Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Pathological crystallography: case studies of several unusual macromolecular crystals

Zbigniew Dauter,[a] Istvan Botos,[b] Nicole LaRonde-LeBlanc[b] and Alexander Wlodawer[b]*

[a]Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, National Cancer Institute and Biosciences Division, Argonne National Laboratory, Argonne, IL 60439, USA, and [b]Protein Structure Section, Macromolecular Crystallography Laboratory, NCI at Frederick, Frederick, MD 21702, USA

Correspondence e-mail: wlodawer@ncifcrf.gov

Received 30 January 2005
Accepted 11 April 2005

Although macromolecular crystallography is rapidly becoming largely routine owing to advances in methods of data collection, structure solution and refinement, difficult cases are still common. To remind structural biologists about the kinds of crystallographic difficulties that might be encountered, case studies of several successfully completed structure determinations that utilized less than perfect crystals are discussed here. The structure of the proteolytic domain of *Archaeoglobus fulgidus* Lon was solved with crystals that contained superimposed orthorhombic and monoclinic lattices, a case not previously described for proteins. Another hexagonal crystal form of this protein exhibited an unusually high degree of non-isomorphism. Crystals of *A. fulgidus* Rio1 kinase exhibited both pseudosymmetry and twinning. Ways of identifying the observed phenomena and approaches to solving and refining macromolecular structures when only less than perfect crystals are available are discussed here.

1. Introduction

Macromolecular crystallography is rapidly gaining the status of a routine technique rather than a separate scientific discipline, largely owing to improvements in the methods of data collection and processing (Kabsch, 1993; Leslie *et al.*, 2002; Otwinowski & Minor, 1997; Pflugrath, 1999), automated phasing of diffraction data (Brunzelle *et al.*, 2003; de La Fortelle & Bricogne, 1997; Ness *et al.*, 2004; Sheldrick, 1997; Terwilliger & Berendzen, 1999) and model building (Perrakis *et al.*, 1999; Terwilliger, 2003), as well as rapid structure refinement (Brünger *et al.*, 1998; Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997; Sheldrick & Schneider, 1997; Tronrud *et al.*, 1987). Such improvements are the key to massive efforts aimed at solving the structures of the thousands of proteins promised by various structural genomics initiatives (Burley *et al.*, 1999; Chance *et al.*, 2004; Terwilliger, 2004), with individual structures often solved within minutes of the completion of measurements of diffraction intensities. However, not all protein crystals allow straightforward approaches to structure solution, although with crystallographic methods becoming more routine and often practiced by structural biologists with only limited knowledge of the theory of the field, the reasons for occasional failures are not always obvious. Since negative results are seldom published, relatively few documented cases of phenomena that involve diffraction from unusual crystals of macromolecules are available in the scientific literature, although several papers describing the solution of crystal structures complicated by the presence of various unusual effects have been published in recent years. Among interesting and challenging cases are pseudomerohedrally twinned crys-

tals of acetyl coenzyme A synthetase with the *C*2 cell having apparent *F*222 symmetry (Lehtiö *et al.*, 2005) and tetartohedrally twinned crystals of the MltA protein (Barends *et al.*, 2005) as well as of the SRP–RNA complex (Rosendal *et al.*, 2004) with their $P3_1$ space group twinned into apparent $P6_422$ pseudosymmetry. Unusual streaking of reflection profiles combined with the hemihedral twinning relation $P3_1/P3_112$ for crystals of the ATP-binding cassette transporter (Yuan *et al.*, 2003) was caused by a specific packing of helical fibers. Crystals of SOD from *Pyrobaculum aerophilum* (Lee *et al.*, 2003) contained 24 molecules in the asymmetric unit of the $P3_2$ space group, were twinned into apparent $P3_212$ symmetry and displayed a high degree of translational pseudosymmetry. Another case of twinning combined with translational pseudosymmetry is the structure of the T-cell ligand T10 (Rudolph *et al.*, 2004). Intricate cases of translational pseudosymmetry combined with pseudomerohedral twinning exist in the structures of Mtd (Warkentin *et al.*, 2005) and α-amino acid ester hydrolase (Barends & Dijkstra, 2003). Pseudosymmetry combined with merohedral twinning has also occurred in crystals of an oligonucleotide (Abrescia & Subirana, 2002).

During ongoing efforts by our laboratory to solve a number of macromolecular structures, we have come across similar phenomena that are apparently not uncommon, yet only infrequently documented. We have also solved a structure from crystals that contained overlapping lattices of different symmetries, a case that has not been previously described for proteins. In the interest of helping other structural biologists who might be faced with similar problems in the future, we present several case studies that illustrate the difficulties in question and discuss ways of circumventing them.

2. Materials and methods

2.1. Crystals used in this study

The case studies described here involve crystals of two unrelated proteins and their mutants, with their only common feature being imperfect diffraction properties. However, all the crystals discussed below diffracted to reasonably high resolution and we will not deal here at all with the general case of limited or poor-quality diffraction.

2.1.1. Proteolytic domain of *Archaeoglobus fulgidus* Lon protease. *A. fulgidus* Lon protease is a multidomain single-chain membrane-associated enzyme consisting of an N-terminal ATPase domain followed by a C-terminal proteolytic domain. With a putative membrane-anchoring motif located in the middle of the ATPase domain, the full-length protease could not be crystallized by us using traditional methods applicable to soluble proteins. The protein used in this study corresponds to the proteolytic domain only, with the initial sequence determined by limited proteolysis and comparisons with the structure of the proteolytic domain of *Escherichia coli* Lon (Botos *et al.*, 2004). Several constructs that differed in their starting point by a few amino acids were cloned and expressed in *E. coli*. The construct that reproducibly yielded the monoclinic/orthorhombic crystals that are discussed below consisted of residues 417–621, beginning with the sequence Lys-Leu-Phe-... Because we were unable to solve the structure with these crystals, we tested other constructs and crystallization conditions and ultimately found that a construct starting with residue 415 yields a new hexagonal crystal form that diffracted to high resolution with a single molecule in the asymmetric unit. This structure, solved by single-wavelength anomalous dispersion (Dauter *et al.*, 2002) using data from a selenomethionine-containing crystal, will be described in detail elsewhere (Botos *et al.*, in preparation). Crystals of the D508A mutant of Lon appeared to be isomorphous with those of the wild-type enzyme as judged by the size and type of the hexagonal unit cell, yet they were later found to be significantly non-isomorphous, as discussed below. The coordinates corresponding to the high-resolution wild-type structure (PDB code 1z0w) were successfully utilized for a molecular-replacement solution of the monoclinic/orthorhombic structures presented here. The term 'Lon' will be used throughout this paper to denote all constructs of the proteolytic domain of the *A. fulgidus* enzyme as discussed above.

