Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Atomic resolution structure of native porcine pancreatic elastase at 1.1 Å

Martin Würtele, Michael Hahn, Kai Hilpert and Wolfgang Höhne*

Institut für Biochemie der Charité, Humboldt-Universität zu Berlin, Monbijoustrasse 2, D-10117 Berlin, Germany

Correspondence e-mail: wolfgang.hoehne@charite.de

Received 4 November 1999
Accepted 5 January 2000

PDB Reference: porcine pancreatic elastase, 1qnj.

A data set from the serine protease porcine pancreatic elastase was collected at atomic resolution (1.1 Å) with synchrotron radiation. The improved resolution allows the determination of atom positions with high accuracy, as well as the localization of H atoms. Three residues could be modelled in alternative positions. The catalytic triad of elastase consists of His57, Asp102 and Ser195. The His57 $N^{\delta1}$ H atom was located at a distance of 0.82 Å from the $N^{\delta1}$ atom. The distance between His57 $N^{\delta1}$ and Asp102 $O^{\delta2}$ is 2.70 ± 0.04 Å, thus indicating normal hydrogen-bonding geometry. Additional H atoms at His57 $N^{\varepsilon2}$ and Ser195 O^{γ} could not be identified in the $F_o - F_c$ density maps.

1. Introduction

The availability of high-intensity X-ray beam sources such as synchrotrons and of efficient X-ray imaging detectors, as well as the introduction of macromolecular cryocrystallography, has led in recent years to the solution and refinement of a series of proteins at atomic resolution (for a review, see Dauter *et al.*, 1997). According to Sheldrick (1990), a protein structure can be solved at atomic resolution when its diffraction data extend to at least 1.2 Å. In such structures, the relatively large data-to-parameter ratio allows anisotropic temperature-factor refinement. This leads to a significant improvement of the calculated phases and to a strong regression of the crystallographic R factor, thus allowing a very accurate calculation of the coordinates as well as the coordinate error of the structure.

Here, we report the refinement of native porcine pancreatic elastase (PPE) at a resolution of 1.1 Å. The digestive enzyme PPE (E.C. 3.4.21.36) is a multi-specific serine protease composed of 240 amino acids in its active state (for a review, see Bode *et al.*, 1989). Since its tertiary structure was elucidated for the first time (Watson *et al.*, 1970), a number of structures of native and complexed PPE have been published, making PPE one of the best structurally studied enzymes so far (1.65 Å, Meyer *et al.*, 1988; 1.6 Å, Wilmouth *et al.*, 1998). Besides extending the accuracy of the PPE structure to the thus far highest resolution, the PPE structure reported here provides insight into the protonation of the catalytic triad of this enzyme in the native state.

2. Experimental

2.1. Crystallization, data collection and data processing

Porcine pancreatic elastase was purchased from Serva (Heidelberg). Crystals were grown at 291 K using the hanging-drop vapour-diffusion method from a solution containing 0.25 M Na_2SO_4. Protein (40 mg ml^{-1}) and precipitant were mixed in a 1:1 ratio. The final pH was determined to be between 7.0 and 8.0. Crystals grew overnight to dimensions of 0.2 × 0.2 × 0.4 mm. Crystals were soaked in 20% glycerol for cryo-cooling prior to X-ray exposure. A high-resolution (20.7–1.1 Å) and a low-resolution (45.6–1.5 Å) data set were collected at 100 K from a single crystal with 0.5° oscillations at the X11 synchrotron beamline, DESY, Hamburg using a MAR345 imaging plate (X-ray Research, Hamburg) and a wavelength of 0.9116 Å.

Data were processed using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Crystals belong to the orthorhombic space group $P2_12_12_1$ and have unit-cell dimensions a = 49.91, b = 57.82, c = 74.27 Å. They contain one elastase molecule per asymmetric unit. Crystal mosaicity was refined to 0.66°. A total of 1 071 726 observations in the resolution range 45.6–1.1 Å were reduced to 81 820 unique reflections. The merged data set contained 92.7% of the reflections expected in this resolution range. The overall $R_{\rm merge}$ was 4.5% (35% in the highest resolution shell; for details see Table 1).

