Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Correction of X-ray intensities from an HslV–HslU co-crystal containing lattice-translocation defects

Jimin Wang,[a]* Seong-Hwan Rho,[b] Hyun Ho Park[b] and Soo Hyun Eom[b]

[a]Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA, and [b]Department of Life Science, Gwangju Institute of Science and Technology (G-JIST), Gwangju 500-712, South Korea

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

Received 21 December 2004
Accepted 24 March 2005

PDB Reference: HslV–HslU, 1yyf, r1yyfsf.

Because of lattice-translocation defects, two identical but translated lattices can coexist as a single coherent mosaic block in a crystal. The observed structure in such cases is a weighted sum of two identical but translated structures, one from each lattice; the observed structure factors are a weighted vector sum of the structure factors with identical unit amplitudes but shifted phases. The correction of X-ray intensities from a single crystal containing these defects of the hybrid HslV–HslU complex, which consists of *Escherichia coli* HslU and *Bacillus subtilis* HslV (also known as CodW), is reported. When intensities are not corrected, a biologically irrelevant complex (with CodW from one lattice and HslU from another) is implied to exist. Only upon correction does a biologically functional CodW–HslU complex structure emerge.

1. Introduction

Lattice-translocation defects are caused by the random translocation of some layers by a fixed constant within the stacked layers of a crystal. When such events result in the fragmentation of the crystal into smaller mosaic blocks, the observed intensities are simply additive from each block and independent of the translocation vector. When such events occur within one single coherent mosaic block, interference in X-ray diffraction is observed, with the intensities modulated by a factor that is a function of this translocation vector and the fractions of the translocated and untranslocated layers (Wang *et al.*, 2005). This phenomenon was first observed around 50 years ago (Bragg & Howells, 1954; Cochran & Howells, 1954). Recently, an equation for the modulation factor caused by these defects was formulated and the observed intensities became correctable using this factor (Wang *et al.*, 2005).

The hallmark of lattice-translocation defects is a continuum distribution, resulting in streaky features in some reflections. When the translocation occurs regularly every second, third, fourth or other nth layers, the length of the unit-cell dimension perpendicular to the layers simply increases by a factor of two, three, four or n, with an extra one, two, three or $n - 1$ reflections observed between two adjacent original Bragg reflections. When the translocation occurs randomly, the number of extra reflections also varies randomly. A spectrum of randomly varying numbers of extra reflections is a continuum distribution, with streaky features in reflections. However, not all original Bragg reflections may have this streaky feature. In crystals of φ29 DNA polymerase the translocation vector is (0, 0, 1/2) and the streaky feature is invisible for all reflections with index $l = 2n$ (Wang *et al.*, 2005). In the crystal of the CodW–HslU complex described in this

report, the translocation vector is (0, 0, 1/3) and the streaky features are invisible for all reflections with index $l = 3n$. In addition to these two examples whose atomic structures are now known, there were three other cases of crystals of biological molecules in the literature whose diffraction patterns also had some streaky and sharp reflections and whose atomic structures remained unknown (Howells & Perutz, 1954; Glauser & Rossmann, 1966; Pickersgill, 1987).

Two orientations of the ATPase HslU have been observed in two different quaternary arrangements in the ATP-dependent ATPase–peptidase HslV–HslU complex (Bochtler *et al.*, 2000*a*; Sousa *et al.*, 2000; Wang, Song, Franklin *et al.*, 2001). In one orientation the interface between HslU and HslV is formed by the largely disordered intermediate domain or domain I of HslU (Bochtler *et al.*, 2000*a*) and in the second type it is formed by the nucleotide-binding domain of HslU (Sousa *et al.*, 2000; Wang, Song, Franklin *et al.*, 2001). Whether both or only one of the two quaternary arrangements is biologically relevant has been the center of a recent debate (Ishikawa *et al.*, 2000; Bochtler *et al.*, 2000*b*; Song *et al.*, 2000; Ramachandran *et al.*, 2002). An analysis using the calculated data as an approximation for the observed X-ray data that were unavailable at the time showed that both quaternary arrangements of the HslV–HslU complexes could coexist in the original data (Wang, 2001). When the observed data became available, a severe violation of systematic extinctions for the space-group assignment was noted (Wang, 2001). This violation was later attributed to previously unreported streaky reflections present in the data (Bochtler *et al.*, 2001; Wang, 2003).

In this report, we describe the structure determination of a homologous HslV–HslU complex from a crystal also containing streaky reflections. This complex consists of *Escherichia coli* HslU and *Bacillua subtilis* HslV and is fully functional (Kang *et al.*, 2001; Rho *et al.*, 2005). This complex structure is crucial for understanding the unorthodox activation mechanism of *B. subtilis* HslV (Kang *et al.*, 2001). We will show that the lattice-translocation defect hypothesis is sufficient and necessary to explain both the observed modulation in intensities and the presence of streaky reflections in the observed data. Even though there is only one biological complex in the structure, we will show that a biologically irrelevant CodW–HslU complex (with CodW from one lattice and HslU from another but translocated lattice) is implied to exist when the intensities are not corrected for these defects. Only upon intensity correction does a biologically relevant complex emerge. Finally, because of the technical difficulties introduced by the large number of streaky reflections, some aspects of data processing (such as in autoindexing, geometry parameter refinement, integration and intensity symmetry scaling) are also described.

2. Sample preparation, enzymatic assay and crystallization

Purification of HslU and HslV from both *E. coli* and *B. subtilis* has been described elsewhere (Kang *et al.*, 2001; Rho *et al.*, 2005). In *B. subtilis*, HslU and HslV are also known as CodX and CodW, respectively. Enzymatic assays for the peptidase–ATPase complexes from *E. coli* (*i.e.* HslV–HslU) and from *B. subtilis* (*i.e.* CodW–CodX) and hybrid complexes (*i.e.* HslV–CodX or CodW–HslU) have also been described elsewhere (Rho *et al.*, 2005). In summary, the descending order of enzymatic activity for ATP-dependent proteolysis is HslV–HslU $\simeq$ CodW–CodX >> CodW–HslU > HslV–CodX. The hybrid CodW–HslU complex in this study is enzymatically active (Kang *et al.*, 2001).

The purified CodW–HslU complex crystallized in the *H*32 space group (*R*32 in the hexagonal lattice setting), with unit-cell parameters $a = b = 181.3$, $c = 530.0$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. An analysis of dissolved crystals shows that there was an approximately 1:1 (not 1:2, see below) molar ratio of HslU and CodW. In co-crystallization, dATP was included in order to capture the activated ATPase–peptidase complex structure, but only the hydrolytic product dADP is observed (see below). A similar hydrolysis of dATP by HslU has also been observed in another complex during crystallization (Wang, Song, Franklin *et al.*, 2001). X-ray diffraction data were collected on beamline 18B of the Photon Factory, Japan using an ADSC Quantum 4R detector. Beginning with a crystal orientation that yielded the image shown in Fig. 1, we collected 200° data in 1° increments. The oscillations were taken along a horizontal axis approximately perpendicular to the l axis. Data were processed using *HKL*2000 (Otwinowski & Minor, 1997).