2.1.2. Complexes of *A. fulgidus* Rio1 kinase with nucleotides. Rio1 kinase is a member of a novel family of serine protein kinases. The enzyme present in *A. fulgidus* consists of a single chain of 258 amino acids and is monomeric in solution (LaRonde-LeBlanc, unpublished data). The first crystal structure for any member of this family was the recently published *A. fulgidus* Rio2 (LaRonde-LeBlanc & Wlodawer, 2004). The structure of Rio1 kinase was subsequently solved by us using monoclinic crystals containing a single molecule in the asymmetric unit. The structure was solved by multi-wavelength anomalous dispersion with a selenomethionine-containing crystal in the absence of added nucleotide (LaRonde-LeBlanc *et al.*, in preparation). However, efforts towards crystallization of complexes of the enzyme with ATP or its non-cleavable analog AMPPNP yielded a different crystal form that was quite distinct from the original one. That crystal form will be discussed here.

2.2. X-ray data collection and analysis

The X-ray diffraction data for Lon and Rio1 utilized in these studies were collected at the Advanced Photon Source insertion-device beamline 22-ID (SER-CAT, Argonne National Laboratory, Argonne, IL, USA) on a MAR225 CCD detector (MAR Research). Data for the hexagonal crystals of the D508A mutant of Lon were collected on a MAR345 detector using a Rigaku H3R rotating-anode X-ray source operated at 50 kV and 100 mA with Cu $K\alpha$ radiation focused by an MSC/Osmic mirror system. All data were processed and scaled with *HKL*2000 (Otwinowski & Minor, 1997). The statistics of data processing are shown in Tables 1 and 2 for crystals of Lon and Rio1, respectively.

2.3. Structure solution and refinement

All structures described here were solved by molecular replacement using standard procedures. The structures of

Table 1
Data-collection statistics for Lon protease.

Values in parentheses relate to the highest resolution shell (∼5% of data).

	Crystal 9†		Crystal 4	Crystal 5
	Orthorhombic	Monoclinic	Orthorhombic	Monoclinic
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P2_1$
Molecules per AU	6	6	6	6
Unit-cell parameters (Å, °)	a = 86.25, b = 90.55, c = 147.95	a = 48.45, b = 86.28, c = 137.97, β = 92.3	a = 86.65, b = 88.69, c = 147.24	a = 48.68, b = 86.11, c = 135.61, β = 94.7
Resolution (Å)	50–3.0	50–3.0	20–2.3	50–2.05
Total reflections	159505	68167	180696	297869
Unique reflections	23841	22682	50449	68204
Completeness (%)	99.8 (100.0)	98.9 (98.6)	97.1 (95.7)	96.5 (76.8)
$\langle I\rangle/\langle\sigma(I)\rangle$	13.2 (6.7)	11.7 (5.7)	9.1 (2.2)	10.4 (1.3)
$R_{\rm merge}$‡ (%)	13.5 (25.4)	10.1 (19.8)	12.9 (48.2)	13.2 (58.2)

† Data from both lattices were processed for crystal 9, whereas only individual orthorhombic and monoclinic data were fully processed for crystals 4 and 5, respectively, although both of these crystals also exhibited some degree of twinning. ‡ $R_{\rm merge} = \sum_h \sum_i |I_i - \langle I\rangle| / \sum_h \sum_i I_i$, where I_i is the observed intensity of the ith measurement of reflection h and $\langle I\rangle$ is the average intensity of that reflection obtained from multiple observations.

Table 2
Data-collection statistics for Rio1 kinase.

Values in parentheses relate to the highest resolution shell (2.02–1.95 Å for crystal 4 and 2.07–2.00 Å for crystal 3).

	Crystal 4		Crystal 3	
	Orthorhombic	Monoclinic	Orthorhombic	Monoclinic
Space group	$P22_12$	$P2_1$	$P22_12$	$P2_1$
Molecules per AU	2	4	2	4
Unit-cell parameters (Å, °)	a = 53.27, b = 80.32, c = 120.88	a = 53.27, b = 80.31, c = 120.88, β = 89.98	a = 53.38, b = 80.30, c = 121.04	a = 53.33, b = 80.29, c = 121.02, β = 89.99
Resolution (Å)	30–1.95	30–1.95	50–2.0	50–2.0
Total reflections	251949	250698	278496	278641
Unique reflections	37214	68759	35941	67986
Completeness (%)	96.3 (84.6)	92.3 (75.0)	99.9 (99.8)	98.1 (96.1)
$\langle I\rangle/\langle\sigma(I)\rangle$	12.6 (2.6)	13.9 (2.4)	16.5 (5.2)	13.4 (3.7)
$R_{\rm merge}$† (%)	22.9 (47.8)	8.9 (38.9)	17.3 (38.9)	9.9 (32.3)
Refined twin fraction		0.845/0.155		0.65/0.35

† The definition of $R_{\rm merge}$ is given in the footnote to Table 1.

Rio1 kinase, the monoclinic variant of wild-type Lon and the D508A mutant of Lon were solved using the program *AMoRe* (Navaza, 1994), whereas the structure of the orthorhombic variant of Lon was solved with *EPMR* (Kissinger *et al.*, 1999). The structures were refined with *REFMAC*5 (Murshudov *et al.*, 1997), *CNS* (Brünger *et al.*, 1998) or *SHELXL* (Sheldrick & Schneider, 1997).

