Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Ultrahigh-resolution study of protein atomic displacement parameters at cryotemperatures obtained with a helium cryostat

Tatiana Petrova,[a,b,c] Stephan Ginell,[a] Andre Mitschler,[c] Isabelle Hazemann,[c] Thomas Schneider,[d] Alexandra Cousido,[c] Vladimir Y. Lunin,[b] Andrzej Joachimiak[a]* and Alberto Podjarny[c]

[a]Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, USA, [b]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino 142290, Russia, [c]Laboratoire de Génomique et de Biologie Structurales, IGBMC, CNRS–ULP–INSERM, 1 Rue Laurent Fries, BP 163, 67404 Illkirch, France, and [d]Institute of Molecular Oncology, Milan, Italy

Correspondence e-mail: andrzejj@anl.gov

Received 30 March 2006
Accepted 4 October 2006

Two X-ray data sets for a complex of human aldose reductase (h-AR) with the inhibitor IDD 594 and the cofactor $NADP^+$ were collected from two different parts of the same crystal to a resolution of 0.81 Å at 15 and 60 K using cold helium gas as cryogen. The contribution of temperature to the atomic B values was estimated by comparison of the independently refined models. It was found that although being slightly different for different kinds of atoms, the differences (δB) in the isotropic equivalents B of atomic displacement parameters (ADPs) were approximately constant (about 1.7 $Å^2$) for well ordered atoms as the temperature was increased from 15 to 60 K. The mean value of this difference varied according to the number of non-H atoms covalently bound to the parent atom. Atoms having a B value of higher than 8 $Å^2$ at 15 K showed much larger deviations of δB from the average value, which might reflect partial occupancy of atomic sites. An analysis of the anisotropy of ADPs for individual atoms revealed an increase in the isotropy of ADPs with the increase of the temperature from 15 to 60 K. In a separate experiment, a 0.93 Å resolution data set was collected from a different crystal of the same complex at 100 K using cold nitrogen as a cryogen. The effects of various errors on the atomic B values were estimated by comparison of the refined models and the temperature-dependent component was inferred. It was found that both decreasing the data redundancy and increasing the resolution cutoff led to an approximately constant increase in atomic B values for well ordered atoms.

1. Introduction

Atomic displacement parameters (ADPs) provide essential information on the dynamic and static disorder of a crystal sample and are determined by X-ray crystallography during the course of structure refinement. In the isotropic approximation they are represented by only one parameter (B value) per atom, while in the more general anisotropic case they are represented by a 3×3 symmetric tensor, *i.e.* they provide six parameters per atom. A set of ADPs is usually attributed to the time- and space-averaged deviation of an atomic position from its mean value (Trueblood *et al.*, 1996), while the ADP values derived from the refinement program are actually cumulative parameters which also include experimental and computational errors.

A set of ADP values provides no indication about the qualitative and quantitative influence of the different factors

involved. Extracting the contribution of temperature should give a quantitative measure of atomic motion (both atomic vibrations and a variety of collective motions). Evaluation of the part related to static disorder can provide an insight into the order in macromolecular crystals. Thus, partitioning of the different factors that contribute to the ADP values would make it possible to obtain more detailed information about the structure. In our study, we make an attempt to separate different contributions to the ADPs.

Since the dependence on temperature should be highly significant for the part related to molecular motion and much less significant for the parts related to static disorder and errors, the dependence of ADPs on temperature has been the subject of several studies. In the field of small-molecule crystallography, a theory of the dependence of ADPs on temperature has been developed (Bürgi & Capelli, 2000; Capelli *et al.*, 2000). For macromolecules, the situation is much more complex owing to the simultaneous presence in the crystal of components with different degrees of order (*e.g.* the active site, the core of the protein, flexible surface loops, ordered and disordered solvent *etc.*).

In a study of ribonuclease A (Tilton *et al.*, 1992), nine X-ray data sets were collected at resolution of 1.5 Å in the temperature range 98–320 K. The average B value for all atoms of the protein increased from 6.6 Å^2 at 98 K to 15.7 Å^2 at 320 K. Based on the observed sensitivity of B values on temperature, the authors reached the conclusion that in the temperature range 98–320 K the atomic B values predominantly reflect molecular motion rather than static disorder. The investigators also observed that individual amino acids exhibit different temperature dependences, *e.g.* linear changes with temperature, virtual temperature independence and biphasic behaviour. In a study of human lysozyme (Joti *et al.*, 2002), seven data sets were collected from a single crystal of lysozyme in the temperature range 113–178 K and at resolutions ranging from 1.35 to 1.48 Å. These data were used for normal-mode refinement to separate the mean-square fluctuations of atoms of the protein into contributions from the internal and external degrees of freedom. The authors observed an inflection in the temperature dependence of the total mean-square fluctuations above 150 K and reached the conclusion that it was mostly related to contributions from external degrees of freedom.

These studies of the dependence of B values on temperature covered a temperature range above 98 K and were performed at resolutions no higher than 1.35 Å. We extended the study to lower temperatures (15–100 K) using cold helium gas as the cooling agent. The use of high-quality diffracting crystals of h-AR (Howard *et al.*, 2004) enabled us to obtain highly accurate structural parameters. In this paper, we focus our analysis mainly on the ordered part of the structure.

The improvements in biochemical methods, the use of synchrotron radiation, the collection of X-ray data at cryogenic temperatures and the increased speed and power of computing has contributed to the solution and refinement of more than 220 protein structures at atomic (1.2 Å) and subatomic (0.85 Å) resolutions. Protein models at these resolutions have a high observation-to-parameter ratio, which permits refinement with minimal (or without) stereochemical restraints and thus leads to values of atomic parameters that are not biased by standard ideal values. These precise ideal values used in restraints were obtained from investigations of small-molecule structures and are not necessarily valid for proteins. Therefore, it has become possible to correctly estimate and understand the contribution of thermal motion to the total atomic B values. It would also be of great interest to determine what factors contribute mainly to B values at 100 K, the most common cryogenic temperature used for data collection.

