Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Correction of X-ray intensities from single crystals containing lattice-translocation defects

Jimin Wang,[a]* Satwik Kamtekar,[a] Andrea J. Berman[a] and Thomas A. Steitz[a,b,c]

[a]Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA, [b]Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA, and [c]Howard Hughes Medical Institute, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

Correspondence e-mail: wang@mail.csb.yale.edu

Received 10 June 2004
Accepted 21 October 2004

In 1954, Howells and colleagues described an unusual diffraction pattern from imidazole methemoglobin crystals caused by lattice-translocation defects. In these crystals, two identical lattices coexist as a single coherent mosaic block, but are translated by a fixed vector with respect to each other. The observed structure is a weighted sum of the two identical but translated structures, one from each lattice; the observed structure factors are a weighted vector sum of the two structure factors with identical unit amplitudes but shifted phases. A general procedure is described to obtain the unit amplitudes of observed structure factors from a realigned single lattice through an X-ray intensity correction. An application of this procedure is made to determine the crystal structure of φ29 DNA polymerase at 2.2 Å resolution using multiple isomorphous replacement and multiwavelength anomalous dispersion methods.

1. Introduction

In 1954, Howells and colleagues obtained an unusual X-ray diffraction pattern from crystals of imidazole methemoglobin (Bragg & Howells, 1954). The normally monoclinic methemoglobin crystal form had been converted to a new form that they called 'statistically orthorhombic'. Howells and colleagues proposed a packing model that accounted for this transformation. They suggested that the packing of molecules within planes perpendicular to the *c* axis was identical in both crystal forms. However, in the statistically orthorhombic form successive layers were displaced relative to each other. The magnitude of the displacement was a definite amount, but its direction varied stochastically and was either parallel or antiparallel with the *a* axis. The observed structure factors were a summation through all layers, which could be mathematically divided into two groups or two lattices, one with a small displacement in the $+a$ direction and the other with the same magnitude of displacement but in the $-a$ direction. We refer to this phenomenon, in which two lattices coexist as a single coherent mosaic block, but are translated by a fixed vector with respect to each other, as lattice-translocation defects.

The lattice-translocation defects observed in imidazole methemoglobin had a dramatic effect on the reflection shapes in diffraction images. While some reflections remained sharp, others could be classified as moderately sharp, diffuse and very diffuse (Bragg & Howells, 1954). This change in shape was a consequence of the fact that successive *c* layers, although they belonged to different lattices, diffracted coherently (in marked contrast to diffraction in merohedrally twinned crystals, where coherent interactions between different lattices are minimal and there is no change in spot shapes).

© 2005 International Union of Crystallography
Printed in Denmark – all rights reserved

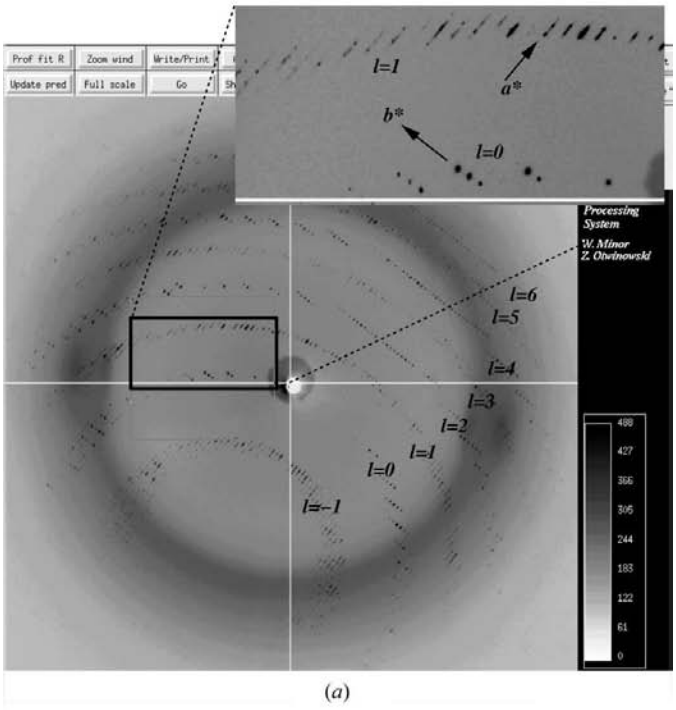

Figure 1
Observed alternating sharp and diffuse reflections as a function of the reciprocal-space index *l*. This feature was present in the nearly 300 crystals we examined in this crystal form for the structure determination. (*a*) A native crystal. (*b*) A Hg-soaked crystal that was used for multiwavelength anomalous dispersion data collection and structure determination.

Table 1
Symbols and definitions.

$\mathbf{t_d}$	Lattice-translocation defect vector
$\mathbf{t_i}$	Intermolecular (particle) vectors within one asymmetric unit
$\mathbf{F}_{\rm unit}$	Unit structure factors without lattice defects
$I_{\rm unit}$	Unit intensities without lattice defects
$\mathbf{F}_{\rm total}$	Total structure factors from two structures from two lattices
$I_{\rm total}$, $I_{\rm obs}$	Observed total intensities
κ, $1-\kappa$	Two fractions of lattice-translocation defects
M, N	No. of layers that make up a lattice-translocation defect
a, b, c, α, β, γ	Real unit-cell parameters
a^*, b^*, c^*, α^*, β^*, γ^*	Reciprocal unit-cell parameters
$\mathbf{h}$	Vector of reciprocal-space index h, k, l
l	Reciprocal-space index l
$\langle\ \rangle_l$	Layer-average measurement for a given index l
Δz	z component of $\mathbf{t_d}$ or $\mathbf{t_i}$
f	Intensity correction factor ($1/f$)
Q	The sum of squared differences between the fitted and observed functions
A, B	An average intensity and its variation
r	The ratio between A and B; $r = B/A$
v	The layer-averaged intensity ratio between indexes l = odd and even
p	The ratio between the Patterson peaks at (0, 0, 1/2) and (0, 0, 0)

Cochran & Howells (1954) were able to provide a quantitative explanation for the change in spot shapes. When the translation between two layers was associated with no phase shift or a phase shift of magnitude $2n\pi$, sharp and strong reflections were observed. However, phase shifts of other values created smeary reflections; in particular, those associated with a phase shift of $(2n + 1)\pi$ were weak and diffuse. Because the shifts between layers occurred stochastically, no single super-lattice with enlarged cell edges could encompass them.