3. Results

3.1. A twinning phenomenon resulting in a single crystal with two different lattices

The most unusual phenomenon to be reported here involves the crystals of Lon protease. These crystals could be grown quite reproducibly and individual crystals appeared by visual inspection to be single. Different crystals provided diffraction patterns extending to 2.05–3 Å and data could be indexed automatically without apparent difficulty. However, some crystals appeared to be orthorhombic, whereas others indexed in a smaller monoclinic lattice. In both cases, prediction of the diffraction patterns for the identified cell resulted in many spots that did not follow the metrics of the chosen lattice. However, when only the non-predicted spots were utilized in the second pass of indexing, they could be indexed unambiguously using the alternate lattice. An example of such a diffraction pattern and its prediction using both lattices is shown in Fig. 1. It is clear that both lattice types are present simultaneously in the crystal, leading to partially overlapping diffraction patterns. The relative intensity of the monoclinic and orthorhombic patterns differs between different crystals, but all crystals grown under these conditions showed the presence of this phenomenon to a varying extent.

Diffraction images from one of the crystals in which both types of lattices were present in a roughly 1:1 ratio (crystal 9) were processed to provide complete data sets for both space groups. The monoclinic data could be indexed in space group $P2_1$, with unit-cell parameters a = 48.45, b = 86.28, c = 137.97 Å, β = 92.3°. The orthorhombic lattice from the same crystal could be indexed in space group $P2_12_12_1$, with unit-cell parameters a = 86.25, b = 90.55, c = 147.95 Å. Both data sets were processed to a resolution of 3 Å, with scaling R factors of 0.101

for the monoclinic set and 0.135 for the orthorhombic set (Table 1). The mosaicity parameters for both lattices varied smoothly between 0.8 and 1.2° in different parts of reciprocal space. It is clear that these *R* factors are much higher than would be expected from normal crystals exhibiting such reasonably strong diffraction, yet they were low enough that an attempt to solve the structure by molecular replacement could possibly succeed.

The structure of this crystal form of Lon was initially solved using another crystal (crystal 4) in which the orthorhombic lattice predominated and for which the monoclinic data, although processable, were incomplete owing to the way in which the crystal was oriented (~60% overall completeness). The orthorhombic pattern was processed to a resolution of 2.3 Å; the unit-cell parameters were a = 86.65, b = 88.69, c = 147.24 Å and the space group was $P2_12_12_1$. The merging R factor was 0.129, with $\langle I\rangle/\langle\sigma(I)\rangle$ = 9.3 for all data (Table 1). With the volume of the asymmetric unit being 280 000 Å^3 and the molecular weight of Lon being 21 815 Da, the likely number of molecules in the asymmetric unit was estimated to be between four (V_M = 3.20 Å^3 Da^{-1}) and six (V_M = 2.14 Å^3 Da^{-1}) (Matthews, 1968). The structure of *A. fulgidus* Lon was previously solved by us in a hexagonal crystal form with one molecule in the asymmetric unit (Botos *et al.*, manuscript in preparation), so all these oligomeric states had to be considered.

Our initial attempts to solve the structure with *AMoRe* (Navaza, 1994) were unsuccessful, with no clear solutions appearing after a series of runs utilizing different resolution and Patterson integration ranges. However, results obtained with *EPMR* (Kissinger *et al.*, 1999) were more encouraging. One molecule of Lon was used as a model. While no apparent solutions were identifiable for a single-molecule search (best correlation coefficient 0.137 and R factor 0.716 for resolution range 12.5–4.5 Å), the best result after searching for two molecules had a correlation coefficient 0.488 and R = 0.576. These parameters improved further for a solution consisting of four molecules (correlation coefficient 0.650, R = 0.489). However, the resulting structure still exhibited significant open areas in the asymmetric unit, indicating the possible presence of more molecules. The final run searching for six molecules resulted in a correlation coefficient of 0.733 and R = 0.432, with very good packing in the unit cell.