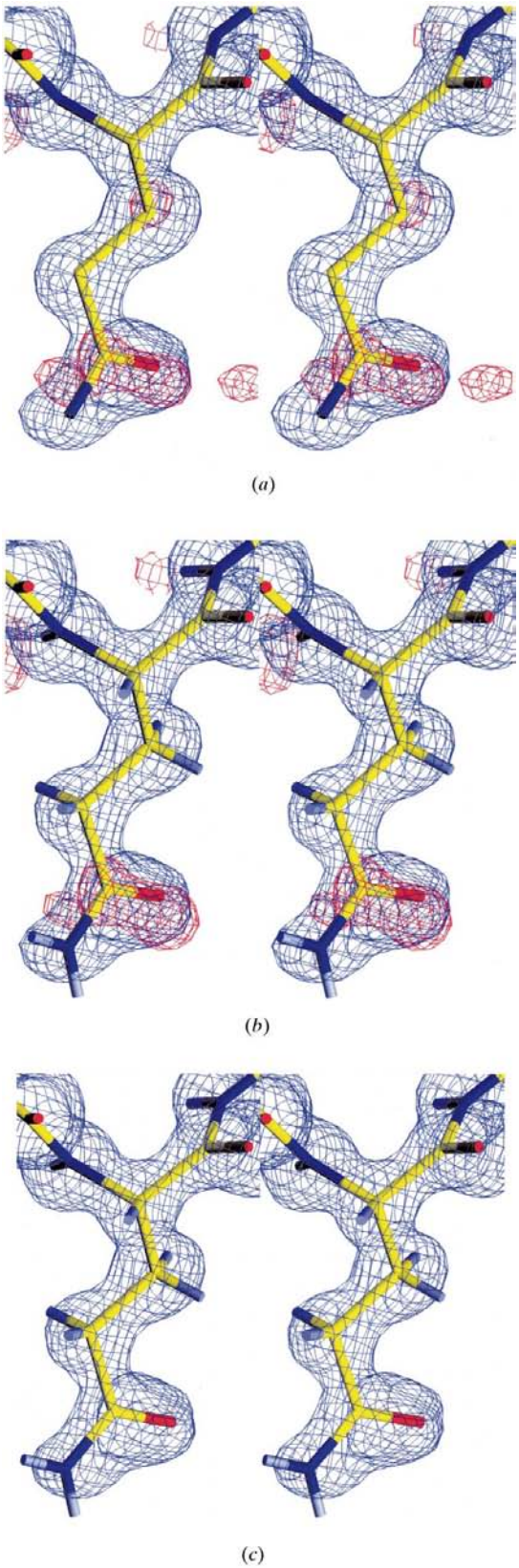

Figure 1
Stereoview of Glu153 showing phase improvement after anisotropic *B*-factor refinement. $2F_o - F_c$ maps are contoured at 1.0σ (blue); $F_o - F_c$ maps are contoured at 2.5σ (red). (*a*) Isotropically refined model. (*b*) Isotropically refined model with riding H atoms. (*c*) Model with anisotropically refined *B* factors. Images were produced with *SETOR* (Evans, 1993).

2.2. Refinement

A first model was calculated after molecular replacement (*AMoRe*; Navaza, 1994) using the 1.6 Å resolution structure of PPE (Wilmouth *et al.*, 1998; PDB entry code 1btu). No significant non-isomorphism was detected compared with the published structure. It was refined in a first step with isotropic *B* factors using the maximum-likelihood method with *REFMAC* (Murshudov *et al.*, 1997). *R* and $R_{\rm free}$ converged at 25.8 and 27.4%, respectively. In a second step, *ARPP* (Lamzin & Wilson, 1993) was used together with *REFMAC* to localize a total of 347 solvent molecules. Water peaks were inserted by *ARPP* according to their distances to hydrogen-bonding partners and omitted from the structure if the peak height in the $2F_o - F_c$ electron-density maps was less than $0.5\sigma_{\rm r.m.s.}$. All water molecules were carefully inspected (and corrected) manually using *O* (Jones *et al.*, 1991) and *XFIT* (McRee, 1993). An isotropic *B*-factor cutoff of 60 Å^2 for solvent molecules was introduced to avoid overfitting. Refinement in this step converged at $R = 19.5\%$ and $R_{\rm free} = 21.5\%$.