3. Sharp and diffuse patterns in X-ray diffraction and intensity integration

The CodW–HslU co-crystals diffracted to 3.6 Å in the l direction and 4.5 Å in the h and k directions (Fig. 1). A striking feature in Fig. 1 is that only one-third of the Bragg reflections remain as well defined sharp spots and the remaining two-thirds of the Bragg reflections have non-Bragg diffraction features with streaky features. These streaky features made the determination of the geometric centroids of the reflections very difficult and also somewhat reduced the intensities of affected reflections. These features have been problematic for all data-processing steps that were dependent on the geometric centroids and intensities of reflections, including steps in autoindexing, geometric parameter refinement, integration, intensity symmetry scaling and post-refinement.

In the presence of lattice-translocation defects in the CodW–HslU co-crystal, autoindexing required a special treatment of spot size and intensity profiles. When a standard procedure was used (for example, using the typical spot size of 0.35 mm for sharp Bragg reflections with 3σ cutoff for the given crystal), streaky reflections (about two-thirds of all reflections) were under-represented in sampling, because they were too weak in intensity or failed to provide accurate centroids. This procedure often resulted in incorrect space groups that could not predict many of these streaky reflections. Only when an artificial large spot size of 0.65–0.70 mm

was used in combination with a low weak level cutoff of 1–2σ was an index in the correct space group $H32$ obtained.

Upon indexing, the direction of the streaky features was determined to be along the l axis, suggesting that layers were stacked along the l axis and translocation of layers occurred during the stacking (Wang *et al.*, 2005). Examination of indices also showed that reflections with index $l = 3n$ were sharp and the remaining two-thirds of the reflections had streaky features (Fig. 1). These systematic streaky features should not introduce additional anisotropy to the data, even though they could visually make the h and k directions appear weak.

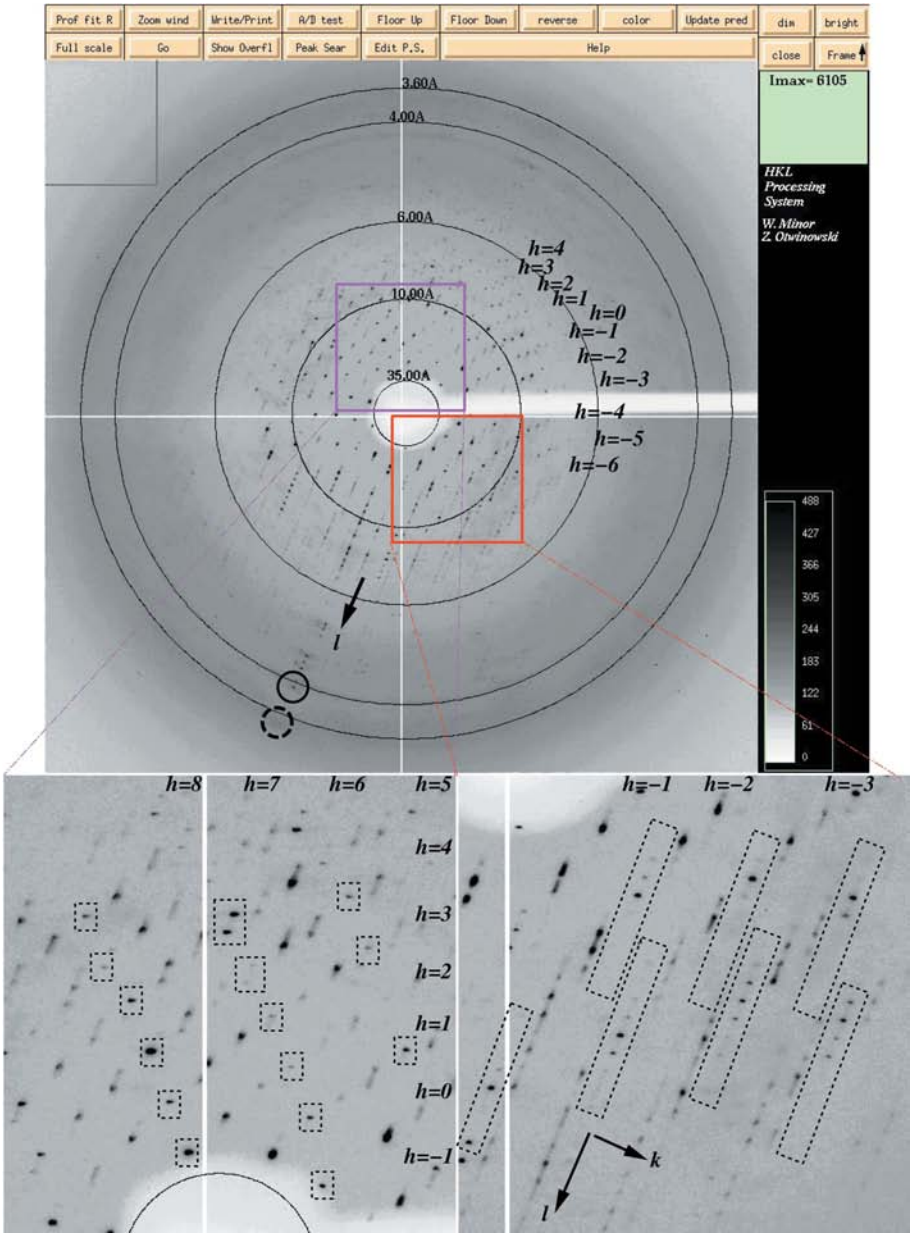

Figure 1
Observed alternating sharp–diffuse reflections. In the l direction, strong reflections are observed beyond 4.0 Å (large concentral circles) and some reflections are observed at 3.6 Å (reflections with small circles); in the h and k directions reflections barely pass 4.5 Å. Only the third of the reflections with index $l = 3n$ (boxed) have typical Bragg spots and the remaining reflections have streaky features along the l direction. Data were processed to 4.16 Å, slightly below an ice ring at 4.15 Å.

Geometric parameter-refinement procedures could not handle streaky reflections well. In the orientation of Fig. 1, the comet-tail-like streakiness along the l-axis direction is clearly visible and distributed in the xy plane of the detector, resulting in apparently 'elongated or ellipsoid profiles' for two-thirds of the reflections. At 90° away from this orientation, the streakiness is no longer visible and is distributed along the direction of the incident X-ray beams, resulting in apparently 'increased mosaicity'. Attempts were made to partially model streaky reflections using apparent ellipsoid profiles (0.35 × 0.70 mm) or increased mosaicity (up to 3.5°), but they were unsuccessful because the centroids of streaky reflections in both orientations were not well defined along at least one direction. With these attempts, many more streaky reflections were indeed well predicted; however, the refinement failed to converge, resulting in a very large χ^2 of about 4. For example, the frame-based mosaicity refinement showed that the refinement had not yet converged for the above-mentioned second orientation even when its mosaicity (over 3.5°) increased to about seven times that of the first orientation. According to the χ^2 values, the best refinement results were obtained when the spot radius was 0.35 mm without using the ellipsoid model and with a fixed mosaicity of 0.648°. Under these conditions, streaky reflections contributed very little to the refinement.