The extent of the lattice-translocation defects and the translational relationship between the two lattices in the 'two-lattice' treatment, as described below, can be established through the native Patterson map calculation. Phase shifts between the two lattices introduce an extra strong–weak modulation of X-ray intensities that is not present in individual lattices, which in turn results in extra peaks in the native Patterson maps. The Patterson peak heights are a function of the magnitude of the modulation or the extent of the lattice-translocation defects; their locations are a function of translation vectors relating the two lattices. Additionally, the possible systematic underestimation of the smeary reflection intensities during intensity integration may further increase the Patterson peak heights. When molecular structures in individual lattices are known, the lattice-translocation defects can in principle be accommodated by using an atomic model that explicitly includes both lattices: the calculated Patterson maps can then be matched with the observed Patterson maps. However, the resulting electron-density maps would still not be interpretable owing to the overlay of the two lattices with different origins, despite a possible perfect agreement between the observed and calculated structure factors in both amplitudes and phases. Therefore, this procedure is not suitable for

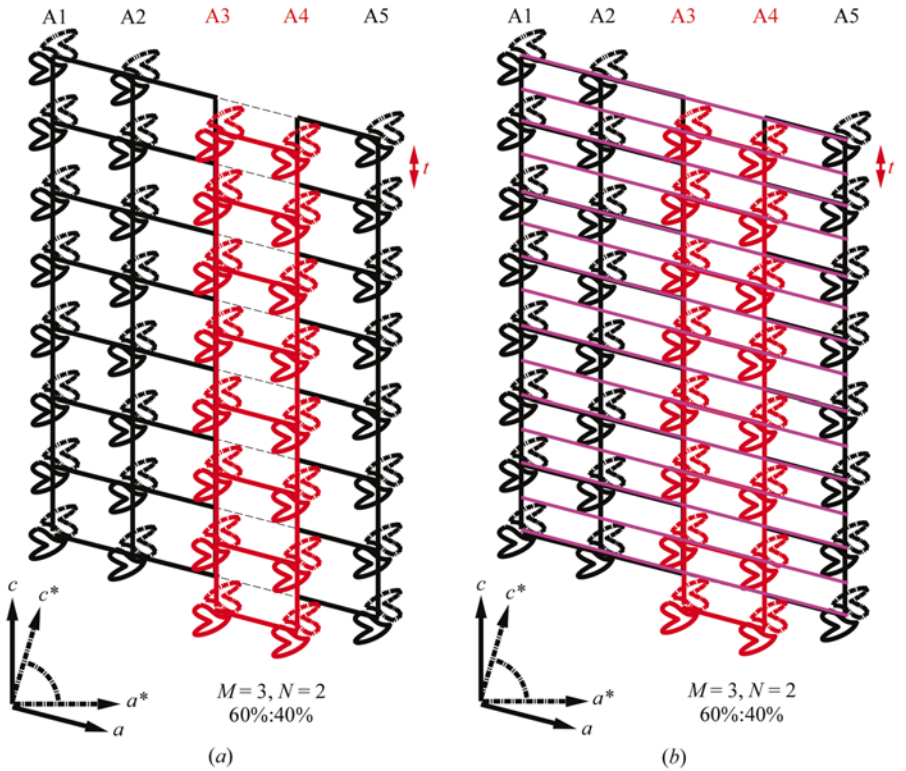

Figure 2
A lattice-translocation defect model. (*a*) The lattice has a layer structure with layers (parallel to the *bc* plane) stacked along the *a** direction. A fraction of layers (N = 2, layers A3 and A4) has a translocation offset t in the c axis with respect to the rest (M = 3, layers A1, A2 and A5). (*b*) A new 'superlattice' (shown in magenta) emerges with ac dimension reduced by half for all reflections with l = even indexes.

solving a new structure that requires map interpretation. An alternative procedure, described here, is to computationally realign the two translocated lattices into a single lattice. This procedure first establishes the vector ($\mathbf{t_d}$) relating the two lattices and the fractions (κ and $1-\kappa$) of each lattice using a native Patterson map calculation. With these parameters, a correction is made to the observed intensities to remove all modulation introduced by the lattice-translocation defects.

Bacillus phage φ29 DNA polymerase is a model enzyme for studying DNA replication initiated at the end of chromosomes (Meijer *et al.*, 2001). It possesses both 5′-3′ DNA polymerase and 3′-5′ exonuclease activities, intrinsic strand-displacement activity and a processive capacity and has the capability of catalyzing the priming reaction using the terminal protein as a substrate. It belongs to the B-family of DNA polymerases and its closest relative of known structure is RB69 DNA polymerase (Wang *et al.*, 1997).