In the next step, anisotropic *B* factors were introduced and the structure was refined using conjugate-gradient least-squares minimization with *SHELXL*97 (Sheldrick & Schneider, 1997). Anisotropic *B* factors of protein atoms were restrained in the atom bond direction (DELU command) and, according to their distances (SIMU command), the anisotropic *B* factors of water molecules were restrained to isotropy with the ISOR command. The respective restraints in the final refinement cycles were set to 0.1, 0.025 and 0.1 as indicated by Merritt (1999). Some water molecules and three amino-acid residues showed split positions and were refined with the sum of their occupational factors constrained to unity; additionally, riding H atoms were introduced into the model. In a final refinement step, a single full-matrix least-squares minimization cycle (LS command) was carried out. In this cycle, all restraints were turned off and *B* factors were not refined (BLOC 1 command). This final model led to $R = 12.67\%$ and $R_{\rm free} = 16.52\%$. Inclusion of the 5% test-set reflections did not improve the *R* factor. The model was evaluated using *PROCHECK* (Laskowski *et al.*, 1993), *WHAT CHECK* (Hooft *et al.*, 1996) and *PARVATI* (Merritt, 1999).

3. Results and discussion

The decisive step in the refinement of atomic resolution structures is the introduction of anisotropic temperature factors. In our case, besides lowering the crystallographic *R* factor, this led to a substantial improvement of the quality of calculated phases, as seen in the electron-density maps. As an example, Fig. 1 shows the $2F_o - F_c$ electron-density map and the $F_o - F_c$ electron-density map of Glu153. Compared with the isotropic starting model (Fig. 1*a*) and an isotropic model with riding H atoms (Fig. 1*b*), the reduction of the 2.5σ contoured $F_o - F_c$ electron density is apparent in the anisotropic *B*-factor model (Fig. 1*c*). This effect is observable throughout the whole protein molecule. In the final model, electron difference peaks at the 2.5σ level are almost restricted to the solvent region.

The resulting model shows excellent agreement between the localization of the polypeptide chain and the $2F_o - F_c$ electron-density maps. Fig. 2(*a*) shows the 3.5σ contoured $2F_o - F_c$ map of the catalytic site. The atomic resolution of most atoms is clearly visible at this contour level. Electron density does smear between some atoms, but only in the case of some conjugated electron systems (*e.g.* peptide bonds). From this map, the expected conformation of the His57 imidazole ring is confirmed by the localization of the larger N atoms. The equal electron distribution of the carboxyl group of Asp102 is also clearly observable.

The high resolution of the elastase was employed to improve several features of the model, such as the correct conformation of most asparagine and glutamine residues according to the expected electron distribution in the amide group. Also, a sodium ion from the precipitant is found in the known calcium-binding site. However, a few residues could not be localized completely. These residues (notably Arg36, Ser36C, Arg125) are generally part of surface loops and have large *B* factors.