The best parameters for the intensity integration of sharp Bragg reflections were found to also be the best ones for the entire data set, including the streaky reflections in the CodW–HslU co-crystal. Unlike the direct optimization of geometric parameter refinement (as judged from the χ^2 values), the optimization of the parameter for integration was only indirect and required the results from the later processing steps, including the post-refinement during intensity symmetry scaling and native Patterson calculation. For this reason, we integrated intensities with fixed (0.35, 0.40 or 0.65 mm) or varying spot sizes and with fixed (retrospectively, 0.648°) or varying mosaicity in eight possible combinations. For each combination, intensities were scaled using post-refinement with an overall fitted mosaicity, group mosaicity per 5° data or individual mosaicity per frame. As judged from the intensity agreement, merging R factors and χ^2 criteria

(and native Patterson calculations below), the optimized parameters for intensity integration were a fixed spot size of 0.35 mm and a fixed mosaicity of 0.648°, which was slightly larger than the post-refinement mosaicity of 0.573° for the entire crystal.

The lattice symmetry was determined to be *H*32 (hexagonal setting of *R*32). An intensity-agreement merging *R* factor for intensity scaling was 17.0% in *H*32, which was only marginally worse than 16.4% in *H*3 (hexagonal setting of *R*3). Below, we have also ruled out the possibility that the dyad of *H*32 was introduced through 'statistical twinning' (Bragg & Howells, 1954), because the two fractions were not at 50% each (see below). Large merging *R* factors of 17.0% for this crystal at 4.16 Å and of 17.1% at 2.8 Å for the φ29 DNA polymerase crystals (Kamtekar *et al.*, 2004; Wang *et al.*, 2005) are a consequence of the presence of streaky reflections for two-thirds and half of the data, respectively.

The resolution of the X-ray data was determined to be 4.16 Å, which was the highest obtainable resolution below an ice ring at 4.15 Å. The data-processing statistics are summarized in Table 1. In the highest resolution shell (4.16–4.31 Å), the average $I/\sigma(I)$ is 2.4, χ^2 is 1.09 and the merging *R* factor was slightly over 100%. Sharpening (see below) increased the effective resolution. Moreover, the resulting unbiased electron-density maps show well resolved β-strands, some side-chain features in α-helices and the bound nucleotides, suggesting that 4.16 Å does not overstate the resolution of the data.

4. Native Patterson peaks and the observed intensity modulation

An outstanding native Patterson peak at (0, 0, 1/3) was mainly a consequence of intensity cancellation by the destructive

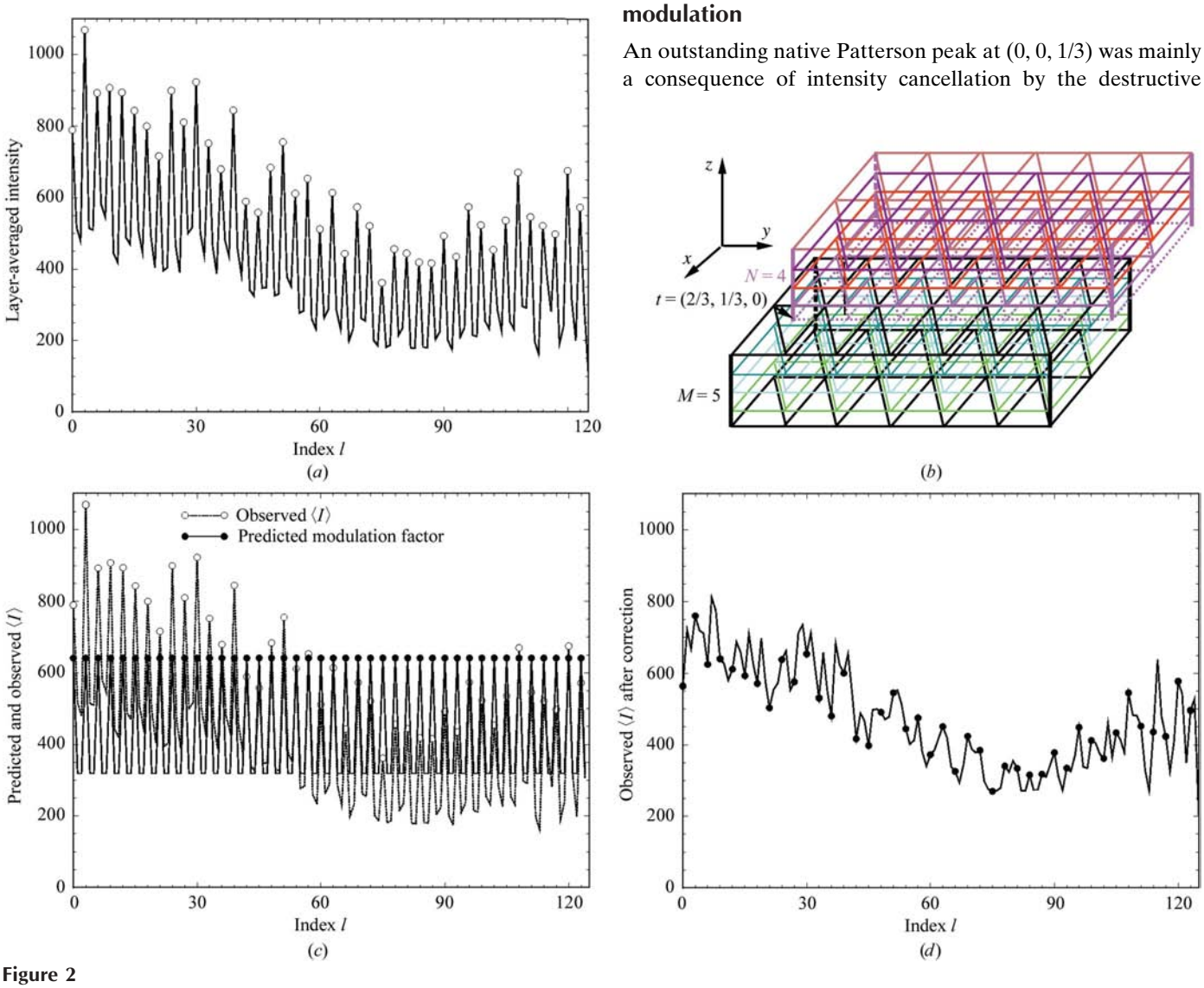

Figure 2
Observed high-frequency variations of layer-averaged intensity as a function of index *l* and intensity correction. (*a*) Observed systematic variations with open circles marked for index $l = 3n$. (*b*) A model for lattice defects in which *N* layers of one lattice domain are translated relatively to *M* layers of another with $\mathbf{t}_d$ = (2/3, 1/3, 0). Fractions of *M* and *N* are the parameters to be fitted. In space group *H*32, a $\mathbf{t}_d$ of (2/3, 1/3, 0) is equivalent to (0, 0, −1/3) when the origin is shifted from (0, 0, 0) to (2/3, 1/3, 1/3) for reference. Note that the modulation factor is identical for both a $\mathbf{t}_d$ of (0, 0, 1.3) and of (0, 0, −1/3). (*c*) Model-predicted modulation factors (solid lines) and observed variations from (*a*) (dotted lines) have the same distribution of minima and maxima. Index $l = 3n$ is marked with circles (open for observation and closed for prediction). (*d*) After correction, the high-frequency variations in layer-averaged intensity a function of index *l* have been completely removed.