Data collection for macromolecular crystals is currently conducted at cryogenic temperatures in order to minimize the effects of radiation damage. Usually, for data collection at synchrotrons cold nitrogen gas is used as a cryogen, mainly because of its low cost. The minimum temperature that can be reached with an open-flow nitrogen-gas cryostat is 78–80 K, but for various technical reasons the sample temperature is typically maintained at 85–100 K. Neon and helium gases allow further lowering of the temperature, with a minimum temperature of 4.5–10 K reachable using an open-flow helium-gas cryostat. During the last decade, new cryotechniques and cryocooling devices using helium have been developed (Hardie *et al.*, 1998; Nakasako *et al.*, 2001; Meserschmidt *et al.*, 2003). In our experiment, the use of helium was dictated for several reasons. Firstly, it was important to obtain the atomic B values at different cryogenic temperatures and to investigate whether new conformational changes occur in the protein structure. Secondly, Hanson and coworkers demonstrated that with helium as a cryogen data quality decays more slowly during data collection than with nitrogen (Hanson *et al.*, 2002, 2003). Therefore, it was anticipated that the use of helium could diminish secondary radiation damage (by better immobilization of some free-radical species) and allow the collection of more redundant subatomic resolution data sets with the same macromolecular crystal. Thirdly, thermal diffuse scattering is significantly reduced at very low temperatures (Goeta & Howard, 2004), leading to more accurate crystallographic data and correctly estimated ADPs. Additionally, it has been observed that when helium is used as the gas stream the background on images is also reduced owing to reduced scattering of the X-rays by the cryogen gas. For example, for charge-density analysis in the field of small-molecule crystallography the data is typically collected at as low a temperature as possible. Given that h-AR is one of the best diffracting macromolecules, it was of great interest to determine the values of the atomic B factors at very low temperatures. Moreover, it was important to obtain extremely accurate crystallographic data and consequently more accurate estimates of the ADPs and an improved molecular model.

In this paper, we provide a detailed comparison of atomic B values for subatomic resolution models of h-AR in the same crystal at two different temperatures, 15 and 60 K. As a reference, we use the 0.66 and 0.93 Å structures of h-AR obtained at 100 K (Howard *et al.*, 2004). To evaluate the contributions of data errors of different origin to the atomic B

Table 1
Experimental setup and data-collection statistics for the data collected at 15, 60 and 100 K.

Values in parentheses are for the highest resolution shell.

	15 K	60 K	100 K
Approximate size of the crystal (mm)	0.800 × 0.400 × 0.150	—	0.600 × 0.400 × 0.200
X-ray energy (keV)	19.5	19.5	13.4
Size of the beam, horizontal × vertical (mm)	0.100 × 0.100	0.100 × 0.100	0.125 × 0.350
Low-resolution run			
Crystal-to-detector distance (mm)	375	375	200
Exposure time (s)	1	1	3
Oscillation width (°)	0.5	0.5	0.5
No. of frames	360 × 2 = 720	360 × 2 = 720	360
High-resolution run			
Crystal-to-detector distance (mm)	100	100	95
Exposure time (s)	3	3	5
Oscillation width (°)	0.2	0.2	0.4
No. of frames	900	900 × 2	450
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters			
a (Å)	49.17	49.21	49.35
b (Å)	66.67	66.70	66.93
c (Å)	47.30	47.31	47.40
α (°)	90.0	90.0	90.0
β (°)	92.26	92.27	92.23
γ (°)	90.0	90.0	90.0
Nominal resolution (Å)	0.81	0.81	0.93
Mosaicity	0.3	0.3	0.3
No. of unique reflections	295088	297530	199848
Completeness (%)	95.8 (93.9)	96.4 (93.0)	99.1 (95.1)
$I/\sigma(I)$	15.2 (5.1)	21.9 (6.2)	21.0 (10.9)
$R_{\rm merge}$ (%)	5.3 (20.4)	4.5 (21.3)	3.7 (9.5)

Table 2
List of the models for which atomic B values were compared.

Model	Data-collection temperature (K)	Crystal diffraction limit (Å)	Refinement resolution (Å)	Comments
T15	15	0.81	0.81	720 LR and 900 HR frames used; PDB code 2i16
T60P	60	0.81	0.81	720 LR and 900 HR frames used
T60F	60	0.81	0.81	720 LR and 1800 HR frames used; PDB code 2i17
C66	100	0.66	0.66	Howard *et al.* (2004); PDB code 1us0
C66A	100	0.66	0.81	C66 model re-refined at 0.81 Å
C66B	100	0.66	0.93	C66 model re-refined at 0.93 Å
C93	100	0.93	0.93	

values and to contrast their contributions with the effect of the temperature, several comparisons were performed. The effect of applying a high-resolution cutoff to computer-derived atomic B values was investigated by comparison of models refined against the same data but to different high-resolution limits: 0.66, 0.81 and 0.93 Å. To understand how experimental inaccuracies influence atomic B values, two models refined with the same data set, in one case using all the data and in the other case using only part of the data (a complete but less redundant data set), were compared. The influence of the quality of the crystal on the B values was studied by a comparison of two models refined in the same resolution range with data sets collected at the same temperature but from different crystals.

2. Data collection and model refinement

2.1. Crystals and data collection

The study was performed on a complex of h-AR with the inhibitor IDD 594 and the cofactor $NADP^+$. This structure has previously been solved and refined to 0.66 Å resolution (Howard *et al.*, 2004; PDB code 1us0). The details of purification and crystallization have been described previously (Lamour *et al.*, 1999; Howard *et al.*, 2004). The results reported in this paper are from two crystals that diffracted to 0.81 and 0.93 Å and used both helium-gas and nitrogen-gas cryostats. Experimental details are summarized in Table 1.

The crystal chosen for the first experiments was large in order to allow the collection of complete data sets (Table 1) from two different regions of the same crystal at temperatures of 15 and 60 K using a CRYOCOOL-LHe helium cryostat (CRYO Industries). In the second experiment, data were collected from a single region of the crystal at 100 K using a nitrogen cryostat. All data were collected using a Q315 CCD detector (ADSC) at the Advanced Photon Source (Argonne, IL, USA) 19ID beamline (Rosenbaum *et al.*, 2006).

The temperature of the helium-gas stream at the crystal position was measured with a thermocouple of approximately the same size as the h-AR crystal and calibrated with ice water and liquid nitrogen. The temperature of the helium-gas stream was also verified using an Si diode with standard curves. In order to obtain a complete diffraction data set, it was necessary to collect the data in two passes. Low-resolution reflections were first collected using a highly attenuated beam, to minimize saturations and radiation damage to the crystal, and the high-resolution data were then collected with a minimally attenuated 100 × 100 µm beam. Data were first collected at 15 K using one position of the crystal. Two data sets were low/medium resolution (LR) and one data set was high resolution (HR). The crystal was then translated by 150 µm to a previously unexposed region and the temperature was increased to 60 K. Two LR and two HR data sets were collected. Even at these low temperatures, special care was taken to minimize radiation-induced damage by using beam attenuators and by evaluating online data-quality parameters (R factors, unit cell and mosaicity stability *versus* frame number). The data collected at 60 K were processed twice: first for all the frames obtained and then for the same number of

Table 3
Refinement statistics.

R factors and e.s.d. values for atomic coordinates for the models refined with the data collected at 15, 60 and 100 K.