Bacillus phage φ29 DNA polymerase was initially crystallized in a monoclinic space group ($P2_1$, unit-cell parameters a = 59.9, b = 169.4, c = 68.6 Å, β = 107.6°) with a severe lattice-translocation defect (Fig. 1). In each diffraction image, nearly half of the data within alternate zones (with $l = 2n + 1$) had smeary features with streaky axes parallel to the *a** direction. This defect severely reduced the quality of integrated X-ray diffraction data (Table 1). Prior to the completion of structure determination (Kamtekar *et al.*, 2004), over 300 crystals had been examined and fewer than five crystals had negligible lattice-translocation defects. On several occasions, 'statistically orthorhombic' lattices were also observed with the new a unit-cell parameter equal to $2a\sin\beta$, exactly like those which led to the original discovery of lattice-translocation defects (Bragg & Howells, 1954). The statistically orthorhombic (or stochastic) crystals of φ29 DNA polymerase and of methemoglobin are fundamentally different from instances in which pseudo-merohedral twinning leads to a conversion from monoclinic to orthorhombic (Rudolph *et al.*, 2004). Owing to coherent diffraction, the stochastic crystals yield diffraction spots of varying shape; the lack of coherent diffraction between mosaic blocks in pseudo-merohedral crystals precludes this effect. The φ29 DNA polymerase crystals diffracted to a resolution of 2.1 Å. There were two molecules per asymmetric unit, related by a non-crystallographic twofold axis that was perpendicular to the crystallographic dyad, with a calculated Matthews coefficient of 2.5 $\text{Å}^3\,\text{Da}^{-1}$ and a solvent content of 51%. The lattice-translocation vector was approximately (0, 0, 1/2), as derived from the native Patterson map calculation.

2. A generalized formulation for correcting lattice-translocation defects

We propose a two-lattice treatment for lattice-translocation defects in which a subset of the layers in a crystal (N layers; symbols and definitions are listed in Table 1) is offset with respect to the remaining layers (M layers) in the lattice by a translocation vector $\mathbf{t_d}$ (Fig. 2). In this treatment, all lattices are made of layers that lie parallel to the *bc* plane and are stacked along the *a** axis. X-rays scattered from the two sets of layers differ by a phase shift of $\exp(2\pi i\mathbf{ht_d})$, where $\mathbf{h}$ is the reciprocal vector and $\mathbf{t_d}$ is the translocation vector. If the structure factor for each layer is $\mathbf{F}_{\rm unit}$, the total structure factor $\mathbf{F}_{\rm total}$ is

$$\mathbf{F}_{\rm total} = [M + N\exp(2\pi i\mathbf{ht_d})]\mathbf{F}_{\rm unit}. \quad (1)$$

Expressing the relative number of M and N layers in terms of their fractions of the crystal κ and $(1-\kappa)$ and reassigning $\mathbf{F}_{\rm unit}$ as a structure factor for one unit cell, the total structure factor of one unit cell, $\mathbf{F}_{\rm total}$, becomes

$$\mathbf{F}_{\rm total} = [\kappa + (1-\kappa)\exp(2\pi i\mathbf{ht_d})]\mathbf{F}_{\rm unit}. \quad (2)$$

The corresponding intensity relationship is

$$I_{\rm total} = (2\kappa^2 - 2\kappa + 1)\left[1 + \frac{2\kappa(1-\kappa)}{2\kappa^2 - 2\kappa + 1}\cos(2\pi\mathbf{ht_d})\right]I_{\rm unit} = A[1 + r\cos(2\pi\mathbf{ht_d})]I_{\rm unit} = fI_{\rm unit}, \quad (3)$$

where $A = (2\kappa^2 - 2\kappa + 1)$, $r = 2\kappa(1-\kappa)/(2\kappa^2 - 2\kappa + 1)$ and f is the extra modulation factor that is introduced into the observed intensities by the lattice-translocation defects. After its determination, the intensity modulation can then be removed upon application of the correction $I_{\rm unit} = (1/f)I_{\rm total}$. (3) has three parameters to be determined: the translocation vector $\mathbf{t_d}$, the fraction κ and the overall scale factor, which combines with the $(2\kappa^2 - 2\kappa + 1)$ term.

The lattice-translocation vector $\mathbf{t_d}$ is determined using the Patterson function method. For example, in data set 3 (Fig. 3) from φ29 DNA polymerase crystals $\mathbf{t_d} = (0, 0, \Delta z)$ with Δz precisely equal to 1/2 and in data set 2 $\Delta z = 0.4735$ (approximately equal to 1/2). X-ray diffraction statistics for data set 3 are given in Table 2. The corresponding peak heights are 27 and 54% of the origin peak, respectively, for data sets 3 and 2. The consequences of the Patterson peak heights (and the derived severity of defect fractions, see below) may appear to be counterintuitive, because data set 3 has robust variations through all l indexes while data set 2 has large variations at both high and low l indexes but small variation at medium l indexes (Fig. 4). However, there is no simple relationship between the Patterson peak heights (p) at a general location and the intensity variation ratio (r) between zones in diffraction. When $\mathbf{t_d} = (0, 0, 1/2)$, $p = r$ as defined in (8). For data set 2, the value Δz (0.4735) becomes 1/2 when data are truncated to very low resolution and its intensity variation ratio (r; for example, between $l = 3$ and $l = 2$) is indeed much larger than for data set 3 (Fig. 4). Many other data sets have slight variations in peak heights and Δz (mostly near 1/2). Data set 1 had its first constructive diffraction power addition at the index l approximately equal to 15 with small layer-averaged intensity variations (Fig. 3). The Patterson map also showed a small peak (6% of the origin peak height, which is consistent with the small layer-averaged intensity variations) at $\Delta z = 0.07$ (which corresponds to index l approximately equal to 15). Such a small Δz vector could not be determined accurately owing to truncation-ripple effects of the origin peak in the Patterson map calculations. Data set 1 was approximated as an ideal data set without (or with minor) lattice-translocation defects. Because $\mathbf{t_d}$ has only a small Δz component with