Table 1
X-ray diffraction data merging *R*-factor statistics.

Resolution range (Å)	$I/\sigma(I)$	χ^2	*R* factor (%)
35.00–8.92	29.81	1.07	6.7
8.92–7.10	26.62	1.26	9.1
7.10–6.21	16.32	1.18	17.2
6.21–5.64	10.17	1.11	27.9
5.64–5.24	7.70	1.07	35.7
5.24–4.93	7.85	1.09	35.2
4.93–4.68	6.72	1.10	40.6
4.68–4.48	5.80	1.09	47.6
4.48–4.31	4.54	1.06	58.1
4.31–4.16	2.40	1.09	>100
All	14.7	1.11	17.0

interference between the translated and untranslated lattices. Its peak height was 26.1% of the origin-peak height and was independent of the resolution ranges used for calculations. When using a small spot size (0.35 mm) during integration, more pixels in streaky reflections were outside the integration box than were pixels for sharp Bragg reflections, which may lead to some systematic underestimation of streaky reflections. However, this peak was only slightly reduced (23–24%) when the native Patterson maps were calculated using data integrated with larger spot sizes (0.40–0.65 mm). This result suggests that the contribution to this peak from a systematic underestimation of streaky reflections is negligible in comparison with the intensity cancellation by the destructive interference between the translated and untranslated lattices. Therefore, there was no advantage in using a large spot size for streaky reflections either during parameter refinement (see above) or during integration.

The observed native Patterson peak is consistent with the observed systematic variations in layer-averaged intensity as a function of index l (Fig. 2*a*). When index $l = 3n$, the layer-averaged intensities are at maxima; when index $l \neq 3n$, they are at minima for which there are no systematic differences between two subsets of reflections with indices $l = 3n + 1$ and $l = 3n + 2$. The layer-averaged intensity ratio between any two adjacent indices (index $l \neq 3n$, weak zones, and $l = 3n$, strong zones) is nearly constant through the entire data set and approximately equals an overall ratio of 0.512 for all weak and strong zones. The observed intensity modulation is fully predictable using the lattice-translocation defect theory (Wang *et al.*, 2005).

5. Correction of intensities using the lattice-translocation defect theory

We propose a lattice-translocation defect model to explain the observed intensity modulation (Fig. 2*b*). As is evident from the direction of the reflection streakiness, layers are stacked along the l axis (Wang *et al.*, 2005). Lattice-translocation defects occur to yield two associated lattice domains, one consisting of M layers and the other N layers. In this specific model (Fig. 2*b*), the translocation vector is (2/3, 1/3, 0). If the reference origin is shifted from (0, 0, 0) to its equivalent new origin (2/3, 1/3, 1/3) in the *H*32 space group, the vector now becomes (0, 0, −1/3). According to previously formulated equations (Wang *et al.*, 2005), the observed total structure factor $\mathbf{F}_{\rm total}$ is related to the unit structure factor $\mathbf{F}_{\rm unit}$ in the two lattice domains by the translocation vector $\mathbf{t}_{\rm d}$, weighted by the numbers of layers (or unit cells) M and N in vector summation,

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

where $\mathbf{h}$ is the index (*hkl*). In the case of the CodW–HslU co-crystal, $\mathbf{t}_d$ is (0, 0, −1/3) and the corresponding total and unit intensities $I_{\rm total}$ and $I_{\rm unit}$ are related by

$$I_{\rm total} = fI_{\rm unit} = [A + B\cos(2\pi l/3)]I_{\rm unit}, \quad (2)$$

where f is the modulation factor and A and B are the parameters to be fitted (Wang *et al.*, 2005).

Upon least-squares fitting, $A = 426.373$, $B = 214.699$ and the ratio $r = B/A = 0.503$ for the CodW–HslU co-crystal. This ratio approximately equals an overall ratio of 0.512 (see above) for layer-averaged intensities between all weak and strong zones. Using the previously formulated equations, this value corresponds to the two fractions of 21% *versus* 79% (not 50% *versus* 50%!) in intensity composition (Wang *et al.*, 2005). As seen in Fig. 2(*c*), the distribution of minima and maxima in the observed data matches reasonably well with the predicted modulation factor. Upon application of the correction factor $1/f$ to the observed data, we obtain the new intensities of the structure factor that now corresponds to one structure in one lattice. This correction removes the high-frequency variations in every third index l in layer averaging of the observed intensities without affecting the low-frequency variations, which are also a function of resolution in general (Fig. 2*d*). When the corrected intensities were used for native Patterson map calculations, the peak at (0, 0, 1/3) was now either completely eliminated or less than the 3σ noise level in all resolution ranges.

In our lattice-translocation defect model, we describe that one lattice domain has M layers and the other has N layers with an abrupt transition between them (Fig. 2). (1) and (2) also apply to the case of a smooth transition with few extra layers between two lattice domains, as long as the number of transition layers is negligible in comparison with M and N and the two lattice domains are within the coherent length of the X-ray scattering. The translocation of layers has a zero component in its vector in the direction perpendicular to the layers in both the case of HslV–HslU (2/3, 1/2, 0) described here and the previously described case (Wang *et al.*, 2005) of φ29 DNA polymerase crystals (0, 0, 1/2). Our model should also be applicable to the three-dimensional lattice-translocation defect with a non-zero component in all directions.

The lattice-translocation defects destroy neither the symmetry of each lattice domain nor the symmetry of the entire crystal. In Fig. 2(*b*), the translocation vector is (2/3, 1/3, 0) for the origin at (0, 0, 0), it is (0, 0, −1/3) for the origin at (2/3, 1/3, 1/3) and so on. If the origin is shifted from one lattice domain to another, the direction of this vector is inverted. Therefore, in this single lattice-translocation model all six symmetry-related peaks observed in native Patterson

Table 2
X-ray diffraction data and refinement statistics.

Values in parentheses are for the highest resolution shell.