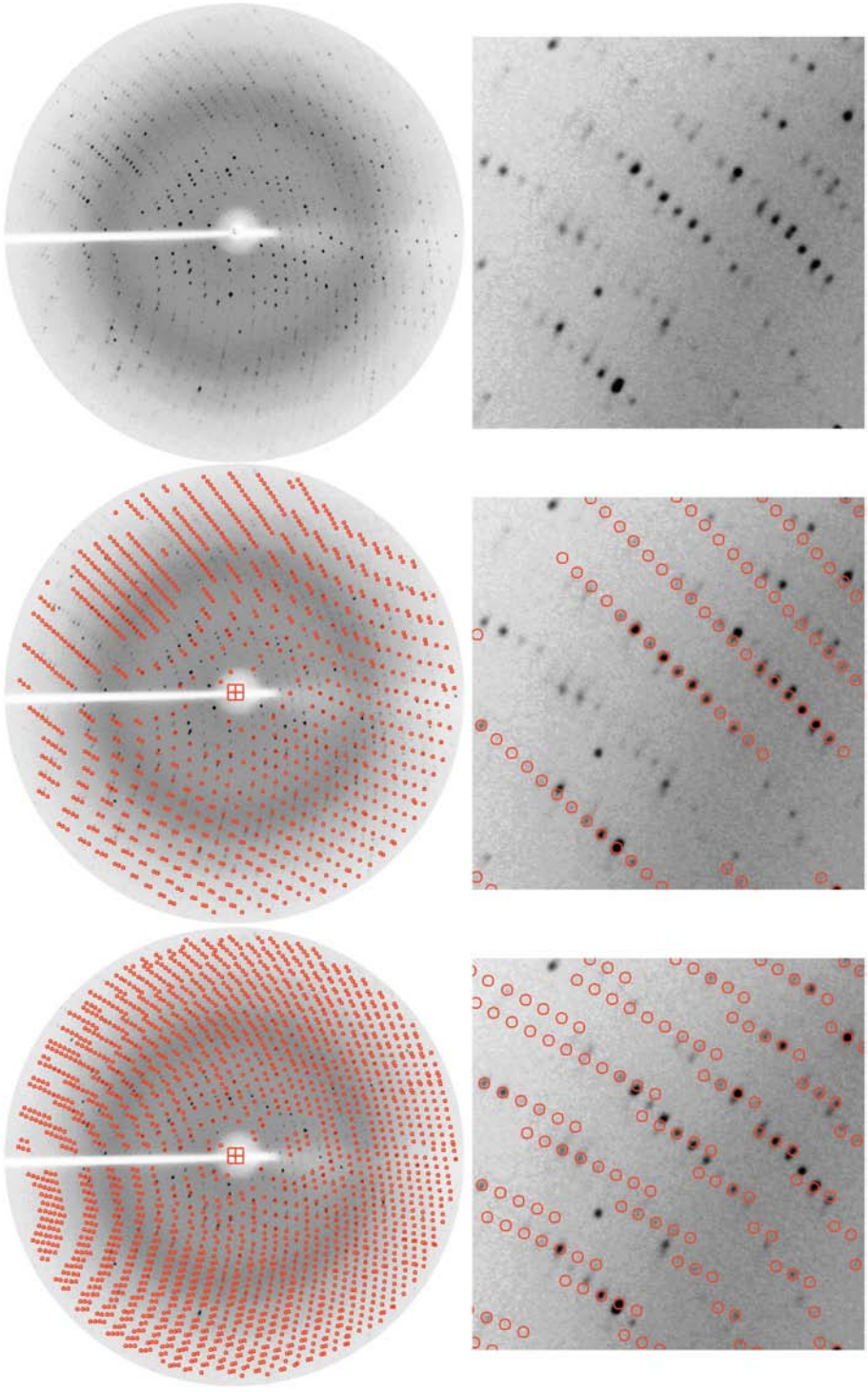

Figure 1
X-ray diffraction pattern from a crystal of the proteolytic domain of *A. fulgidus* Lon. The resolution on the outer edge of the image is 3.0 Å. The upper left panel shows a single 0.5° oscillation frame, the middle panel shows the interpretation of the pattern in the monoclinic cell and the lower panel illustrates the interpretation according to the orthorhombic cell. The panels on the right show zoomed-in portions of the corresponding diffraction patterns.

The six molecules in the asymmetric unit are arranged as a non-crystallographic hexamer with the sixfold axis parallel to the *b* axis of the unit cell (Fig. 2*a*). The orthorhombic cell contains two such hexameric layers, with their axes offset by a translation of 20 Å along the *c* axis; that is, 0.135 of its length. Owing to the symmetry of the unit cell, the third parallel layer