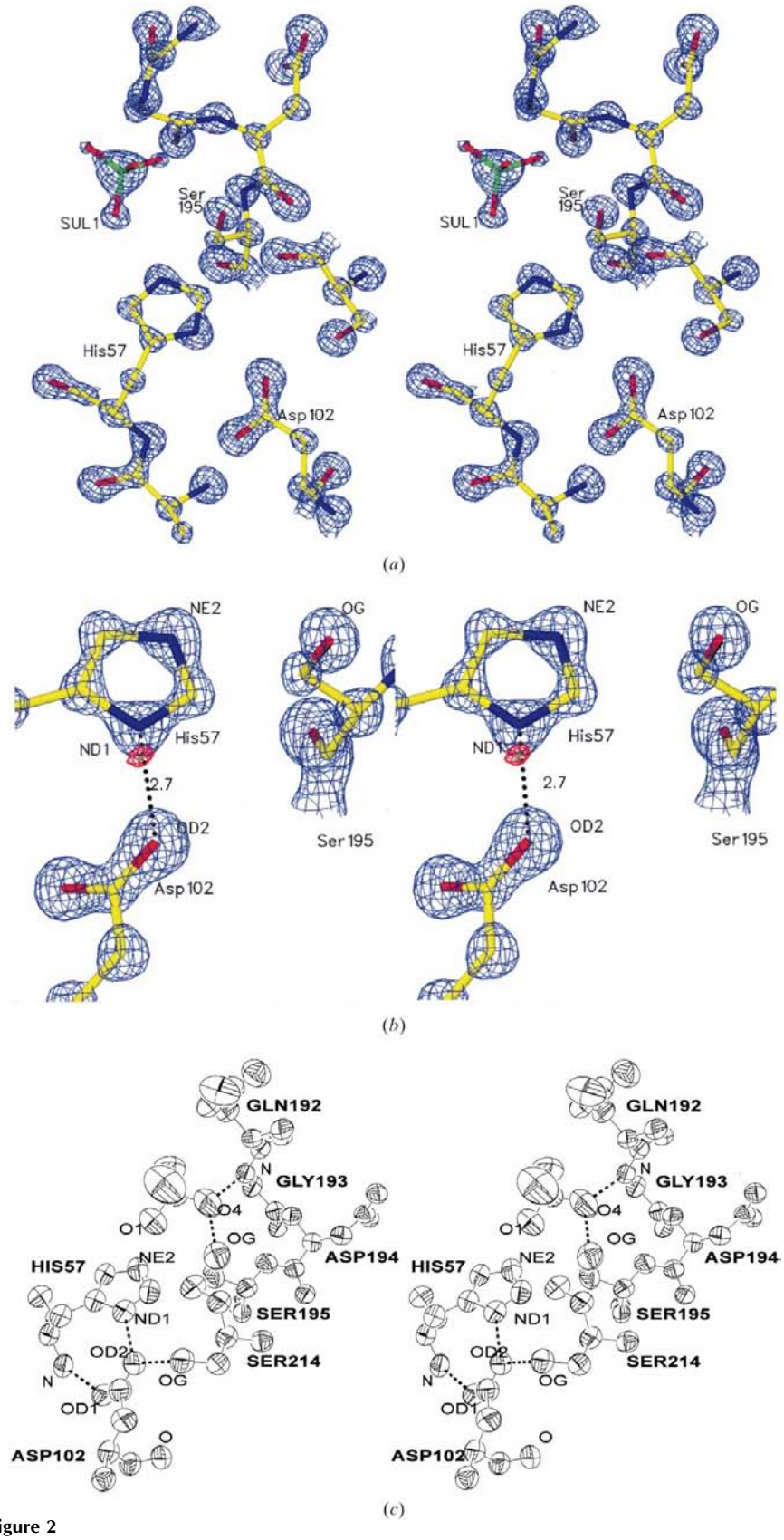

Figure 2
The catalytic triad of PPE. (*a*) $2F_o - F_c$ map contoured at 3.5σ (blue) shows atomic resolution of active-site residues of the final model (SUL1: sulfate). (*b*) $2F_o - F_c$ map of anisotropic *B*-factor model with the $HN^{\delta1}$ omitted contoured at 2.5σ (blue). $F_o - F_c$ map is contoured at 2.5σ (red). (*c*) Thermal ellipsoid representation of active-site residues was made with *ORTEX* 7e (McArdle, 1995; Burnett & Johnson, 1996). Putative hydrogen bonds are shown as dotted lines.

Phase improvement led to the localization of at least three residues with alternative positions (Val83, Ser217, Arg223). In general, our resulting model superposes very well with the 1.6 Å PPE model of an enzyme–inhibitor complex soaked with 20% glycerol (Wilmouth *et al.*, 1998) and the 1.65 Å model of native PPE in 70% methanol (Meyer *et al.*, 1988). Compared with the starting model of Wilmouth *et al.* (1998), our structure has an r.m.s. deviation of 0.43 Å for all protein atoms and 0.14 Å for the C^{α} atoms. The corresponding values for the PPE structure reported by Meyer *et al.* (1988) are 0.82 and 0.29 Å, respectively. The higher values in the latter case might be caused by the different cryo-buffers used. In both cases, the highest deviations correspond to atoms of the side chains of surface amino acids that have a different conformation in our structure (*e.g.* for Arg61 NH2, 5.7 Å deviation from the model of Wilmouth *et al.*, 1998 and 8.2 Å deviation from the structure of Meyer *et al.*, 1988). The highest deviation of a Cα atom corresponds to a slightly shifted surface loop (Gly36B Cα, 0.7 Å deviation from the model of Wilmouth *et al.*, 1998; Ser36C Cα, 1.0 Å deviation from the model of Meyer *et al.*, 1988).