	T15	T60P	T60F	C66A	C66B	C93
R factor with $F > 4\sigma(F)$ (%)	7.94	7.91	7.67	7.79	7.74	7.98
R factor using all reflections (%)	8.69	8.44	8.14	7.96	7.79	8.13
⟨E.s.d.⟩ for CA atoms	0.0161	0.0155	0.0143	0.0123	0.0167	0.0187
⟨E.s.d.⟩ for CA atoms of the active site	0.0151	0.0096	0.0091	0.0079	0.0110	0.0116
⟨E.s.d.⟩ for O atoms of the active site	0.0079	0.0073	0.0071	0.0063	0.0100	0.0120
No. of atoms with e.s.d. < 0.005 Å	167	248	368	677	205	87

Table 4
Mean values of e.s.d.s in *B* values (Å^2).

	T15	T60P	T60F	C66A	C66B	C93
⟨E.s.d.⟩ in *B* factors for CA atoms in single conformation						
For atoms with $B < 10$ Å^2	0.09	0.08	0.08	0.07	0.08	0.10
All atoms	0.11	0.10	0.10	0.09	0.12	0.12
⟨E.s.d.⟩ for CB atoms in single conformation						
For atoms with $B < 10$ Å^2	0.10	0.08	0.08	0.07	0.08	0.10
All atoms	0.14	0.13	0.12	0.11	0.15	0.15
⟨E.s.d.⟩ in *B* factors for O atoms of water molecules with occupancy 1.0						
For atoms with $B < 10$ Å^2	0.12	0.08	0.09	0.08	0.09	0.09
All atoms	0.67	0.62	0.61	0.60	0.63	0.68
⟨E.s.d.⟩ in *B* factors for CA atoms of residues in double conformations	0.44	0.40	0.38	0.33	0.38	0.53

frames as were collected at 15 K, *i.e.* using the frames of one HR scan instead of two. The approximate cost of the helium needed to collect an ultrahigh-resolution data set at 15 K with low- and high-resolution passes of 360° each on aldose reductase is $50. This value does not include the liquid helium used during setup, crystal screening *etc*. The experimental setup and data analysis will be discussed in more detail in a separate publication (Ginell *et al.*, in preparation).

One further data set was collected with another crystal using a 125 × 350 µm X-ray beam at 100 K with a nitrogen-gas cryostat. Data processing, integration and merging of different data sets were performed with the program *HKL*-2000 (Otwinowski & Minor, 1997). A summary of data collection and processing is given in Table 1. Programs from the *CCP*4 suite were used for data analysis (Collaborative Computation Project, Number 4, 1994).

2.2. Refinement

The models used for the analysis are listed in Table 2. All these models were obtained in the course of refinement based on the previously solved structure C66 (Howard *et al.*, 2004). As shown previously (Afonine *et al.*, 2004), the ADP values obtained during the course of refinement have an additional systematic shift depending on the resolution zone used in refinement (with the highest resolution leading to lowest values of this shift; see §4.3). Therefore, for accurate comparison it is worthwhile to use the ADP value resulting from refinement at the same resolution. In order to adapt the initial model to the data resolution, some preliminary re-refinement of model C66 was performed against the same initial data set (but with a decreased high-resolution cutoff) before starting the refinement against newly collected data sets for models T15, T60P, T60F and C93. The procedure of re-refinement was performed step by step by using pregressively fewer and fewer data from 0.66 to 0.81 Å. The model C66A refined at 0.81 Å resolution was further used as a starting model for the refinement against 15 and 60 K data sets.

The refinement was carried out using the program *SHELXL* (Sheldrick & Schneider, 1997). The manual rebuilding of the model was performed using the programs *XtalView*/*Xfit* (McRee, 1999) and *PyMOL* (DeLano, 2006). Fourier syntheses were calculated with the program *SF*2*CNS* (Urzhumtsev & Urzhumtseva, 2002). Residues that were in a single conformation were refined without any restraints. For the other residues of the protein, distance, planarity and chiral volume restraints and DELU and SIMU restraints for ADPs were applied. The final *R*-factor values and the estimated standard deviations (e.s.d.) in atomic coordinates and *B* values are presented in Tables 3 and 4. E.s.d.s were obtained by inversion of the blocks of full matrix, where each block consisted of about 60 overlapping residues. The refinement against the data collected at 60 K was performed twice, first using all data (T60F) and then only part of the data processed with the same number of frames as for the 15 K data (T60P) (Table 2). ADPs were refined anisotropically. The total number of refined parameters for the T15 model was 34 437 (25 444 restraints) and for the T60F and T60P models was 34 454 (25 482 restraints).

3. Analysis of the temperature-dependent part of the ADP value

3.1. Increase in *B* values for protein atoms that are in a single conformation

Histograms of the distribution of atomic *B* values for models T15, T60F and C66A are presented in Fig. 1. X-ray data for these models were collected at 15, 60 and 100 K, respectively, and these models were refined in the same resolution range. These histograms and the data presented in Table 5 clearly demonstrate a general trend, *i.e.* a decrease in atomic *B* values with a lowering of the temperature of data collection. Because of this effect on atomic *B* values, one would expect to observe an increasing number of peaks above noise corresponding to H atoms in the final density map. Indeed, for parent atoms of residues in single conformations, peaks corresponding to almost all H atoms are seen in

Table 5
Mean values of atomic B values (Å^2) for residues in a single conformation for all models used in analysis.

	T15	T60P	T60F	C66	C66A	C66B	C93
Mean B	4.49	6.17	5.45	5.55	5.71	5.90	6.76
Mean B for atoms with B(T15) < 10 Å^2	3.92	5.42	4.78	4.91	5.03	5.20	5.93
Mean B for CA atoms	4.14	5.79	5.06	5.05	5.24	5.43	6.21
Mean B for CA atoms with B(T15) < 10 Å^2	3.90	5.54	4.81	4.82	5.00	5.20	5.99

$F_{obs} - F_{calc}$ H-omit difference maps. For example, in 15 and 60 K maps a peak can be observed corresponding to a H atom near the CE1 atom of His110. The result is significant because this H atom was not present above noise in the difference map calculated for the C66 model (see Fig. 2). However, in general, a comparison of 15, 60 and 100 K difference maps shows no significant increase in the number of peaks corresponding to H atoms upon lowering the temperature from 100 to 15 K. Further investigations are needed in order to understand which factors (completeness of the data, other systematic and random data errors, incorrectly determined anisotropic parameters of a parent atom) along with the B values of the closest parent atom determine the possibility of identifying H atoms at subatomic resolution.