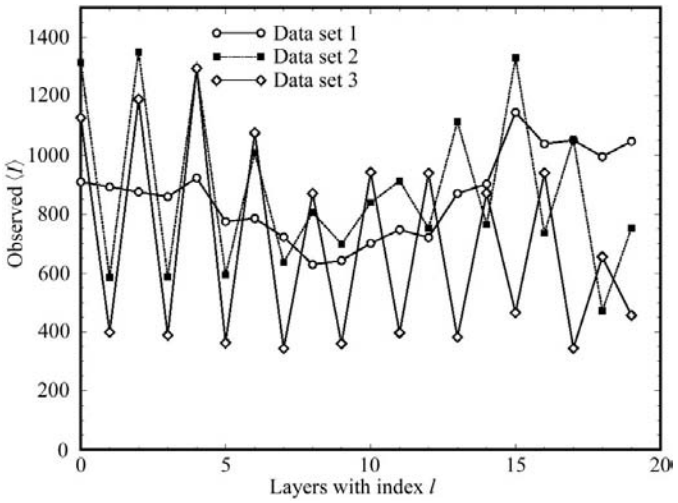

Figure 3
Observed variations of layer-averaged intensity in three data sets as a function of the reciprocal-space index l. These three sets are representatives of many data sets that we have collected. Data set 1 had a negligible lattice defect, with little variation in the plot.

Table 2
X-ray diffraction statistics of crystals with and without lattice-translocation defects.

(*a*) A typical crystal with defects. A full 180° X-ray diffraction data set from one crystal was collected with 0.5° per image at Advanced Photon Source, Chicago. Data were integrated with 0.5 mm spot radius and 5° three-dimensional profiling (ten images in this case) using *HKL*-2000 (Otwinowski & Minor, 1997). Mosaicities varied from 0.8 to 1.5° with an average of 1.15°. With this spot radius, data were 100% complete. Other spot radii were used in test integration runs, but failed to improve the statistics. Large spot radii led to apparent overlaps. This data set was used in analysis as data set 3 in Figs. 3 and 4.

Resolution(Å)	$I/\sigma(I)$	$R_{\rm merge}$
30.0–7.56	11.4	0.117
7.56–6.02	9.85	0.129
6.02–5.26	9.41	0.131
5.26–4.78	9.04	0.122
4.78–4.44	9.09	0.121
4.44–4.18	8.79	0.129
4.18–3.97	8.43	0.141
3.97–3.80	7.82	0.160
3.80–3.65	7.40	0.172
3.65–3.53	6.99	0.185
3.53–3.42	6.27	0.211
3.42–3.32	5.85	0.222
3.32–3.23	5.88	0.258
3.23–3.15	5.10	0.301
3.15–3.08	4.51	0.340
3.08–3.02	3.88	0.394
3.02–2.96	3.38	0.442
2.96–2.90	2.96	0.492
2.90–2.85	2.72	0.516
2.85–2.80	2.37	0.568
Overall, 30–2.8	7.55	0.171

(*b*) A crystal without defects. One crystal that suffered from negligible defects was obtained after the structure of φ29 DNA polymerase had been determined and partially refined using intensity-corrected data sets. This crystal also diffracted to 2.2 Å resolution. X-ray diffraction data from this crystal were then used for the final structure refinement (Kamtekar *et al.*, 2004) and are available from the Protein Data Bank (code 1xi1). Although the lattice-translocation defects strongly affected the quality of data, there was no direct correlation between the resolution limit of a crystal and the extent of the defects it suffered from.

Resolution(Å)	$I/\sigma(I)$	$R_{\rm merge}$
50.0–4.74	19.2	0.065
4.74–3.76	14.6	0.079
3.76–3.29	11.4	0.099
3.29–2.99	9.80	0.125
2.99–2.77	8.11	0.166
2.77–2.61	5.72	0.237
2.61–2.48	4.59	0.304
2.48–2.37	3.68	0.361
2.37–2.28	2.58	0.440
2.28–2.20	1.98	0.538
Overall, 50–2.2	12.4	0.097

$\Delta x = \Delta y = 0$, the intensity modulation caused by such lattice defects is simply a function of the reciprocal-space index l as shown in layer-averaged observed intensity (Fig. 3). The average intensity (on an arbitrary scale with an overall scale factor) for all layers can be represented by the parameter A as shown in (3) and the amplitude of variations between layers can then be represented by the parameter Ar. In this case, the maximum of average intensity is $(A + Ar)$ and the minimum is $(A - Ar)$. If we let $B = Ar = 2\kappa(1 - \kappa)$, (3) becomes a two-parameter linear equation

$$I_{\rm total} = fI_{\rm unit} = [A + B\cos(2\pi l\Delta z)]I_{\rm unit} \qquad (4)$$

or

$$\langle I_{\rm obs}\rangle_l = f\langle I_{\rm unit}\rangle_l = [A + B\cos(2\pi l\Delta z)]\langle I_{\rm unit}\rangle_l. \qquad (5)$$

The parameters A and B can be obtained using a standard least-squares minimization with the function Q to represent the sum of the squared differences,

$$Q = \sum_l \{\langle I_{\rm obs}\rangle_l - [A + B\cos(2\pi l\Delta z)]\}^2. \qquad (6)$$

When A and B are determined, the intensity-correction factor f can be applied to the observed data as $1/[A + B\cos(2\pi l\Delta z)]$ for both measured intensities and their errors. After applying the correction, the layer-dependent variations of intensity in data sets 2 and 3 have been completely eliminated (Fig. 4). Amplitude differences between data sets 2 and 1 and between data sets 3 and 1 are also reduced by 3.7 and 7.0%, respectively, with a maximal reduction of about 10% (Fig. 5). In this particular example, data sets 2 and 3 came from crystals that were soaked with the heavy atoms Pt and Hg, respectively. However, this correction cannot eliminate any intensity differences arising from non-isomorphism that may exist between data sets.