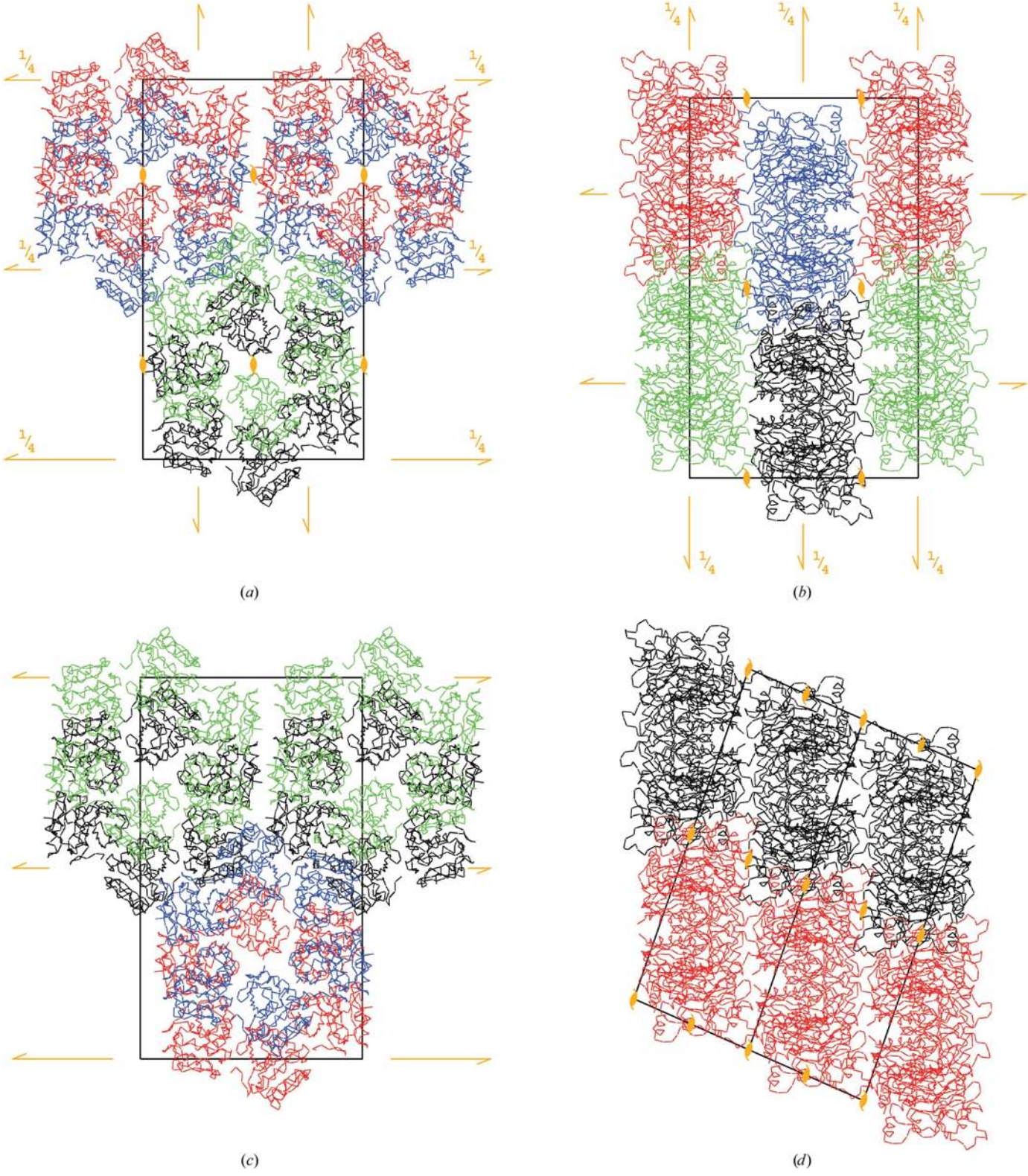

Figure 2
Packing diagrams for the crystals of *A. fulgidus* Lon. Orthorhombic cell projected (*a*) onto the *bc* plane and (*b*) onto the *ac* plane. Monoclinic cell projected (*c*) onto the plane containing the *b* axis and the longer diagonal between the *a* and *c* axes and (*d*) onto the *ac* plane. Two cells and six asymmetric units are shown.

Table 3
Refinement statistics for Lon protease.

	Crystal 9		Crystal 4	Crystal 5
	Orthorhombic	Monoclinic	Orthorhombic	Monoclinic
R† (%)	19.0 (21.3)	21.4 (24.8)	19.5 (21.3)	21.3 (27.0)
R_{free}‡ (%)	35.2 (42.8)	31.0 (35.2)	32.8 (37.7)	29.9 (36.7)
R.m.s.d. bond lengths (Å)	0.073	0.041	0.045	0.033
R.m.s.d. angles (°)	5.29	3.28	3.24	2.49
B factor, protein (Å^2)	26.87	18.38	12.75	37.76
B factor, solvent (Å^2)	34.55	12.14	24.19	44.44
No. of protein atoms	8778	8742	8778	8742
No. of solvent molecules	1092	503	1095	552
PDB code	1z0t	1z0v	1z0g	1z0e

† $R = \big||F_o| - |F_c|\big|/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively, calculated for all data. ‡ R_{free} as defined in Brünger (1992).

is identical to the first one and thus this crystal form can be described as a series of layers placed in a zigzag fashion and with the 2_1 axis parallel to the b direction connecting hexamers in every other layer (Fig. 2*b*). Since the sixfold local symmetry of the Lon hexamer also includes the twofold rotation, the relation between a pair of hexamers from two layers in effect corresponds to pure translation. Indeed, the native Patterson map shows a very prominent peak at 0.0, 0.5, 0.135, with a height equal to 50% of the origin peak.

Another crystal (crystal 5) provided monoclinic data extending to 2.05 Å resolution. The unit-cell parameters were $a = 48.68$, $b = 86.11$, $c = 135.61$ Å, $\beta = 94.7^\circ$ (space group $P2_1$). Diffraction data were 96.5% complete, the scaling R factor was 0.132 and $\langle I\rangle/\langle\sigma(I)\rangle$ was 12.6 for the whole data set (Table 1). The structure was solved starting from the same starting point as the orthorhombic structure, namely from the coordinates of a single molecule from hexagonal crystals. In the monoclinic case, *AMoRe* provided an unambiguous solution for six molecules in the asymmetric unit. Although the steps taken to obtain solutions of the two crystal forms were different and independent of each other, the resulting hexamer of Lon is virtually identical, while the packing is very similar. The two adjacent layers of hexamers (in two neighboring unit cells) are arranged in a manner identical to the orthorhombic crystal form (Fig. 2*c*). Indeed, when only two layers are compared, the packing is the same in both the monoclinic and orthorhombic lattices. The difference is only seen in the next layer (Fig. 2*d*), which is offset from the previous one in exactly the same manner as layer 2 from layer 1, rather than corresponding to layer 1 in a zigzag fashion characteristic for the orthorhombic lattice. As a consequence, the monoclinic cell is twice smaller than the orthorhombic one, although in both the asymmetric units contain the complete hexamer of Lon.

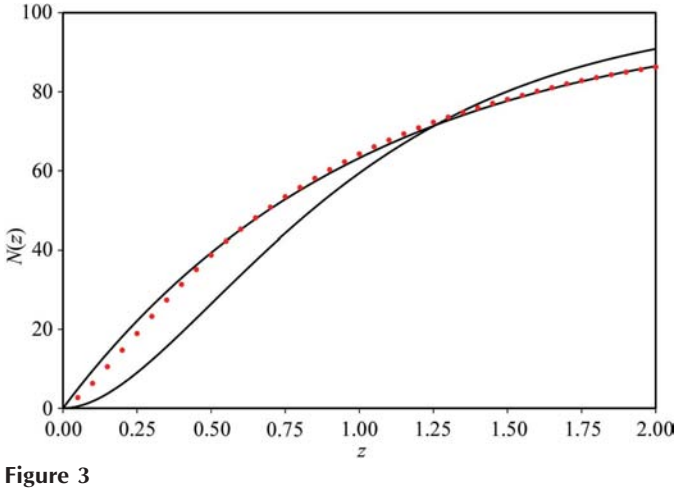

Figure 3
Cumulative intensity distribution for the crystal of Rio1 kinase calculated with the *TRUNCATE* module of *CCP*4. The distribution appears to be close to normal and does not indicate the significant degree of twinning.

Since the structures corresponding to the two crystal systems were initially solved from data collected from separate crystals, both data sets obtained from the same specimen (crystal 9) were refined independently, starting from the solutions described above. Both 3 Å data sets refined reasonably well, to R factors of 19.0 and 21.4 (Table 3) for the orthorhombic and monoclinic cell, respectively. The results of scaling and refinement are obviously not as good as would be expected for crystals that did not show the imperfections that were present in this case, yet they indicate that the overall structures were solved correctly.