Space group	*H*32
Unit-cell parameters (Å, °)	$a = b = 181.286$, $c = 529.967$, $\alpha = \beta = 90$, $\gamma = 120°$
Resolution (Å)	35–4.16 (4.31–4.16)
No. of observations	740249
No. of rejections	5918 [about 0.7%]
No. of unique reflections	25592 (2538)
$R_{\rm merge}$ (%)	17.0
$I/\sigma(I)$	14.7 (2.4)
Completeness (%)	100 (100)
Refinement resolution (Å)	35–4.16 (4.38–4.16)
R factor (%)	28.0 (30.6)
Free *R* factor (%)	34.6 (37.5)
Subunit content	2 HslU plus 2 dADP, 2 CodW
No. of non-H atoms	9216
R.m.s.d. bonds (Å)	0.033
R.m.s.d. bond angles (°)	2.9
R.m.s.d. torsion angles (°)	11.1
PDB code†	1yyf

† The originally integrated intensity data, the corrected and sharpened observed structure-factor amplitudes and the calculated structure factors, as well as the coordinates, are available from this single entry.

maps, (0, 0, 1/3), (0, 0, 2/3), (2/3, 1/3, 2/3), (1/3, 2/3, 1/3), 1/3,2/3, 0) and (2/3, 1/3, 0), are accounted for. The predicted modulation of intensity is identical for all symmetry-related vectors and independent of which symmetry-related translocation vector is chosen.

Trame & McKay (2001) have recently reported another case of lattice-translocation defects in crystals of HslU. They independently treated all symmetry-related translocation vectors with an equal fraction in their model. Their treatment explains part, but not all, of the observed intensity modulation. As argued above, it is not necessary to treat all symmetry-related translocation vectors as independent fractions. It is possible that a better correction for their observed data might be obtainable with a model in which the two fraction *M* and *N* are allowed to vary.

6. Structure determination and refinement of the HslV–HslU complex

The structure determination was carried out using molecular replacement as implemented in the program *CNS* (Brünger *et al.*, 1998). The protein-alone search models were hexameric HslU and dedocameric HslV generated from the coordinates of 1g4a (Wang, Song, Franklin *et al.*, 2001). Each correct solution in cross-rotation and translation searches was at least 2σ above the next highest noise peaks when using the corrected intensity data between 15 and 4.5 Å. Similar outstanding solutions were obtained when only two subunits of HslU and HslV were used as search models. These solutions were robust and independent of the resolution ranges of the data used. When HslU and HslV solutions were combined for the determination of the common crystallographic origin, only the biologically relevant CodW–HslU complex was obtained (Fig. 4). With this solution, the overall *R* factor and free *R* factor were both about 51% after one body per subunit rigid-

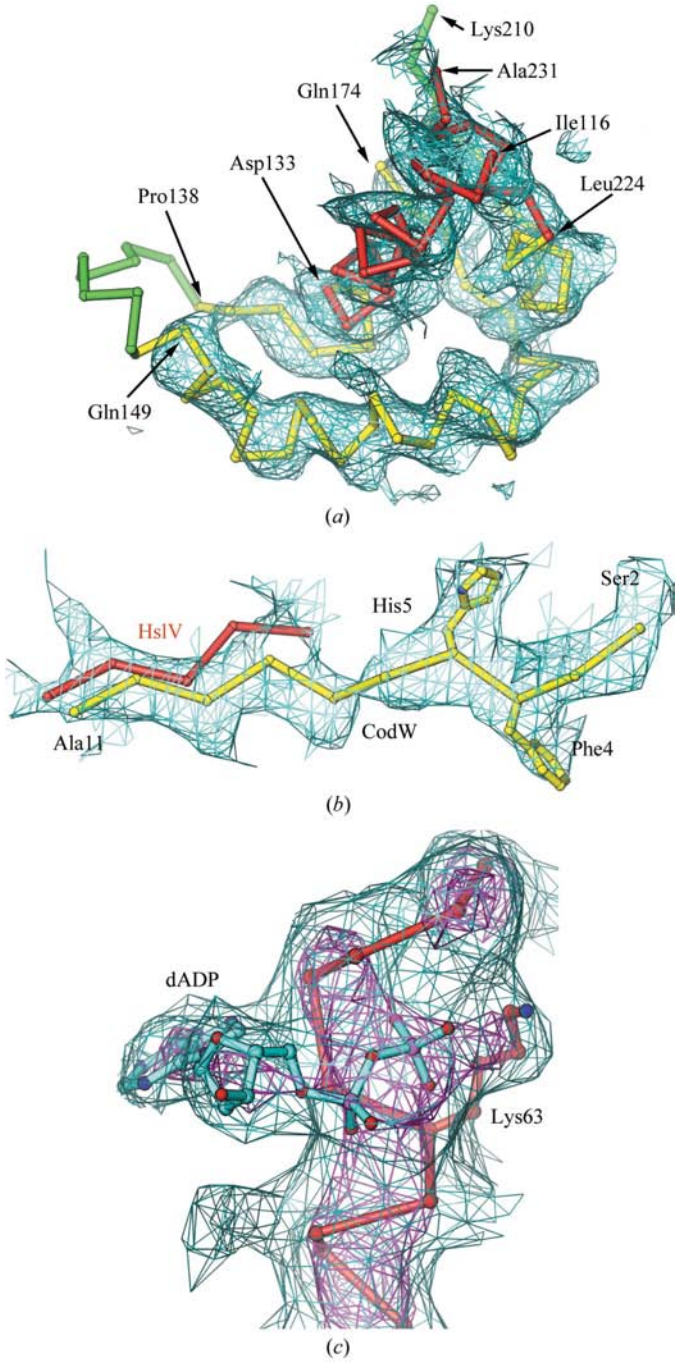

Figure 3
Unbiased electron-density maps. Electron-density maps were calculated as follows (also see text for details): (i) molecular-replacement solutions using HslV and HslU as search models from the coordinates of 1g4a (Wang, Song, Franklin *et al.*, 2001), (ii) rigid-body refinement ($R_{\rm free}$ = 45% for all data between 35 and 4.16 Å) and (iii) twofold non-crystallographic symmetry averaging with auto-masking. This model did not contain HslU domain I residues 134–223, dADP and six extra N-terminal residues of the *B. subtilis* CodW that were missing in the search model of *E. coli* HslV. Electron-density maps are contoured at 1.0σ (cyan) and 3.5σ (magenta). All models that were included in phase calculations are in red. (*a*) Portions of HslU domain I (residues 134–223) are ordered. Residues 139–148 and 210–212 (green) are visible at the 0.5σ level (not shown). (*b*) CodW N-terminal residues in a reasonable well defined β-strand are clearly visible (for example, Phe4 and His5). (*c*) The electron-density maps show that the bound nucleotide is dADP, the hydrolytic product of dATP that was included in crystallization.