Rearranging the equation $r = 2\kappa(1 - \kappa)/(2\kappa^2 - 2\kappa + 1)$ and solving the quadratic equation in κ, we can estimate the lattice-defect fraction κ from the fitted parameters A and B (and their ratio $r = B/A$) as follows

$$\kappa = 1/2 \pm 1/2[(1 - r)/(1 + r)]^{1/2}. \qquad (7)$$

The equations can be simplified for the special situation in which $\Delta z = 1/2$, as in the case of data set 2 (to a very good approximation to the 3.2 Å resolution limit), in which the layer-averaged intensity constant is different if index l = odd or l = even. We define their ratio as $v = \langle I\rangle_{l=\rm odd}/\langle I\rangle_{l=\rm even}$, which is independent of index l (and also the resolution) of the data. The peak ratio (p) of the Patterson functions between

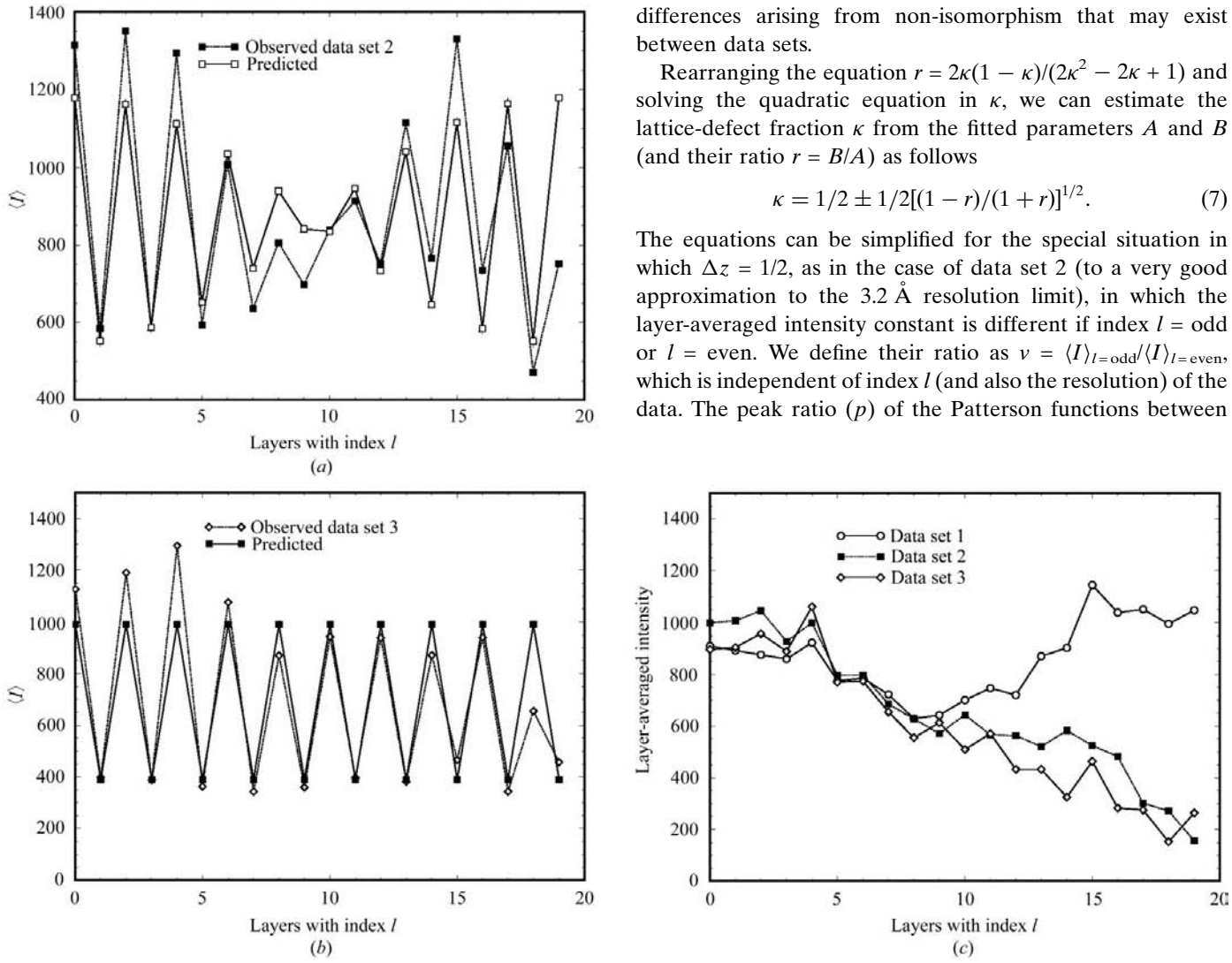

Figure 4
Intensity correction for the lattice-translocation defect. (*a*) The observed (filled squares) and predicted (open squares) layer-averaged intensity variations using the formula $f = [A + B\cos(2\pi l\Delta z)]$ for data set 2. The defect fractions are 75.1:24.9% and the translocation vector is $\Delta z = 0.4735$. (*b*) The observed (open diamonds) and predicted (filled squares) layer-averaged intensity variations for data set 3. The defect fractions are 83.3:16.7% and the translocation vector is $\Delta z = 1/2$. (*c*) Layer-averaged intensity after application of the correct factor $1/f$. Data set 1 had its first constructive diffraction power addition occurring at index l approximately equal to 15 (with a small translocation vector of about $\Delta z = 0.06$–0.07, as explained in the text). No correction was applied to data set 1, because its defect fraction was negligibly small and its $\mathbf{t_d}$ value could not be accurately determined because it was too close to the origin of the native Patterson function.