The relationship between the two lattices became quite clear once the structures were solved. Not surprisingly, their unit-cell parameters bear close relationship, with the monoclinic b axis and the orthorhombic a axis having the same length within the error of measurement and the longer ac diagonal of the monoclinic cell having the same length as the orthorhombic c axis. It can be assumed that the crystals grow as layers of hexagons. If all layers maintain a constant offset they create monoclinic crystals, but if they alternate the crystals become orthorhombic. It is reasonable to postulate that such switching could be accomplished many times within a single crystal, resulting in the unusual twinning phenomenon observed by us, especially since the contacts between layers are not very numerous (Figs. 2*b* and 2*d*). There is some resemblance of this situation to the structure of graphite (Bernal, 1924), with its known property as a lubricant resulting from the possibility of sliding between neighboring layers of conjugated systems of hexagonally arranged C atoms.

Whereas most individual reflections resulting from the two lattices do not overlap, some do (Fig. 1), and since the standard data-processing software is not capable of extracting individual intensities from overlapped reflection profiles, this

Table 4
Statistics of data intensities for Rio1 kinase (acentric reflections only).

		Theoretical	
	Observed	Untwinned	50% twinned
Wilson ratios			
$\langle I^2\rangle/\langle I\rangle^2$	2.060	2.0	1.5
$\langle F\rangle^2/\langle F^2\rangle$	0.784	0.785	0.885
Yeates statistics†			
$\langle H\rangle$	0.164	0.5	0.0
$\langle H^2\rangle$	0.048	0.333	0.0
Padilla & Yeates‡ statistics			
$\langle\|L\|\rangle$	0.442	0.5	0.375
$\langle L^2\rangle$	0.268	0.333	0.2

† According to Yeates (1997). ‡ According to Padilla & Yeates (2003).

leads to systematic errors in their determination and to comparatively high scaling *R* factors. In principle, once both structures have been solved it should be possible to reprocess the data using the experimentally determined 'twinning ratio' or alternatively to refine the model jointly using both data sets, along the lines used previously in joint X-ray/neutron refinement (Wlodawer & Hendrickson, 1982). However, the extent of effort necessary to accomplish such a task was not justified in this particular case, although development of such software might become possible sometime in the future.

3.2. A crystal that is both pseudo-symmetric and twinned

The crystals of Rio1 kinase could be easily indexed in an orthorhombic cell, with unit-cell parameters a = 53.27, b = 80.32, c = 120.88 Å (Table 2). However, scaling these data as orthorhombic proved to be highly problematic, with a scaling *R* factor of 22.9%. Attempts to scale these data in the monoclinic system, assuming each axis in turn to be the unique b axis, resulted in acceptable scaling statistics (R = 8.9%) in the case of b = 80.32 Å and the monoclinic angle β refining to 89.98°. Similar results were obtained for other crystals as well (Table 2). Although not very common, monoclinic cells with almost right angles have been reported in the past. The Wilson ratios and the $N(z)$ cumulative intensity statistics (Fig. 3, Table 4) did not suggest a significant degree of twinning. The application of a more recently introduced approach (Padilla & Yeates, 2003) based on the local intensity differences showed that the crystal could be twinned by about 8%. However, the *H*-test (Yeates, 1997), using only the subset of relatively strong reflections, indicated a substantial degree of twinning of above 30%.

The structure of Rio1 that is discussed here was solved by molecular replacement with *AMoRe* using a dimer of molecules from the previously solved monoclinic structure, with data in the range 10–3.5 Å. Placement of two dimers in the monoclinic cell gave a correlation coefficient of 0.37 and an *R* factor of 51.2% after rigid-body refinement. The packing of four molecules in the cell is quite regular, approximating the constellation of the orthorhombic space group $P2_12_12_1$. The content of two monoclinic unit cells is illustrated in Fig. 4, with each set of crystallographically independent molecules represented in a different color. The molecules are arranged in layers perpendicular to the b axis and related by the crystallographic twofold screw axes. Two such layers are drawn in the figure, the lower one with thicker lines and pale colors and the upper one with thinner lines and darker colors. The molecules within each layer are to a good approximation related by the screw axes parallel to a and c directions and there is an additional set of approximate 2_1 axes parallel to b in between of the crystallographic axes. This pseudo-orthorhombic symmetry elements and the corresponding cell are marked in Fig. 4 in red. The pseudo-orthorhombic cell is half as large as the proper monoclinic cell and contains only one molecule in the asymmetric unit.

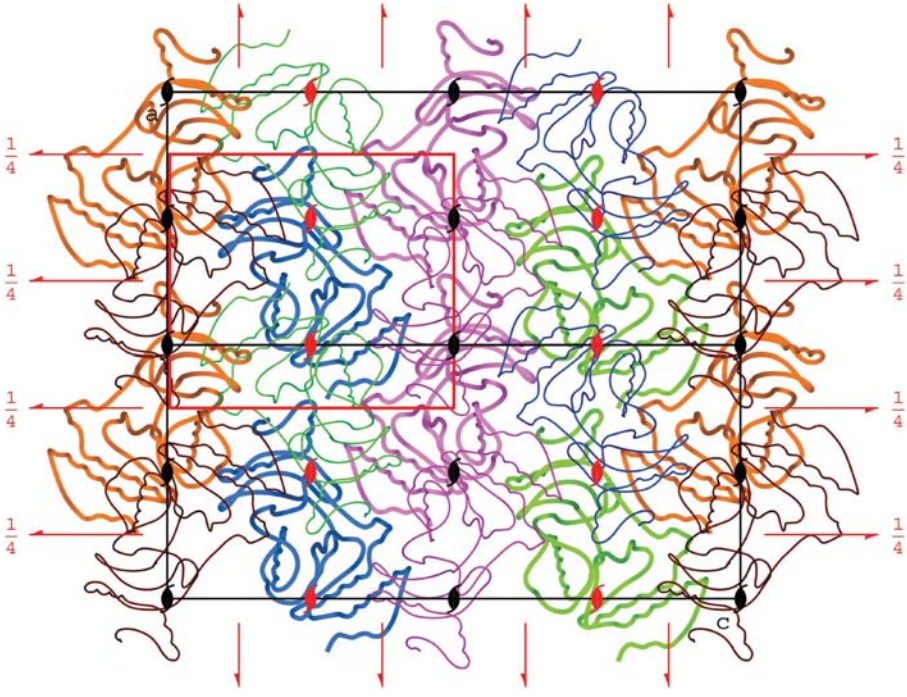

Figure 4
Schematic packing diagram for the crystal of *A. fulgidus* Rio1 kinase (PDB code 1zp9). The four crystallographically independent molecules present in the asymmetric unit of the monoclinic cell are colored differently. The molecules related by the crystallographic 2_1 symmetry axes, packed in the lower layer, are colored similarly, but in lighter shades. The pseudosymmetric orthorhombic cell with corresponding approximate twofold screw axes is shown in red.