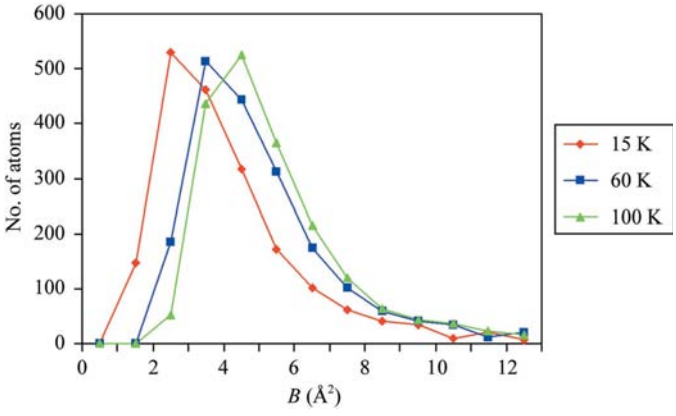

Figure 1
The distribution of the atoms according to their B values. Only atoms of residues in a single conformation are shown.

Firstly, we compared the B values for the T15 and T60P models. The data for the T15 and T60P models were collected from the same crystal at different temperatures and have the same redundancy. Therefore, the differences in the B values may be attributed to increasing thermal motion as the temperature increases from 15 to 60 K, with a possible small contribution from a radiation-damage component. The overall R factor between these two data sets is equal to 3.8% (17.2% in the highest resolution shell 0.83–0.81 Å); the difference in the Wilson B factors is equal to 1.7 Å^2 (the Wilson B factor is 3.0 and 4.8 Å^2 for the less redundant 15 and 60 K data sets, respectively). The difference between the data sets is more pronounced at resolutions higher than 0.95 Å (see Fig. 3).

The differences $\delta B = B$(T60P) $- B$(T15) for CA atoms of residues in a single conformation are presented in Fig. 4. For CA atoms and in most of the cases below, δB values were analyzed as a function of B values at 15 K. For initial B values below 7 Å^2, the δB values are almost identical (approximately 1.65 Å^2). A similar increase in δB values was observed for CB (Fig. 5), C, O and N atoms. This suggests that for the well ordered atoms of the main chain, increasing the temperature from 15 to 60 K causes an approximately equal increase in the B values. For atoms with initial B values above 7 Å^2, the mean value of δB is similar; however, the deviations of δB from the mean value are much larger. In model T15, all atoms with B values above 7 Å^2 are located on the surface of the molecule. These relatively large deviations in ADP values as the temperature increases could indicate either the appearance of alternative conformations or incorrectly determined occupancy values for these atoms, implying disorder.

The mean increase in the B values for all atoms of residues in a single conformation is 1.71 Å^2, with a standard deviation σ of 0.26 Å^2. However, a more detailed analysis revealed that this value varies slightly with the type of atom. The mean value

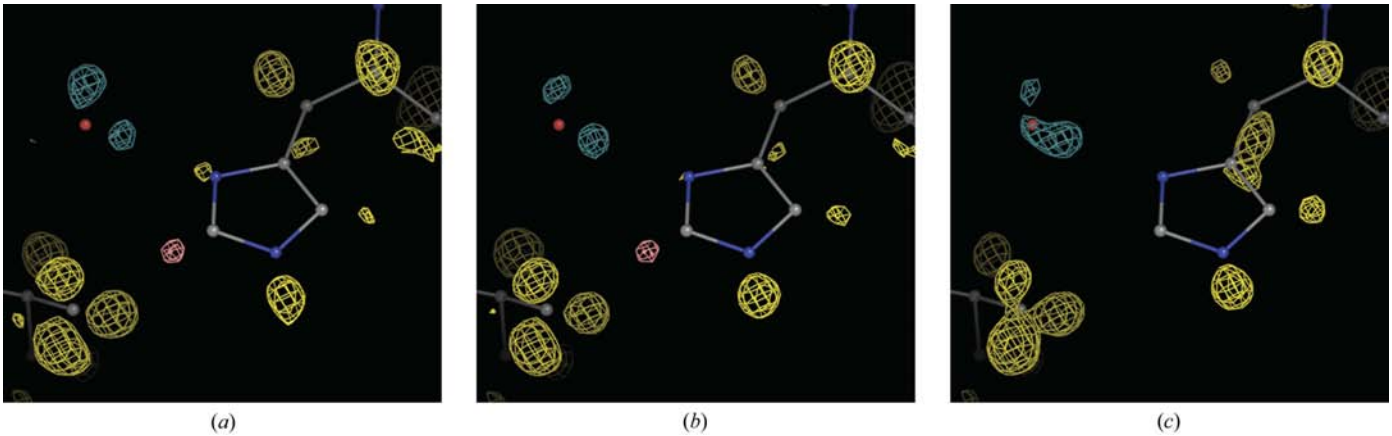

Figure 2
$F_{obs} - F_{calc}$ H-omit maps in the vicinity of His110 at 2.5σ. Protein models obtained at (*a*) 15 K, (*b*) 60 K and (*c*) 100 K. An electron-density peak corresponding to the H atom near CE1 is observed in the 15 and 60 K maps but is not present in the 100 K map.

of δB is 1.67 Å^2 (σ = 0.14 Å^2) for CA atoms, 1.71 Å^2 (σ = 0.14 Å^2) for CB atoms, 1.70 Å^2 (σ = 0.15 Å^2) for O atoms and 1.66 Å^2 (σ = 0.07 Å^2) for N atoms. This difference becomes more pronounced in the case of terminal atoms, *i.e.* atoms covalently bound to only one non-H atom. In Fig. 6, δB values for terminal C atoms are presented. The mean value of δB in this case is 1.87 Å^2, not 1.67 Å^2, as obtained for the main-chain atoms. The larger value of $\langle\delta B\rangle$ reflects the higher mobility of the terminal atoms, which results in a larger amplitude of the motion as temperature increases.

3.2. Increase in *B* values for protein atoms of residues in double conformations

Analysis of the B values for residues in double conformations gives approximately the same mean value of δB as for residues in single conformations. For CA atoms, it is equal to 1.67 Å^2 (σ = 0.30 Å^2), which is the same as the value obtained for CA atoms in single conformations. However, in this case the deviations of δB from the mean value are significantly larger (Fig. 7), most likely as a result of inaccurate estimation of the occupancies of residues in double conformations.