(0, 0, 1/2) and (0, 0, 0), each of which can be Fourier summed through the two groups with either index l = odd or l = even (including $l = 0$), is also independent of resolution,

$$p = (1 - v)/(1 + v) = r. \quad (8)$$

Therefore, the Patterson peak ratio ($p = r$) itself directly becomes the coefficient in front of the $\cos(2\pi\mathbf{ht_d})$ term of (3). The lattice-defect fraction κ and the Patterson peak ratio p are related to each other in the same way as κ and r in (7). The defect fraction κ and the intensity ratio v are related by

$$\kappa = 1/2 \pm 1/2v^{1/2}. \quad (9)$$

After fitting the parameters derived from Fig. 4 into (7–9), the two defect fractions for data set 2 are 75.1 and 24.9%; they are 83.3 and 16.7% for data set 3. Upon application of the intensity correction to the individual data sets, the native Patterson peaks were almost completely eliminated, as predicted from $p = r$ (8) and from eliminating the variation in the layer-averaged intensity (Fig. 4). While the large intensity modulations owing to lattice-translocation defects are correctable using (5) and (6), small modulations owing to the changes in spot shapes associated with the defects are partially correctable by intensity integration using appropriate integration shapes and sizes for each image (Fig. 6). Finally, the formulation of intensity correction described in this work should be generally applicable with any three-dimensional $\mathbf{t_d}$ as long as the layer averaging can be carried out for observed intensities in planes perpendicular to that vector.

3. Application of the correction formulation to the case of φ29 DNA polymerase

We have demonstrated above the application of an intensity correction to the individual data sets of φ29 DNA polymerase crystals. Applying the correction, we have identified one Hg and one Pb derivative in addition to an SeMet-substituted protein derivative. An example of an experimental electron-density map generated using phases derived from these derivatives after intensity correction is shown in Fig. 7. The phasing was carried out sequentially from the single isomorphous replacement with anomalous scattering (SIRAS), multiple isomorphous replacement with anomalous scattering (MIRAS) and multiwavelength anomalous dispersion (MAD) methods, followed by density modification. Using this map, the structure was readily interpreted and the model refined to a free R factor of 27.8% at 2.2 Å (Kamtekar *et al.*, 2004). This provided the first structure of a protein-primed DNA polymerase, a project which began in our laboratory nearly 15 years ago. When the intensity correction was not applied, the heavy-atom structures corresponding to a single lattice could not be determined. Even when the correct coordinates of the heavy-atom structures of one lattice were supplied, the heavy-atom parameters could not be refined and the correct experimental phases could not be obtained because of instability in the refinement caused by the presence of large intensity contributions from another lattice.

A packing analysis of the lattice of φ29 DNA polymerase crystals shows that a translation of $\Delta z = 1/2$ for the entire layer of molecules is indeed stereochemically permissible. This translation occurred randomly. If the translation occurred in every layer, a new lattice would emerge with a new β value. If the translation occurred in every other layer, the unit-cell edge would double ($2a\sin\beta$) in the direction perpendicular to the layers (a^*) with a new orthogonal β angle; in every third layer, the unit-cell edge would triple ($3a\sin\beta$) and so on. In all these cases, only sharp Bragg reflections would be observed. Similarly, only sharp Bragg reflections would be expected if a crystal contained only two distinct coherent lattice domains. The hallmark of the lattice-translocation defect is the random distribution of the translated layers and the presence of the