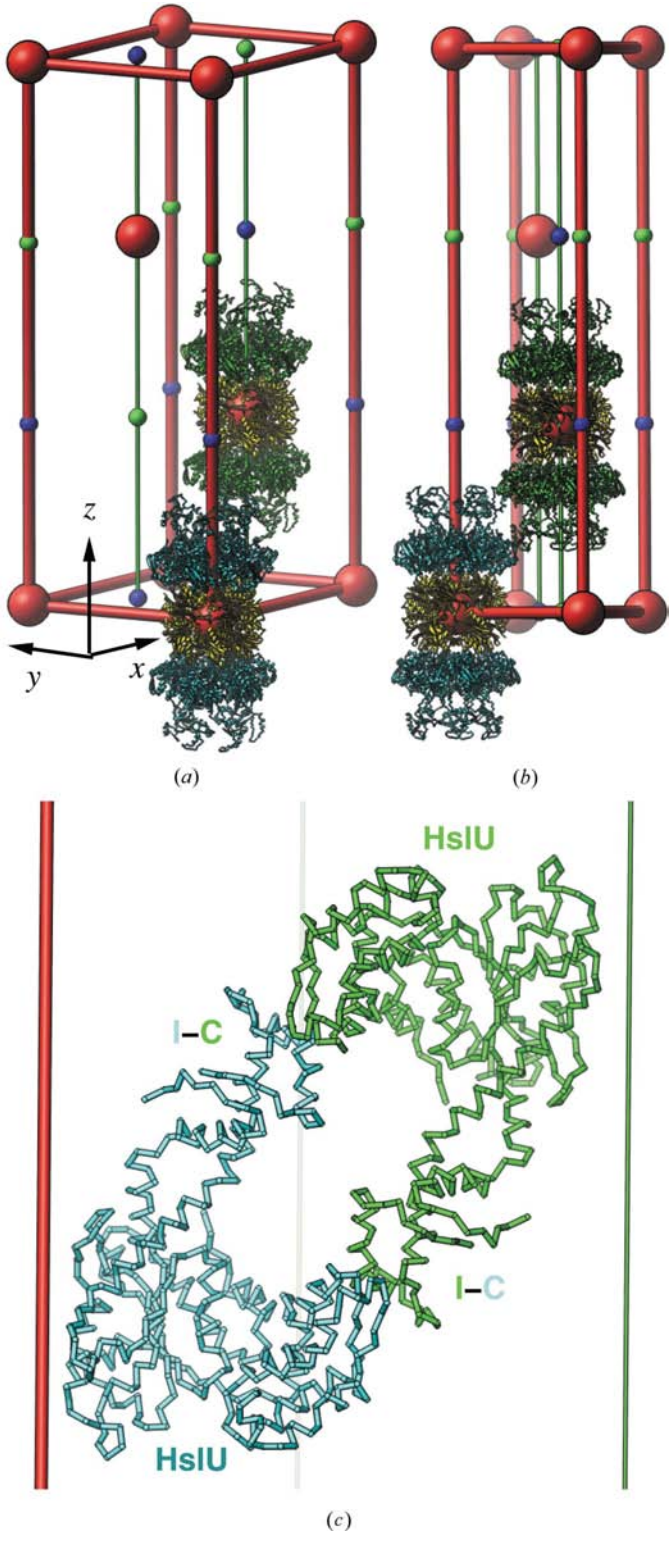

Figure 4
Packing of the HslV–HslU particles in the lattice. (*a*) CodW–HslU particles with pseudo-62 point symmetry sit on crystallographic origins (big red spheres) with *H*32 symmetry. Two subunits of each HslU and CodW are present in one asymmetric unit. Points of interest with (0, 0, 1/3) added to the crystallographic origin are shown as small blue spheres and with (0, 0, 2/3) added in small green spheres. (*b*) Packing of particles is mediated by crystallographic dyads with three HslU subunits per ring. (*c*) A close-up view of dyad-related interactions.

body refinement using the data between 10 and 4.16 Å. The *R* factor and the free *R* factor fell further to 44 and 45%, respectively, after ten-body rigid-body refinement. This crystal had an estimated solvent content of 72% with this solution.

Phase improvement was carried out using the program *RESOLVE* with twofold noncrystallographic symmetry (NCS) averaging (Terwilliger *et al.*, 2000). During the process, the amplitudes (along with their measurement errors) were sharpened by $B = -50$ Å^2 and the averaging NCS mask was autogenerated. This procedure provided an unbiased view of the parts of structures that were not included in the molecular-replacement model.

While the HslV–HslU complex in the presence of dADP was used for the molecular replacement (Wang, Song, Franklin *et al.*, 2001), the structure solution of the CodW–HslU complex was not substantially model-biased. The search model had no nucleotide and lacked parts of HslU (domain I at residues 134–223) and HslV (six amino-terminal residues of CodW). The resulting electron-density maps for these structures were completely unbiased (Fig. 3). These maps show that part of the missing domain I became ordered (Fig. 3*a*) and mediated crystallographic dyad-related particle–particle interactions in the lattice. These maps also show that the six extra amino-terminal residues at CodW were ordered. Finally, these maps show that the bound nucleotide was dADP, the hydrolytic product of dATP, not the originally added dATP. The bound dADP is also consistent with overall HslU conformation (Wang, Song, Seong *et al.*, 2001), but failed to provide an explanation as to how the ATPase activates the peptidase. Unbiased maps also show some side chains on α-helices (Fig. 3*a*) and β-strands as well as well resolved β-strands (Fig. 3*b*), suggesting that the resolution of the structure at 4.16 Å is not an overestimation.

The molecular-replacement solution was refined reasonably well. When the refined CodW structure at 2.5 Å became available (Rho *et al.*, 2005), HslV was replaced with CodW in the search model for molecular replacement, ultimately lowering the *R* factors by an additional 6%. The refinement was carried out using the *CCP*4 program *REFMAC* (Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997) with the sharpened data between 35 and 4.16 Å. The sharpened *B* factor was empirically determined when the refinement results were compared using unsharpened data and sharpened data with $B = -50$ Å^2 or -100 Å^2. With the chosen $B = -50$ Å^2, the best overall better quality of maps and refinement statistics were obtained (Table 2). After rigid-body refinement, limited positional and *B*-factor refinement with tight twofold NCS restraints (with the exception of domain I, which deviated NCS substantially) and translation–liberation–screw (TLS) refinement, the final model had an *R* factor of about 28% and a free *R* factor of 35% with reasonable geometry (Table 2). Considering the low resolution and poor quality of the data, no further refinement was attempted. The relatively large crystallographic model *R* factors in this structure are in part a consequence of the relatively large merging *R* factors for the observed data arising from the presence of the streaky reflections.

7. Packing of CodW–HslU particles in the lattice

The CodW–HslU particle sits at the crystallographic origin (0, 0, 0) with the particle's 62 point symmetry coinciding with the crystallographic 32 point symmetry (Fig. 4). Of the 12 subunits of HslU and CodW in the CodW–HslU complex, only two subunits of each protein are independent in the asymmetric unit. In the lattice, each CodW–HslU particle interacts with six nearest neighboring particles using the crystallographic dyads, relating an inter-particle HslU–HslU interaction (Fig. 4*b*). This same pattern of the inter-particle HslU–HslU interaction has also been seen in other HslU and HslV–HslU structures (Bochtler *et al.*, 2000*a*; Wang, Song, Franklin *et al.*, 2001; Wang, Song, Seong *et al.*, 2001).