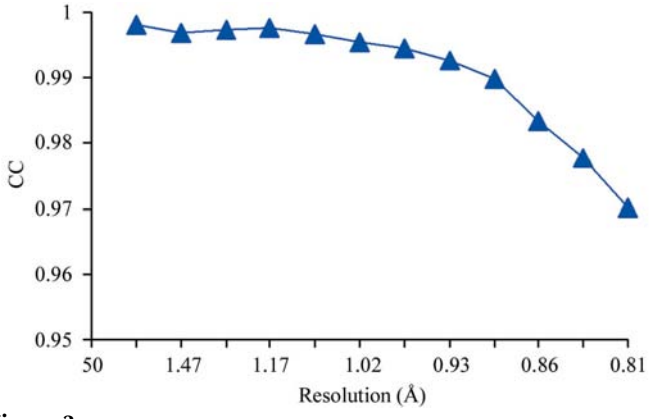

Figure 3
Correlation coefficient (CC) of structure-factor magnitudes between 15 and 60 K data sets.

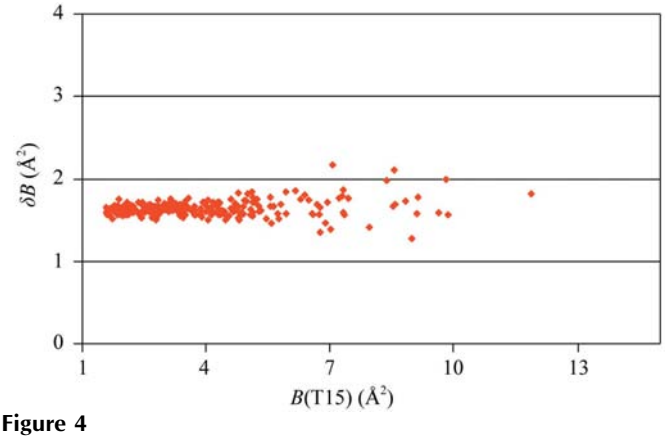

Figure 4
δB = B(T60P) − B(T15) for CA atoms of residues in a single conformation. All values are sorted according to the B values at 15 K. $\langle\delta B\rangle$ = 1.67 Å^2, σ = 0.14 Å^2.

3.3. Increase in *B* values for ordered water molecules

In macromolecular crystals, the bound solvent molecules are usually the least ordered part of the structure. Water molecules are often interpreted as possessing partial occupancy and occupying alternate positions. Nevertheless, our data show that for the O atoms of water molecules, the dependence of the δB values on temperature increase is similar to that observed for atoms of protein amino-acid side chains that are in a single conformation. The mean value of δB for all water molecules is 1.87 Å^2, σ = 1.52 Å^2 (in Fig. 8 δB is only shown for water molecules with B values less than 20 Å^2). Large fluctuations of B values for atoms with B values above 8 Å^2 may indicate a masked partial occupancy of corresponding water molecules.

3.4. Changes in anisotropy with increasing temperature

The anisotropy of an individual atom A is defined as the ratio of the minimum and maximum eigenvalues of the 3 × 3 matrix of ADPs (Trueblood *et al.*, 1996), $A = E_{\rm min}/E_{\rm max}$. For

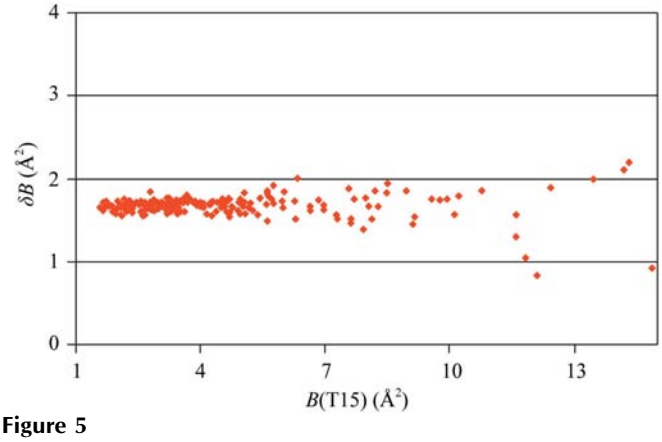

Figure 5
δB = B(T60P) − B(T15) for CB atoms of residues in a single conformation. All values are sorted according to the B values at 15 K. $\langle\delta B\rangle$ = 1.71 Å^2, σ = 0.14 Å^2.

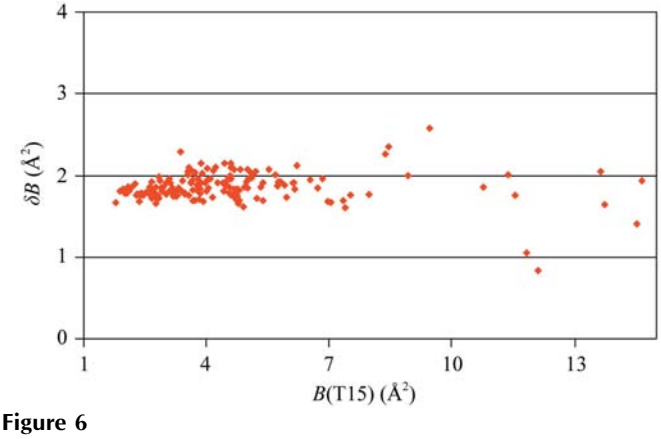

Figure 6
δB = B(T60P) − B(T15) for terminal C atoms of residues in a single conformation. All values are sorted according to the B values at 15 K. $\langle\delta B\rangle$ = 1.87 Å^2, σ = 0.27 Å^2.

the perfectly isotropic case, the value of A is equal to 1.0. As anisotropy increases, the value of A decreases.

For well ordered atoms of protein residues in a single conformation, refinement of models T15 and T60P was performed without any stereochemical restraints. In particular, no restraints on ADPs were used. The A values for CA atoms in a single conformation for models T15 and T60P are presented in Fig. 9(a). The eigenvalues for the ADP matrices were calculated using the open-source cctbx libraries (Grosse-Kunstleve *et al.*, 2002; http://cctbx.sourceforge.net). We observe that for both T15 and T60P the higher the atomic B value is, the lower the A values are on average. This tendency was expected because h-AR has a globular shape. Atoms with low B values are located close to the centre of mass, while atoms with high B values are located near the surface. As shown previously for globular proteins, the radial component of atomic displacements is approximately constant, while the tangential component of atomic displacements increases with increase of the distance from the centre of mass (Schneider, 1996). More surprisingly, the anisotropy of ADPs for individual atoms decreases as the temperature increases from 15 to 60 K (Fig. 9a). This effect could be explained by the isotropic (or almost isotropic) character of temperature-dependent differences in ADP tensors. An increase in the isotropic part of the ADP tensor leads to a reduction of anisotropy and an increase in the A value. To test this hypothesis, we calculated matrices del$U$ = U(T60P) − U(T15), their eigenvalues and anisotropy parameters for individual atoms. The A values are presented in Fig. 9(b). For most well ordered atoms the A values are relatively high, which confirms the isotropic character of the changes in ADPs with temperature.