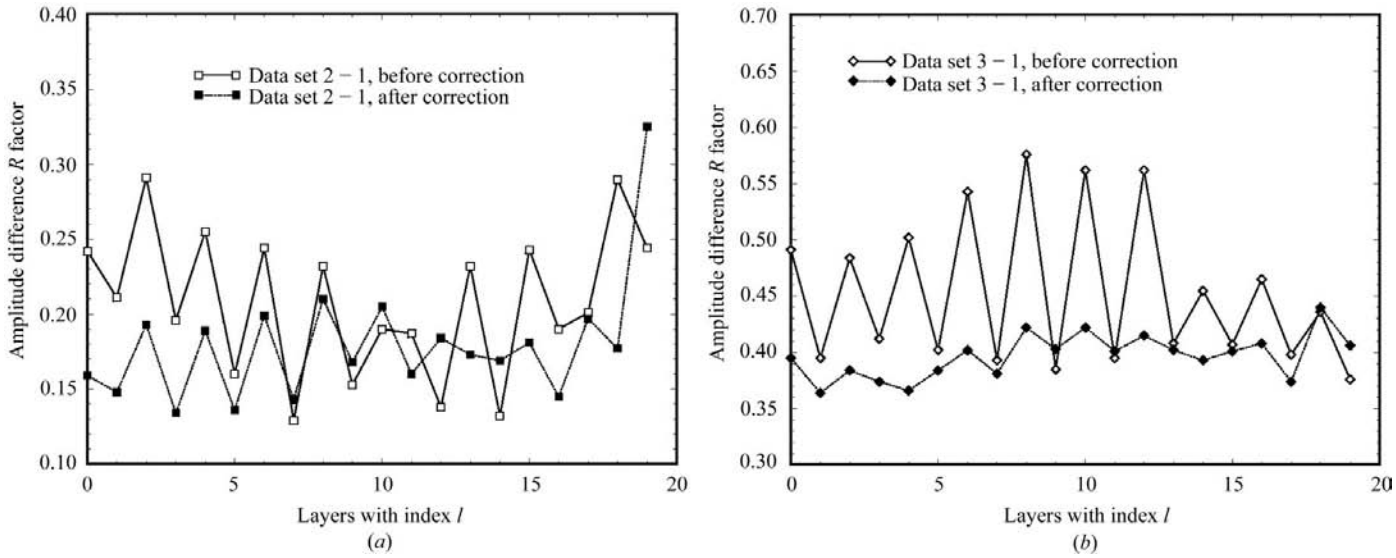

Figure 5
Reduced isomorphous differences between data sets. (*a*) Amplitude differences (as measured by R factor) between data sets 2 and 1 before (open squares) and after (filled squares) intensity correction. The overall reduction of the difference R factor is 3.7%. (*b*) Amplitude differences between data sets 3 and 1 before (open diamonds) and after (filled diamonds) intensity correction. The overall reduction of the difference R factor is 7.0%.

observable streaky reflections (for an example, see Fig. 1). The two-lattice treatment described here does not necessarily imply that two coherent crystals coexist in a physical sense. Instead, it purely provides a mathematical means to treat lattice-translocation defects (Bragg & Howells, 1954) in which all layers with the $+a$ displacement are grouped into one lattice and all layers with the $-a$ displacement are grouped into another.

The lattice-translocation defect theory also provides a basis for the poor diffraction seen for some macromolecular crystals. When a translocation occurs with $\Delta z = 1/2$ as in the case of φ29 DNA polymerase crystals, half of the data have reduced intensities and thus reduced effective resolution owing to both the streaky feature of reflections and the cancellation of diffraction powers between the two lattice domains. When a translocation occurs with $\Delta z = 1/3$, as in the case of HslU–CodW crystals, a detailed analysis of which will be described elsewhere, two-thirds of the data have reduced intensity and resolution. If both the translocations ($\Delta z = 1/2$ and $\Delta z = 1/3$) were to occur in a single crystal, five-sixths of the data would have reduced intensity and resolution and only one-sixth of the data would not be affected. Therefore, the combined effects of multiple lattice-translocation defects in one crystal could lead to a dramatic cancellation of overall diffraction power and result in streakiness for the majority of the data.

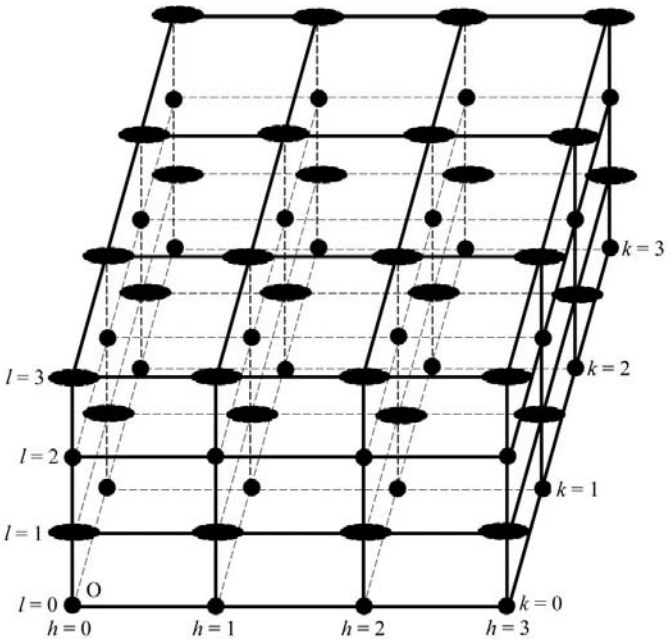

Figure 6
Reciprocal-space lattice in the presence of the lattice-translocation defect of Fig. 3 with $\Delta z = 1/2$. When the defect occurs with low frequency ($\kappa < 5\%$), the streaky features are limited to slightly elliptically shaped lattice points. Such a limited defect can be partially corrected during intensity integration using an elliptical spot-shape option when the crystal is rotated along the a^* axis during data collection, but it cannot when the axis of the crystal rotation is perpendicular to a^*. In the latter case, streaky features would appear only in one orientation and many extra Bragg reflections would not be predictable with a standard definition of the mosaicity in a second orientation that is 90° away from the first orientation.

The reason that the lattice-translocation defect has not completely destroyed diffraction power and is correctable in φ29 DNA polymerase crystals is that there might be only one permissible translocation of $\Delta z = 1/2$. Therefore, the prevention and reduction of the lattice-translocation defects might be an effective way to improve the diffraction quality of poorly diffracting macromolecular crystals.