No additional particles can be accommodated in this lattice. For example, when an additional particle is placed at (0, 0, 1/3) there is an about 20 Å side-by-side inter-particle penetration within an *xy* plane and 55–60 Å tail-to-tail inter-particle penetration along the *z* axis between two symmetry-related HslU hexamers. Such packing may occur only at the edge of one lattice. Indeed, it occurred at the lattice interface where overlapped hexamers are now laterally translocated to form a new lattice, resulting in the lattice-translocation defects.

8. Application of lattice-translocation defect theory to atomic models

Huber and colleagues have unexpectedly observed an extra HslV particle in their crystal lattice that was related by a native Patterson vector (Bochtler *et al.*, 2000*a*). We show below that there an extra HslV particle could also partially fitted in our crystal lattice when the intensity correction is not applied. Upon correction of the observed intensity data, the native Patterson peaks were eliminated (so was the associated extra HslV particle in the lattice) in our case. One might ask whether or not the absence of extra HslV in the lattice is caused by over-correction of the intensity data, because our procedure has one more freedom in fractions as variables than, for example, the treatment of Trame & McKay (2001).

In order to ascertain that we have not over-corrected the intensity data, we will show that the same lattice-translocation theory could be directly applied to atomic models without the need of correction for the observed data, *i.e.* the lattice-translocation defect theory is necessary and sufficient for the observed intensity modulation. Of course, this procedure, without correction of observed data, is not suitable for structure determination and is only suitable for a retrospective verification of the correct model. Using the intensity-corrected data, we have shown above one structure that corresponds to the correct quaternary structure. Using the uncorrected data, we will show below another structure that is implied to exist. Depending on whether or not an extra HslV particle in the lattice (or depending on whether or not the observed intensity data are corrected using the lattice-translocation defect theory), two different quaternary structures of the HslU–CodW complex could be obtained from one crystal. We refer to the application of the theory to the atomic model as a two-lattice model, because two overlapping structures from two lattices are composed, and we refer to the application of the theory to observed data as a one-lattice model, because the two lattices are computationally realigned to make one.

The two-lattice composite atomic model can account for the observed intensity modulation in the uncorrected data. (1) and (2) are applicable to both observed and calculated structure factors. The total calculated structure factor from the composite model, $\mathbf{F}_{\rm combined}$, is related to the structure factor $\mathbf{F}_{\rm calc}$ of a same complex structure in the translated and untranslated lattices by M and N, A and B, or f,

$$|\mathbf{F}_{\rm combined}| = |M + N\exp(2\pi i\mathbf{ht}_d)| \times |\mathbf{F}_{\rm calc}| = f^{1/2}|\mathbf{F}_{\rm calc}|. \quad (3)$$

Upon the application of the modulation factor $f^{1/2}$ to the calculated structure factors, the combined amplitudes now have an identical intensity modulation (data not shown) as the observed data. In this case, the new calculated data scaled well with the uncorrected observed data without any systematic differences in scaling factors. When the correction factor is not applied, large crystallographic R factors are observed (Fig. 5). When it is applied to either observed or calculated data, the R factors fell by as much as 9% in some resolution shells (Fig. 5). Striking similarities between two curves of R factors (Fig. 5), one upon correction of the observed data and the other upon modification of the calculated data, exclude the possibility of over-correcting the observed modulation.

While the one-lattice atomic model explain the intensity-corrected data well, it could not account for the observed

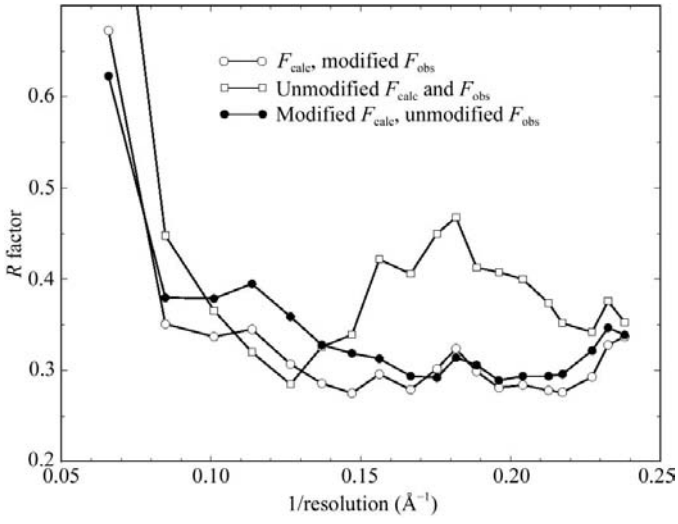

Figure 5
Crystallographic R factors as a function of resolution. Structure factors calculated from the refined atomic model ($F_{\rm calc}$) agree with the observed ones after intensity correction (modified $F_{\rm obs}$, open circles, case I), but do not agree well with the uncorrected structure factors (open squares, case II). Inverse modification of the calculated structure factors (filled circles, case III) without modifying the observations results in a similar R-factor distribution as in case I. R factors in the two lowest resolution bins were excluded in this plot owing to too few reflections. The *CCP*4 program *SCALEIT* was used in R-factor calculation (Collaborative Computational Project, Number 4, 1994). The crystallographic R factors calculated using *REFMAC* were 2–3% better than those calculated using *SCALEIT* (Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997).

intensity modulation in the uncorrected data. Layer-averaged amplitudes of calculated structure factors are flat when they are plotted against index l, because there is no translational relationship of (0, 0, 1/3) in the model. Systematic differences exist between the calculated and observed structure-factor amplitudes: calculated amplitudes are always larger than observed ones with indices $l \neq 3n$ and they are smaller than the observed amplitudes with indices $l = 3n$. These systematic differences are the source of the apparently good fitting of a one HslU/two HslV model, as described below.

9. Uncorrected data led to an incorrect quaternary HslV–HslU structure

Are the correct structural solutions obtainable when the lattice-translocation defects are not corrected? In order to answer this question, we repeated the molecular-replacement calculations and calculated $2F_o - F_c$ electron-density maps using the uncorrected data.

Using either the HslU hexamer or two HslU subunits from the hexamer as a model for molecular replacement, the translation search consistently resulted in two solutions when using the uncorrected data. One corresponded to the correct one and the second was related to the first one by the vector (0, 0, 1/3). We could exclude the second HslU solution because it could not be packed in the unit cell owing to side-to-side and tail-to-tail overlaps with both symmetry-related HslU and with the first HslU hexamer. Therefore, the correct HslU structural solution was obtainable even when using the uncorrected data.

However, the situation was different for HslV. A translational search using two subunits of HslV failed to find any meaningful solutions when using the uncorrected data. Using an HslV dodecamer, the translation search also consistently resulted in two solutions, one corresponding to the correct one and the second related to the first one by the vector (0, 0, 1/3). These solutions were much less outstanding than those of HslV using the corrected intensity data and those of HslU using either corrected or uncorrected data. When sorted in descending Patterson correlation order, the correct HslV solution was actually below the incorrect one in many resolution ranges used for the calculation. Moreover, both HslV solutions could simultaneously pack in the unit cell. The closest distance between two HslV particles in the side-by-side interactions is about 8.5 Å for side chains and 13 Å for main chains. Therefore, (i) an incorrectly positioned HslV particle could be implied to exist in the lattice and (ii) both correctly and incorrectly positioned HslV particles might be implied to coexist in the lattice when using the uncorrected data.