4. Contribution of data errors to atomic B values

The thermal motion of atoms determines only part of the atomic displacement parameters. Other contributions to B values include data inaccuracy, imperfect order in the crystal, incompleteness of the set of reflections and other sources of errors. Using our data, we attempted to estimate the contributions of different errors to atomic B values in order to compare them with the contribution from the temperature.

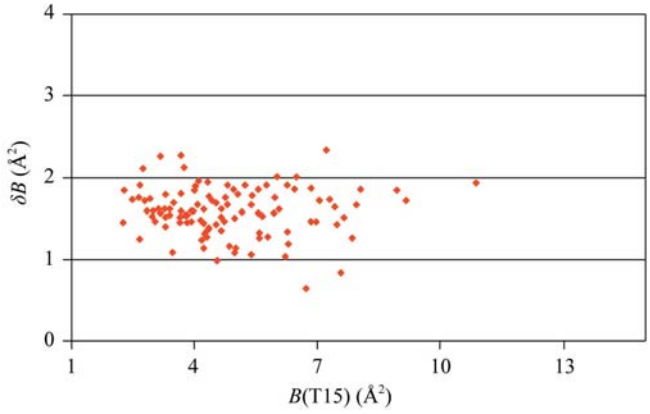

Figure 7
δB = B(T60P) − B(T15) for CA atoms of residues in double conformations. All values are sorted according to the B values at 15 K. $\langle\delta B\rangle$ = 1.67 Å^2, σ = 0.30 Å^2.

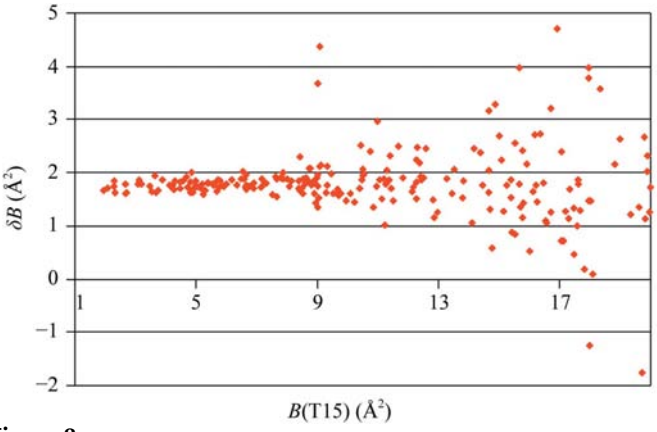

Figure 8
δB = B(T60P) − B(T15) for O atoms of well ordered water molecules. All values are sorted according to the B values at 15 K. δB is shown only for water molecules with B values less than 20 Å^2. $\langle\delta B\rangle$ = 1.87 Å^2, σ = 0.65 Å^2 for atoms with B values less than 20 Å^2.

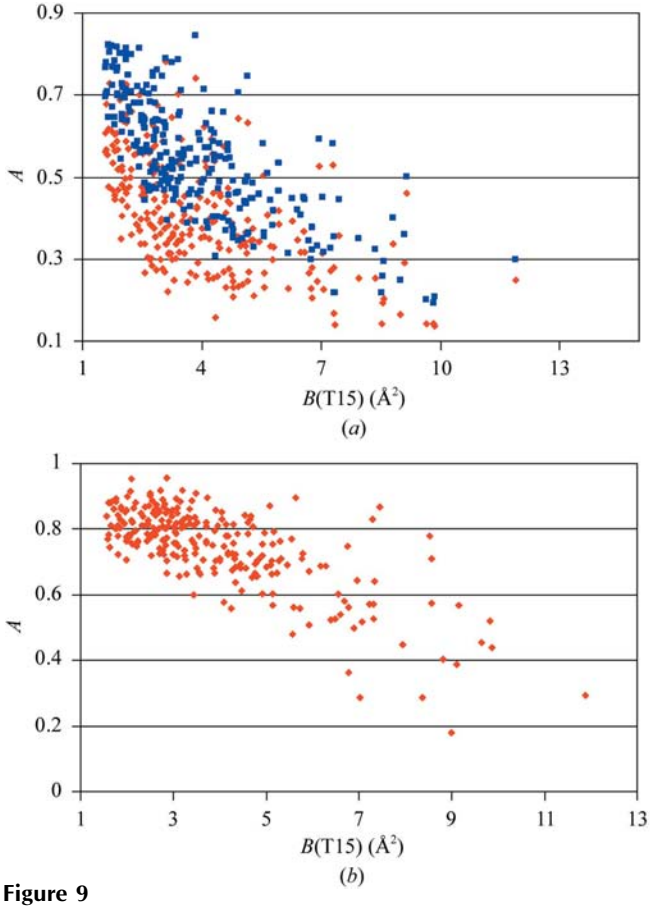

Figure 9
(a) Individual atomic A values for CA atoms of residues in a single conformation for T15 (red) and T60P (blue) models. All values are sorted according to the B values at 15 K. (b) Individual atomic A values were calculated for the matrix del$U$ = U(T60P) − U(T15) for CA atoms of residues in a single conformation. All values are sorted according to B(15 K).

4.1. Contribution of data redundancy

The refinement of the 60 K model was performed twice, firstly using a complete data set of high redundancy (model T60F) and secondly a data set of low redundancy (model T60P). In this section, we focus on a comparison of the atomic B values of these two models, T60P and T60F. A decrease in the data redundancy can result in less accurate structure-factor magnitudes and, as a consequence, in less accurate values of the atomic coordinates and increased ADP values. For the less redundant data set, only one high-resolution scan was collected, resulting in an overall redundancy of 3.6. For the full 60 K high-resolution data set, the mean value of the redundancy is approximately twice as high. The overall R factor between two data sets is 2% (10% in the highest resolution shell, 0.82 to 0.83 Å) and the difference in Wilson B factors is 0.7 Å^2 (the Wilson B factor is 4.1 and 4.8 Å^2 for the full data set and the less redundant data set, respectively).

Fig. 10 gives the δB values for CA atoms caused by a decrease in the data redundancy and hence by a decrease in the data accuracy. It shows an approximately constant increase in B values ($\langle\delta B\rangle = 0.75$, $\sigma = 0.06$ Å^2). For atomic B values less than 8 Å^2, only negligible deviations of δB from the mean value are observed. Surprisingly, the influence of the redundancy of the data is significant and is comparable with the influence of the thermal motion. The mean decrease in the atomic B values as the temperature decreases from 60 to 15 K is approximately twice as large as that caused by a twofold increase in the redundancy of the high-resolution scan data at 60 K.

Note that the plot calculated for terminal atoms is very similar to that in Fig. 10, with almost the same mean value of δB. While the δB values caused by the temperature increase (Figs. 4, 5 and 6) are different for the main-chain and 'terminal' atoms, the reduction in data redundancy causes the same increase in ADP values for all types of atoms.