4. Lattice basis for varying reflection strength and shape

Howells and colleagues have already described the changing shape of reflections as a function of phase shifts (Bragg & Howells, 1954; Cochran & Howells, 1954). Because most of our data sets manifest $\Delta z = 1/2$, we have the simpler situation in which phase shifts have only two values, either $2n\pi$ or $(2n + 1)\pi$. For the X-ray diffraction lunes with index l = even or zero (Fig. 1), the lattice-translocation defect has no apparent effect on the strength and shape of reflections, because scattering waves before and after the translocation interface travel in-phase. Entire mosaic blocks of the crystal behave like a 'perfect' lattice and we observe typical sharp Bragg reflections (Fig. 1). This is equivalent to reducing the unit-cell dimension in the c direction by a factor of two with a new perfect lattice. The fractional translation of half a unit cell in the original unit cell becomes unity in the new unit cell (Fig. 2). This new unit cell explains the sharp Bragg reflections

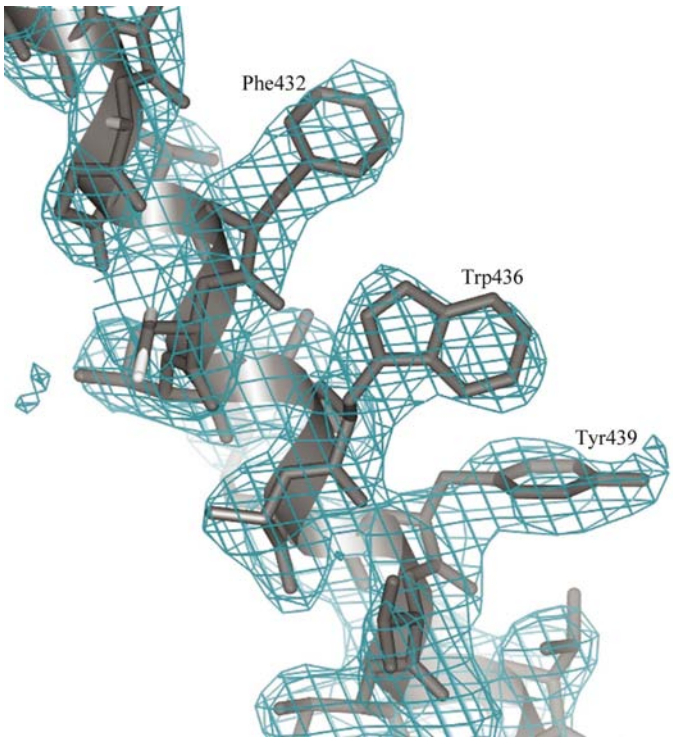

Figure 7
An experimental electron-density map of φ29 DNA polymerase at 2.5 Å using phases initially derived from Hg derivatives after their X-ray intensities were corrected, followed by density modification and twofold non-crystallographic symmetry averaging. This map is contoured at 1.0σ and superimposed with the finally refined model.

with l = even or zero, even though the occupancy of molecules on lattice points is not evenly distributed.

For lunes with the index l = odd (Fig. 1), however, the lattice-translocation defect has a dramatic effect on both the strength and shape of reflections. The amplitude of scattered waves cancel out each time a defect interface is crossed owing to a change in diffraction phase of π and $(2n + 1)\pi$. When two fractions (the translocated and untranslocated blocks) of the defected lattice are equal ($\kappa = 1 - \kappa = 50\%$), two cases are possible. If such translocation occurs periodically in every other layer, the defective lattice re-merges as a new perfect lattice with the unit-cell length doubled ($2a\sin\beta$) in the a^* direction without any smeary features for these reflections (precisely zero intensity). However, if the translocation occurs randomly (as observed in some of the φ29 DNA polymerase crystals), the size of the unit cell is only statistically doubled and half of the reflections (all l = odd reflections in the original lattice) remain extremely smeary, cannot be integrated and would statistically have zero intensity.

An alternative view of the defects for l = odd lunes is that they effectively reduce the size of the crystal mosaic blocks. The smaller the block size, the broader the reflections. Under an extreme condition with only a single layer of the lattice (*i.e.* a two-dimensional crystal), a continuous streaky diffraction perpendicular to this plane will be observed. Therefore, a relative loss of intensity in these reflections arises from a redistribution of scattering powers away from expected Bragg lattice points.

The streaky patterns are slightly more complicated when Δz is not precisely equal to 1/2; for example, in data set 2 (Fig. 3) where $\Delta z = 0.4735$. For reflections with l = odd and $l < 6$ (or with l = even and $l > 12$), scattering amplitudes were partially cancelled out and the shape of reflections was substantially broadened. In contrast, reflections with l = odd and $l > 12$ (or with l = even and $l < 6$) are not affected by these translocation defects. Other reflections with index $6 < l < 12$ had variable widths of reflection spot size and varying degrees of diffuseness.

5. Concluding remarks

Lattice-translocation defects have been observed in this laboratory in at least four separate systems, including RuvC, T7 RNA polymerase, HslU–CodW (a complex homologous to HslU–HslV) and φ29 DNA polymerase. We believe that this crystal defect has been an underreported crystallographic problem during the past half-century because no solution to the problem was available. The approach we have described here provides a general method for computationally correcting lattice-translocation defects and makes structure determination with crystals containing these defects more tractable.

We thank Professor P. B. Moore for interesting discussion on the subject and comments on this manuscript. This work was supported in part by NIH grant R01-GM57510.

References

Bragg, W. L. & Howells, E. R. (1954). *Acta Cryst.* **7**, 409–411.
Cochran, W. & Howells, E. R. (1954). *Acta Cryst.* **7**, 412–415.
Kamtekar, S., Berman, A., Wang, J., Lazaro, J. M., deVega, M., Blanco, L., Salas, M. & Steitz, T. A. (2004). *Mol. Cell*, **16**, 609–618.
Meijer, W. J. J., Horcajadas, J. A. & Salas, M. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 261–287.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rudolph, M. G., Wingren, C., Crowley, M. P., Chien, Y. H. & Wilson, I. A. (2004). *Acta Cryst.* D**60**, 656–664.
Wang, J., Sattar, A. K., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997). *Cell*, **89**, 1087–1099.