Ambiguities in the molecular-replacement solutions for HslV persisted even when the correctly positioned HslU particle was included. In reference to HslU, one solution of HslV corresponded to a biologically functional complex as described above and the other to a biologically irrelevant complex in which HslU domain I was used as an interface to bind HslV. The determination of the common origin between HslU and each solution of HslV showed that both biologically relevant and irrelevant complexes could exist. When HslU was combined with either HslV solution, there was no difference in crystallographic R factors even after rigid-body refinement for both solutions (both about 45%).

An unusual one HslU/two HslV model appeared to fit the uncorrected data better than the correct one HslU/one HslV model, purely owing to better numeric scaling, even though this model does not agree with the molar ratio analysis of dissolved crystals. In this model, there were some partial side-by-side HslU–HslV overlaps between two adjacent particles in the lattice. We could not exclude this model on the basis of these overlaps, because the conformation of HslU in the structure might be different from that used for molecular replacement. Moreover, the crystallographic R factors with this one HslU/two HslV model were better than the correct one HslU/one HslV model by 3%. Upon this study, we now know that the dATP that was included in the co-crystallization has hydrolyzed to dADP and the conformation of HslU in the structure is identical to that used for molecular replacement. With this HslU conformation, we could exclude the one HslU/two HslV model in retrospect. In this special case, the improved R factors (including the free R factor!) do not imply

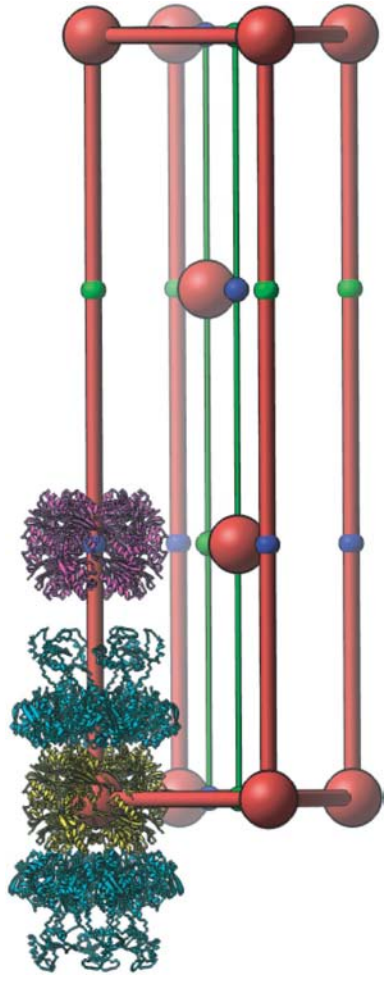

Figure 6
Incorrect model containing an extra HslV particle (magenta) at (0, 0, 1/3) fits the uncorrected observed data better than the correct model.

a better model, but instead they result from better numeric scaling between the calculated and uncorrected observed data.

Model-based Fourier and residual Fourier maps using the uncorrected intensity data consistently implied the existence of the incorrect one HslU/two HslV model in the lattice. Three types of $2F_o - F_c$ maps were calculated using the uncorrected observed amplitudes. In the first type, model phases were generated using one HslU and one HslV in the correct quaternary structure, and in the second type, model phases were from the incorrect quaternary structure. In both cases, we consistently observed the electron-density features for a second HslV that differs from the first HslV by the vector (0, 0, 1/3). Additionally, there was substantial density near each HslV surface, implying that HslV might be fixed in place by some interactions of unknown structures. The extra density could actually come from the model of the adjacent lattice domain. In the third type of map calculation, model phases were taken from the correct structure using the corrected observed amplitudes followed by extensive solvent flattening in order to ensure that there was no density at (0, 0, 1/3) in calculated phases. Even with these phases, the $2F_o - F_c$ maps using the uncorrected amplitudes continued to reveal densities for HslV at (0, 0, 1/3). To conclude, an incorrect quaternary of HslV–HslU has persistently been implied to exist in the lattice (Fig. 6) when the lattice-translocation defects were not corrected.

JW acknowledges Yale University for financial support of this work. We thank for Dr S. Kamtekar and Ms A. J. Berman for commenting on this manuscript.

References

Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000*a*). *Nature (London)*, **403**, 800–805.
Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000*b*). *Nature (London)*, **408**, 668.
Bochtler, M., Song, H. K., Hartmann, C., Ramachandran, R. & Huber, R. (2001). *J. Struct. Biol.* **135**, 281–293.
Bragg, W. L. & Howells, E. R. (1954). *Acta Cryst.* **7**, 409–411.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cochran, W. & Howells, E. R. (1954). *Acta Cryst.* **7**, 412–415.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Glauser, S. & Rossmann, M. G. (1966). *Acta Cryst.* **21**, 175–177.
Howells, E. R. & Perutz, M. F. (1954). *Proc. R. Soc. London Ser. A*, **225**, 308–314.
Ishikawa, T., Maurizi, M. R., Belnap, D. & Steven, A. C. (2000). *Nature (London)*, **408**, 667–668.
Kamtekar, S., Berman, A., Wang, J., Lazaro, J. M., deVega, M., Blanco, L., Salas, M. & Steitz, T. A. (2004). *Mol. Cell*, **16**, 609–618.
Kang, M. S., Lim, B. K., Seong, I. S., Seol, J. H., Tanahashi, N., Tanaka, K. & Chung, C. H. (2001). *EMBO J.* **20**, 734–742.
Murshudov, G. N., Vagin, A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pickersgill, R. W. (1987). *Acta Cryst.* A**43**, 502–506.
Ramachandran, R., Hartmann, C., Song, H. K., Huber, R. & Bochtler, M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 7396–7401.
Rho, S. H., Park, H. H., Lim, Y. J., Kang, M. S., Lim, B. K., Seong, I. S., Chung, C. H., Wang, J. & Eom, S. H. (2005). Submitted.
Song, H. K., Hartman, C., Ramachandran, R., Bochtler, M., Behrendt, R., Moroder, L. & Huber, R. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14103–14108.
Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S. & McKay, D. B. (2000). *Cell*, **103**, 633–643.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Trame, C. B. & McKay, D. B. (2001). *Acta Cryst.* D**57**, 1079–1090.
Wang, J. (2001). *J. Struct. Biol.* **134**, 15–24.
Wang, J. (2003). *J. Struct. Biol.* **141**, 7–8.
Wang, J., Kamtekar, S., Berman, A. J. & Steitz, T. A. (2005). *Acta Cryst.* D**61**, 67–74.
Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H., Seong, I. S., Lee, C. S., Chung, C. H. & Eom, S. H. (2001). *Structure*, **9**, 177–184.
Wang, J., Song, J. J., Seong, I. S., Franklin, M. C., Kamtekar, S., Eom, S. H. & Chung, C. H. (2001). *Structure*, **9**, 1107–1116.