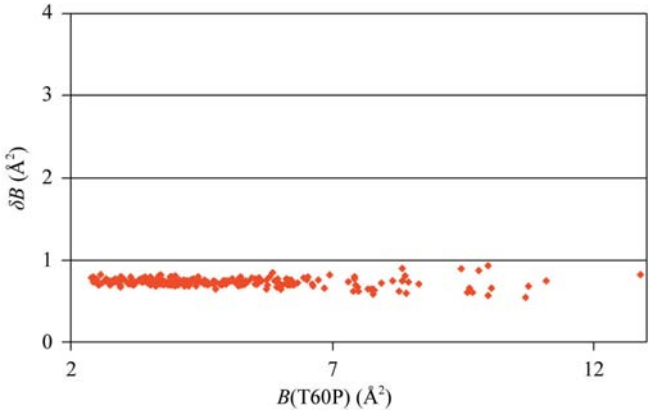

Figure 10
$\delta B = B$(T60P) − B(T60F) for CA atoms of residues in a single conformation. All values are sorted according to the B values of the atoms of the T60P model. $\langle\delta B\rangle = 0.75$ Å^2, $\sigma = 0.06$ Å^2.

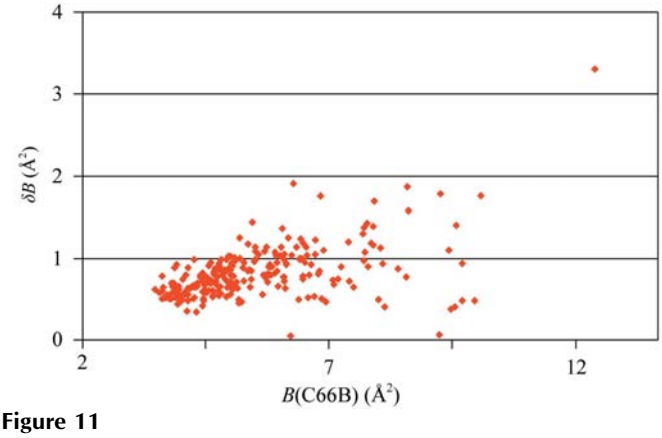

Figure 11
$\delta B = B$(C93) − B(C66B) for CA atoms of residues in a single conformation. All values are sorted according to the B values of the atoms of the C66B model.

4.2. Contribution of crystal quality

Crystal quality is a further factor that significantly influences ADP values. Reduction of order causes an increase in ADP values and a decrease in the diffracting power of the crystal. The crystal quality can be evaluated to some extent by the value of the outer resolution limit of diffraction. The high-resolution limit of diffraction data cannot be determined unambiguously since separate reflections are sometimes found well outside the resolution limits. However, criteria such as '50% of reflections must have intensity greater than 3σ' can be used to be objective. The analysis of the ADPs obtained from crystals of different quality reveals a more complicated behaviour of δB values than described above. Fig. 11 gives the δB values caused by different quality of the crystals. Model C93 was refined against data collected from a crystal with an upper resolution limit of 0.93 Å. Model C66B was refined against data collected from a crystal of a higher quality (that diffracted to 0.66 Å resolution). For comparison with the C93 model, only data to 0.93 Å were used in the refinement. As one can see from the graph (i) the average value of the δB values slightly increases with increasing B values and (ii) the

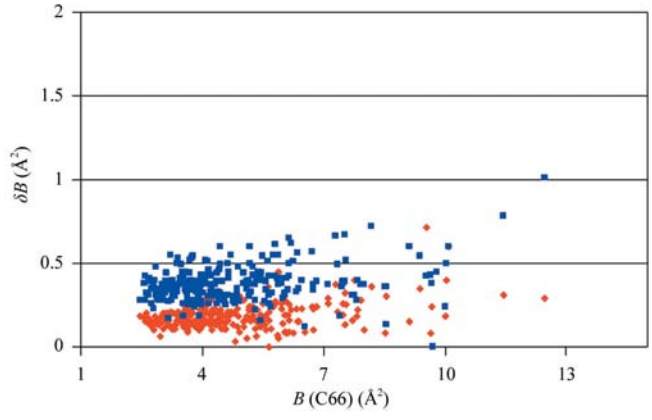

Figure 12
$\delta B = B$(C66A) − B(C66) (red, $\langle\delta B\rangle = 0.22$ Å^2, $\sigma = 0.08$ Å^2) and $\delta B = B$(C66B) − B(C66) (blue, $\langle\delta B\rangle = 0.38$ Å^2, $\sigma = 0.10$ Å^2) for CA atoms of residues in a single conformation. All values are sorted according to the B values of the atoms of the C66 model.

δB values have large deviations from the average value (Fig. 11).

4.3. Contribution of a high-resolution cutoff

As mentioned above, changing the upper limit of the resolution cutoff for reflections involved in refinement introduces a systematic shift in ADP values. This shift in ADP values may be caused by the fact that the larger the number of data used in refinement, the lower the numerical errors are for the atomic coordinates obtained during refinement. Fig. 12 shows the results of a comparison of ADP values for three models obtained in refinement against the same experimental data set (100 K, upper diffraction limit 0.66 Å). In one case the entire resolution range (C66 model), in the second case only data to 0.81 Å (C66A model) and in the third case only data to 0.93 Å (C66B model) were used during refinement. The δB values are almost constant for the atoms with B values lower than 7–8 Å^2. For the atoms possessing higher B values, larger deviations of the δB values from their mean value are observed. The differences in ADP values are relatively small compared with the factors discussed above; however, they can significantly change the ultrahigh-resolution Fourier maps (Afonine *et al.*, 2004).

5. Discussion

In the experiments with cold helium gas, two data sets at subatomic resolution were collected from the same crystal at two different temperatures, 15 and 60 K. Without taking into account the shift of the crystal between data collections, we believe that the final models T15 and T60P, which were built from data sets of equal redundancy, differ only by temperature and therefore we estimate from them the δB values corresponding to the increase of the temperature from 15 to 60 K. The average value of δB for CA atoms is about 1.7 Å^2. Thus, we conclude that at 60 K the contribution of thermal motion to the B values of CA atoms exceeds 1.7 Å^2.