The presence of an additional pseudo-orthorhombic cell with one of the dimensions half the size of the original monoclinic cell means that molecules are packed in a highly parallel fashion. The native Patterson shows a significant peak at 0, 0, 0.48 (Fig. 5) with a height of 11σ (6% of the origin peak). The value of the twinning factor refined by *SHELXL* is 0.35, but

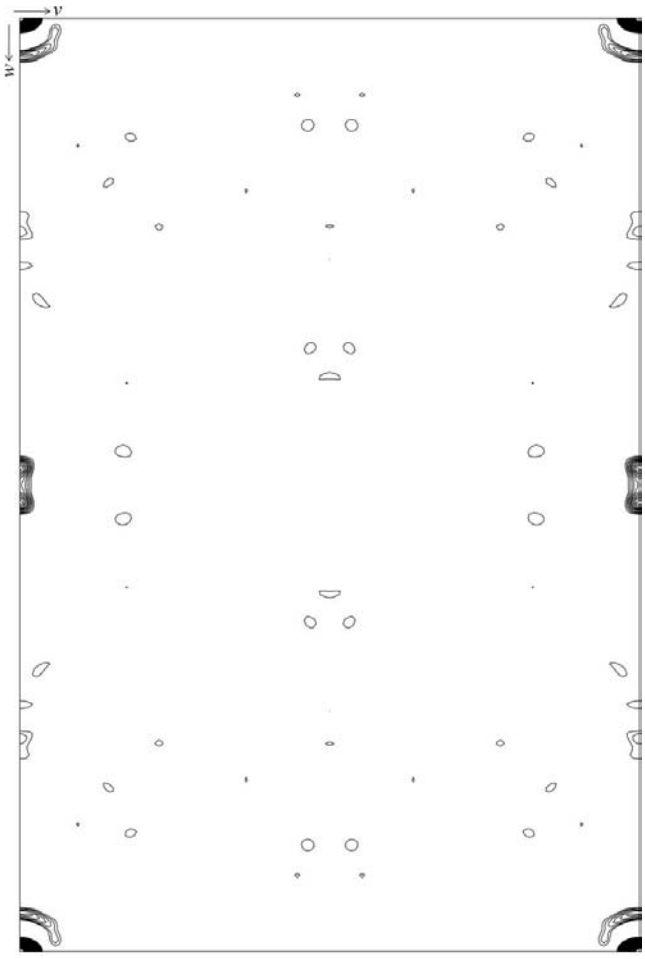

Figure 5
The $v = 0$ Harker section of the native Patterson synthesis for Rio1 kinase, contoured at every 1σ starting with 2σ. The prominent peak with coordinates 0, 0, 0.48 corresponds to the translation vector between pairs of Rio1 molecules arranged in an almost parallel fashion within the asymmetric unit of the crystal cell.

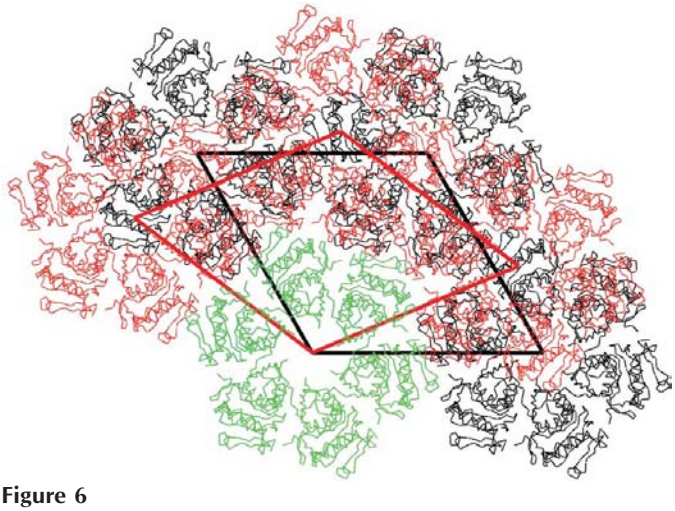

Figure 6
Packing of the wild-type Lon protease (in black) and its D508A mutant (in red) in their respective unit cells. The constellation of molecules around the individual 6_5 helical axis overlaps perfectly (in green) after rotation by $\sim$23°. In contrast, interactions between the neighboring (poly)hexamers are completely different.

the statistics of intensities did not reveal such a high degree of twinning. This is understandable in the light of the presence of the non-crystallographic parallel translation within the structure. The effect of twinning on the statistics of intensities is manifested by lower fractions of very weak and very strong reflections. In contrast, the presence of non-crystallographic translations causes some class of reflections to be weaker than expected. If a crystal is twinned and translationally pseudo-symmetric, both tendencies cancel out and the resulting intensity statistics may look almost normal, as in the case of Rio1.

