i\frac{\pi}{2}k\right), \qquad (6)$$

where hkl are Miller indexes. The structure factors corresponding to the whole structure are in this case

$$\mathbf{F}(hkl) = \mathbf{F}_1(hkl) + \mathbf{F}_2(hkl) = \mathbf{F}_1(hkl) + \mathbf{F}_1(\bar{h}k\bar{l})\exp\left(i\frac{\pi}{2}k\right) \qquad (7)$$

and

$$\mathbf{F}(\bar{h}k\bar{l}) = \mathbf{F}_1(\bar{h}k\bar{l}) + \mathbf{F}_1(hkl)\exp\left(i\frac{\pi}{2}k\right) = \left[\mathbf{F}_1(hkl) + \mathbf{F}_1(\bar{h}k\bar{l})\exp\left(-i\frac{\pi}{2}k\right)\right]\exp\left(i\frac{\pi}{2}k\right). \qquad (8)$$

2.2. Obstacles to determination of the space group

If reflection index k is odd, *i.e.* $k = 2m + 1$, where m is an integer, then the magnitudes $F(hkl)$ and $F(\bar{h}k\bar{l})$ are generally different,

$$F(hkl) = |\mathbf{F}_1(hkl) + i(-1)^m\mathbf{F}_1(\bar{h}k\bar{l})|, \qquad (9)$$

$$F(\bar{h}k\bar{l}) = |\mathbf{F}_1(hkl) - i(-1)^m\mathbf{F}_1(\bar{h}k\bar{l})|, \qquad (10)$$

but the two magnitudes are equal if k is even,

$$F(hkl) = F(\bar{h}k\bar{l}) = |\mathbf{F}_1(hkl) + (-1)^m\mathbf{F}_1(\bar{h}k\bar{l})| \quad \text{if } k = 2m. \qquad (11)$$

It follows from (11) that in the considered example half of the reflections indicate symmetry typical of space groups $P2$ or $P2_1$, while the crystal structure does not possess this symmetry. The partial reciprocal-space symmetry (11) can artificially lower the $R_{\rm sym}$ value and provoke the choice of a higher symmetry space group than the crystal really has.

2.3. Extinctions

Similarly to the crystallographic screw axis, a noncrystallographic screw axis may cause extinctions for some reflections. For example, it follows from (7) that

$$F(0k0) = 0 \quad \text{if } k = 4n + 2, n \text{ is an integer}. \qquad (12)$$

Nevertheless, in contrast to cases in which a pseudo-translation is present, the number of such reflections is relatively small and is hardly ever able to influence statistics significantly.

2.4. Statistics for twinned intensities

Let it now be supposed that in addition to the conditions formulated in §2.1 the studied specimen is perfectly hemihedrally twinned with respect to axis **b**, so that the observed (twinned) intensities are

$$I_{\rm ptwin}(hkl) = \frac{1}{2}[I(hkl) + I(\bar{h}k\bar{l})] = \frac{1}{2}[F^2(hkl) + F^2(\bar{h}k\bar{l})]. \qquad (13)$$

Here,

$$I(hkl) = F^2(hkl) \qquad (14)$$

are untwinned intensities that obey Wilson statistics. It follows from (11) that

$$I_{\rm ptwin}(hkl) = I(hkl) \quad \text{if } k = 2m. \qquad (15)$$

This last equation shows that in the considered example the twinned intensities are equal to untwinned intensities (and so obey the Wilson distribution) for half of all reflections. The cumulative distribution function $N_{\rm obs}(z)$ calculated for the whole set of twinned intensities is biased towards the Wilson distribution in this case.

2.5. *L* test

The L test suggested by Padilla & Yeates (2003) is designed to distinguish between two cases: all intensities are distributed with the Wilson distribution (1) (but possibly with different values of the local mean) or all intensities are distributed with distribution (3) that corresponds to perfect twinning. In the considered example, the intensities of half of all reflections are distributed with one distribution, while the other half have the second. As a result, one can expect that the distribution of the observed L statistics and its moments lie between theoretical values corresponding to untwinned and perfectly twinned specimens.

2.6. General case of screw symmetry translation

In contrast to the crystallographic screw axis, the translational part of a noncrystallographic screw axis is not constrained by the crystal lattice to discrete values and may be arbitrary. If in the conditions of

§2.1 the translational part of the noncrystallographic screw axis has relative coordinates (0, t, 0) then (6) should be replaced by

$$\mathbf{F}_2(hkl) = \mathbf{F}_1(\bar{h}k\bar{l})\exp[i\pi k(2t)] \qquad (16)$$

and condition (15) becomes

$$I_{\rm ptwin}(hkl) = I(hkl) \quad \text{if } 2tk \text{ is an integer.}$$

Depending on the value of t, the last condition may be satisfied for a smaller fraction of reflections than in the case of $t = 1/4$ or may never be satisfied. For example, in the case of $t = 1/6$ the subset of reflections with $k = 3m$ has a Wilson intensity distribution, *i.e.* one third of the total number of reflections. If t is an irrational number then formally the condition in (17) is never satisfied. Nevertheless, even in such a case it is possible to have a $2tk$ value that is close to an integer for some values of the index k. The twinned intensities will be close to untwinned intensities for such reflections and the corresponding statistical distribution may have some bias towards the Wilson distribution.

3. Conclusion

A noncrystallographic screw axis parallel to a potential twinning axis can corrupt intensity statistics and cause difficulties in determination of the crystal space group and in detection of twinning.

The work was supported by RFBR grant 10-04-00254-a.

References

Abrescia, N. G. A. & Subirana, J. A. (2002). *Acta Cryst.* D**58**, 2205–2208.
Dauter, Z. (2003). *Acta Cryst.* D**59**, 2004–2016.
Dauter, Z., Botos, I., LaRonde-LeBlanc, N. & Wlodawer, A. (2005). *Acta Cryst.* D**61**, 967–975.
Lebedev, A. A., Vagin, A. A. & Murshudov, G. N. (2006). *Acta Cryst.* D**62**, 83–95.
Lee, S., Sawaya, M. R. & Eisenberg, D. (2003). *Acta Cryst.* D**59**, 2191–2199.
Lunin, V. Y., Lunina, N. L. & Baumstark, M. W. (2007). *Acta Cryst.* D**63**, 1129–1138.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Parsons, S. (2003). *Acta Cryst.* D**59**, 1995–2003.
Rees, D. C. (1980). *Acta Cryst.* A**36**, 578–581.
Rees, D. C. (1982). *Acta Cryst.* A**38**, 201–207.
Stanley, E. (1955). *Acta Cryst.* **8**, 351–352.
Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Yeates, T. O. (1988). *Acta Cryst.* A**44**, 142–144.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.
Yeates, T. O. & Fam, B. C. (1999). *Structure*, **7**, R25–R29.
Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N. & Adams, P. D. (2008). *Acta Cryst.* D**64**, 99–107.