We can also estimate an upper limit of the temperature contribution to the atomic B values of CA atoms. For the T60F model, the smallest atomic B value for CA atoms is 2.4 Å^2 (CA of Asp43). As the mean difference between the atomic B values for the models refined at 0.66 (C66) and 0.81 Å (C66A) is about 0.2 Å^2, we estimate that at 0.81 Å resolution at least 0.2 Å^2 of the atomic B value can be attributed to the contribution of the high-resolution cutoff. Therefore, at 60 K the thermal contribution to the atomic B values of CA atoms (T60P) is greater than 1.7 Å^2 and less than 2.2 Å^2. An analysis of the histogram of distribution of B values for the T60F model suggests that temperature provides the main contribution to the atomic B value for the best ordered atoms (about one third of the total). For the other two thirds, factors other than temperature, most probably static disorder, make the main contribution to the ADPs.

Additional experiments are needed to determine the contribution of thermal motion to atomic B values at 100 K, the temperature used most often for data collection. Nevertheless, we already can make some predictions concerning this value. For each CA atom of the T60F model, the lower limit of the contribution of temperature is about 1.7 Å^2. At 100 K, the contribution of thermal motion should be greater. For the CA atoms of the C66 model, the minimal atomic B value is 2.47 Å^2. Therefore, we can conclude that the contribution of thermal motion to atomic B values at 100 K for well ordered CA atoms is greater than 1.7 Å^2 and less than 2.5 Å^2.

These limits vary depending on the type of atom, as the amplitude of thermal motion increases slightly more for atoms linked by more than one covalent bond to neighbouring non-H atoms. Based on our results (Figs. 4, 5 and 6), we expect that the value of the temperature-dependent part of the ADP is approximately equal for all well ordered atoms in a similar chemical environment.

Atoms with B values less than approximately 8 Å^2 at 15 K exhibit equivalent responses to different sources of inaccuracy, while for larger B values the response becomes unequal. However, for double conformations, an unequal response to increasing temperature is observed for atoms in all ranges of B values. This observation, along with analysis of electron-density maps for water molecules, leads us to think that this irregularity is probably an indication of incorrectly determined occupancy values and the presence of masked alternative conformations for the corresponding residues. We estimate the upper limit of B values for the well ordered atoms both for the protein and solvent region as 8 Å^2. Note that this value is about twice as high as the mean atomic B value (see Table 5). However, the only way to confirm this hypothesis would be to build additional conformations, refine a model and check the final value of the R factor.

Not only thermal motion but also different systematic errors occurring in the process of structure determination contribute to ADPs. An analysis was performed for two sources of errors in B values: experimental errors in the determination of $F_{\rm obs}$ and the high-resolution cutoff of the resolution of the data revealed a similar behaviour of the ADPs. For well ordered atoms, the δB values are almost equal for all atoms as redundancy and resolution cutoff vary. As shown above, the contribution of redundancy exceeds the contribution of the high-resolution cutoff at subatomic resolution. Much larger differences in ADPs may arise from differing crystal quality. The comparison of different contributions to atomic B values leads us to believe that the quality of crystal (*i.e.* representing the disorder of atoms) determines the main part of the ADP values obtained. It is worthwhile to mention that the study of the dependence of ADPs on crystal quality encounters problems, as it is difficult to eliminate other sources that influence these values. It may be possible that statistical analysis of the distribution of B values in PDB structures would provide a deeper insight into this subject.

We thank the Institute of Diabetes Discovery for providing the inhibitor IDD594. We thank Zbignew Dauter for revision of the manuscript and thoughtful discussion. We thank the staff of the SBC for their help with data collection, the staff of

the IGBMC for their help with computing facilities and Nickolay Chirgadze for thought-provoking discussions. We thank Pavel Afonine, Irina Dementieva and Youngchang Kim for helpful discussions. This project was supported in part by the US Department of Energy, Office of Biological and Environmental Research under contract No. W-31-109-Eng-38, by the Centre National de la Recherché Scientifique (CNRS), by the CNRS–DFG collaboration (CERC3), by the Institut National de la Santé et de la Recherché Médicale and the Hôpital Universitaire de Strasbourg (HUS). TP and VYL were partially supported by RFBR grant 04-07-90402-b. The submitted manuscript has been created by the University of Chicago as Operator of Argonne National Laboratory ('Argonne') under Contract No. W-31-109-Eng-38 with the US Department of Energy.

References

Afonine, P., Lunin, V. Y., Muset, N. & Urzhumtsev, A. (2004). *Acta Cryst.* D**60**, 260–274.

Bürgi, H.-B. & Capelli, S. C. (2000). *Acta Cryst.* A**56**, 403–412.

Capelli, S. C., Förtsch, M. & Bürgi, H.-B. (2000). *Acta Cryst.* A**56**, 413–424.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

DeLano, W. L. (2006). *The PyMOL Molecular Visualization System.* DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org.

Goeta, A. E. & Howard, J. A. K. (2004). *Chem. Soc. Rev.* **33**, 490–500.

Grosse-Kunstleve, R. W., Sauter, N. K., Moriarty, N. W. & Adams, P. D. (2002). *J. Appl. Cryst.* **35**, 126–136.

Hanson, B. L., Harp, J. M., Kirschbaum, K., Schall, C. A., DeWitt, K., Howard, A., Pinkerton, A. A. & Bunick, G. J. (2002). *J. Synchrotron Rad.* **9**, 375–381.

Hanson, B. L., Schall, C. A. & Bunicke, G. J. (2003). *J. Struct. Biol.* **142**, 77–87.

Hardie, M. J., Kirschbaum, K., Martin, A. & Pinkerton, A. A. (1998). *J. Appl. Cryst.* **31**, 815–817.

Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A. & Podjarny, A. (2004). *Proteins*, **55**, 792–804.

Joti, Y., Nakasako, M., Kidera, A. & Go, N. (2002). *Acta Cryst.* D**58**, 1421–1432.

Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorselaer, A., Podjarny, A. & Moras, D. (1999). *Acta Cryst.* D**55**, 721–723.

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Meserschmidt, M., Meyer, M. & Luger, P. (2003). *J. Appl. Cryst.* **36**, 1452–1454.

Nakasako, M., Sawano, M. & Kawamoto, M. (2001). *Rigaku J.* **18**, 47–53.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rosenbaum, G., Alkire, R. W., Evans, G., Rotella, F. J., Lazarski, K., Zhang, R. G., Ginell, S. L., Duke, N., Naday, I., Lazarz, J., Molitsky, M. J., Keefe, L., Gonczy, J., Rock. L., Sanishvili, R., Walsh, M. A., Westbrook, E. & Joachimiak, A. (2006). *J. Synchrotron Rad.* **13**, 30–45.

Schneider, T. R. (1996). *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 133–144. Warrington: Daresbury Laboratory.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Tilton, R. F. Jr, Dewan, J. C. & Petsko, G. A. (1992). *Biochemistry*, **31**, 2469–2481.

Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Cryst.* A**52**, 770–781.

Urzhumtsev, A. & Urzhumtseva, L. (2002). *J. Appl. Cryst.* **35**, 750.