3.3. Unexpected non-isomorphism

The crystals of wild-type Lon protease that were utilized in the successful structure determination of this enzyme by SAD were grown in the hexagonal space group $P6_5$, with one molecule in the asymmetric unit and unit-cell parameters $a = b = 83.74$, $c = 41.23$ Å. It is not surprising that when we subsequently grew crystals of the D508A mutant of Lon with unit-cell parameters $a = b = 81.77$, $c = 41.58$ Å and the same apparent space group, we assumed that these crystals must be almost isomorphous (the difference in the length of the a/b axes is only 2.3%). However, we were unable to scale the wild-type and mutant data assuming either of the two possible orientations of the polar axis. When the wild-type structure was utilized as a molecular-replacement (*AMoRe*) search model against the mutant data and refined at 1.55 Å resolution (PDB code 1z0c), we found that the molecules were rotated by $\sim$23° around the sixfold screw axis, explaining the non-isomorphism (Fig. 6). This rotation did not change the intermolecular interactions owing to the operation of this axis, yet significantly changed the interactions away from that axis. Since the presence of a conserved pseudo-hexamer was considered to be an important indication of the relevance of the quaternary structure of Lon (Botos *et al.*, 2004), conservation of such assemblies despite variation in the crystal contacts is an important result. This is a clear example showing that isometry of the unit cell is not necessarily an indication of true isomorphism.

4. Discussion

The crystals discussed here presented us with different problems and difficulties and the question might be raised whether the effort required to solve these structures was justified in the first place or whether it would not have been more prudent to simply abandon these problematic crystals and search for conditions that would yield different forms. To a certain extent that is what we have been doing, but crystallographers do not always have the luxury of following that approach. The crystals of Lon protease discussed above were the first ones to be grown and we attempted to solve the structure both by molecular replacement using the available coordinates of *E. coli* Lon (Botos *et al.*, 2004) (unsuccessfully, not surprisingly in view of significant differences between these enzymes, as will be discussed elsewhere) or *de novo*

using selenomethionine-labeled protein. The latter approach was abandoned when non-twinned hexagonal crystals became available and the structure was solved in a straightforward manner. Nevertheless, some information gleaned during the analysis of the orthorhombic/monoclinic crystals was extremely valuable. These crystals contained hexamers similar to those of the previously described hexameric assemblies of *E. coli* Lon, providing very strong evidence that the quaternary structures of both enzymes are very similar. This type of information could not be obtained through analysis of the hexagonal crystals, since they did not contain closed hexamers but rather infinite helices created by shifting each molecule within a hexamer by ∼6.5 Å. Another valuable piece of information obtained in the study described here is the confirmation of the persistently inactive conformation of the active site. This phenomenon was first noted in the high-resolution hexagonal structure and it was important to ascertain whether it was an isolated case or reproducible. Since the conformation of the active site has now been shown to be very similar in the three crystal forms, we must look for its explanation beyond postulating a simple crystallographic artifact related to growing crystals under a particular single set of conditions. We thus feel that the effort expanded on solving these structures was completely justified.

All available crystals of the nucleotide complexes of Rio1 kinase were similar to those described above and in this case we did not have the option of refining the structures using any other crystal forms. However, the degree of twinning did differ between various crystals and the crystal used in the final refinement was slightly less twinned than the crystal discussed here. However, even the best crystals showed some degree of twinning, whereas pseudosymmetry was obviously present in all of them.

Part of the diffraction data used in this study was collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID, located at the Advanced Photon Source, Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38.

References

Abrescia, N. G. A. & Subirana, J. A. (2002). *Acta Cryst.* D**58**, 2205–2208.
Barends, T. R. M., de Jong, R. M., van Straaten, K. E., Thunnisen, A.-M. W. H. & Dijkstra, B. W. (2005). *Acta Cryst.* D**61**, 613–621.
Barends, T. R. M. & Dijkstra, B. W. (2003). *Acta Cryst.* D**59**, 2237–2241.
Bernal, J. D. (1924). *Proc. R. Soc. London, Ser. A*, **106**, 749–773.
Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. & Gustchina, A. (2004). *J. Biol. Chem.* **279**, 8140–8148.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
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.
Brunzelle, J. S., Shafaee, P., Yang, X., Weigand, S., Ren, Z. & Anderson, W. F. (2003). *Acta Cryst.* D**59**, 1138–1144.
Burley, S. K., Almo, S. C., Bonanno, J. B., Capel, M., Chance, M. R., Gaasterland, T., Lin, D., Sali, A., Studier, F. W. & Swaminathan, S. (1999). *Nature Genet.* **23**, 151–157.
Chance, M. R., Fiser, A., Sali, A., Pieper, U., Eswar, N., Xu, G., Fajardo, J. E., Radhakannan, T. & Marinkovic, N. (2004). *Genome Res.* **14**, 2145–2154.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z., Dauter, M. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
LaRonde-LeBlanc, N. & Wlodawer, A. (2004). *Structure*, **12**, 1585–1594.
Lee, S., Sawaya, M. R. & Eisenberg, D. (2003). *Acta Cryst.* D**59**, 2191–2199.
Lehtiö, L., Fabrichniy, I., Hansen, T., Schönheit, P. & Goldman, A. (2005). *Acta Cryst.* D**61**, 350–354.
Leslie, A. G., Powell, H. R., Winter, G., Svensson, O., Spruce, D., McSweeney, S., Love, D., Kinder, S., Duke, E. & Nave, C. (2002). *Acta Cryst.* D**58**, 1924–1928.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Ness, S. R., de Graaff, R. A., Abrahams, J. P. & Pannu, N. S. (2004). *Structure*, **12**, 1753–1761.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Rosendal, K. R., Sinning, I. & Wild, K. (2004). *Acta Cryst.* D**60**, 140–143.
Rudolph, M. G., Wingren, C., Crowley, M. P., Chien, Y. H. & Wilson, I. A. (2004). *Acta Cryst.* D**60**, 656–664.
Sheldrick, G. M. (1997). *Methods Enzymol.* **276**, 628–641.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Terwilliger, T. C. (2003). *Methods Enzymol.* **374**, 22–37.
Terwilliger, T. C. (2004). *Nature Struct. Mol. Biol.* **11**, 296–297.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.
Warkentin, E., Hagemeier, C. H. S. S., Thauer, R. K. & Ermler, U. (2005). *Acta Cryst.* D**61**, 198–202.
Wlodawer, A. & Hendrickson, W. A. (1982). *Acta Cryst.* A**38**, 239–247.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.
Yuan, Y.-R., Martsinkevitch, O. & Hunt, J. F. (2003). *Acta Cryst.* D**59**, 225–238.