Based on a cutoff radius of 1 Å, 85% of all water molecules in the model of Wilmouth *et al.* (1998) and 78% of all water molecules in the model of Meyer *et al.* (1988) can be considered as conserved in our model. This is especially valid for the internal water clusters of PPE as described by Meyer *et al.* (1988). Moreover, the higher resolution of our data allowed the localization of more water molecules than in these previous structures.

The coordinate error of the final model was calculated as the estimated standard deviation (e.s.d.) of the coordinates by a final full-matrix least-squares minimization. This procedure yields e.s.d.s between 0.01 and 0.05 Å for about 80% of all protein atoms. E.s.d. values are highly correlated with the temperature factor. The e.s.d.s of His57 $N^{\delta1}$ and $N^{\varepsilon2}$, Asp102 $O^{\delta1}$ and $O^{\delta2}$ as well as of Ser195 O^{γ} are of the magnitude of 0.02 Å. The calculated distances between some of these atoms together with their error margins are shown in Table 2. The anisotropy ($A = E_{\rm min}/E_{\rm max}$, the quotient of the minimum and maximum eigenvectors of the anisotropic displacement parameter matrix as defined by Trueblood *et al.*, 1996) of all protein atoms has a mean value of 0.527 ($\sigma = 0.148$) and is, as expected, strongly correlated with the corresponding distance

Table 1
Crystallographic overview of data and model.

Values corresponding to the highest resolution shell are shown in parentheses.

Resolution (Å)	45.6–1.1 (1.14–1.10)
Number of reflections	1071726
Number of unique reflections	81820
Completeness (%)	92.7 (60.4)
I/σ	14.9 (2.02)
R_{merge} (%)	4.5 (35.0)
R_{work} (%)	12.67
R_{free} (%)	16.52
R.m.s. bond lengths (Å)	0.016
R.m.s. bond angles (°)	2.291

Table 2
Distances between atoms of the catalytic site of PPE.

Accuracy is given as the sum of the e.s.d.s of the coordinates of both atoms.

Residues	Distance (Å)	Accuracy (Å)
Ser195 O^{γ}–His57 $N^{\varepsilon 2}$	3.17	0.044
His57 $N^{\delta 1}$–Asp102 $O^{\delta 2}$	2.70	0.039
Ser195 O^{γ}– SUL1 O4	2.60	0.056
SUL1 O4–Gly193 N	2.89	0.055
His57 $N^{\varepsilon 2}$–SUL1 O1	2.71	0.050
Asp102 $O^{\delta 2}$–Ser214 O^{γ}	2.69	0.032
Asp102 $O^{\delta 1}$–His57 N	2.82	0.038

of the atoms from the centre of mass of the molecule.

In the final model, the r.m.s. deviation of all atom distances from ideality is 0.016 Å and the r.m.s. of all binding angles is 2.291°. No residues have forbidden conformations in the Ramachandran plot, 87.4% of all non-glycine and non-proline residues being in the most favoured conformation and 12.6% in the additionally allowed region. These are roughly the same values as in the previous 1.6 Å structure of Meyer *et al.* (1988). Like in other atomic resolution structures (Dauter *et al.*, 1997; Longhi *et al.*, 1997), we see a significant deviation of the peptide torsion angle ω from planarity. This value ranges from 156.4° (Pro198) to 191.9° (Ser170). The mean value of this angle is 178.6°, with a standard deviation of 6.0°.

An interesting feature of atomic resolution structures is the ability to deduce the position of H atoms from $F_o - F_c$ maps. In the case of PPE, we were able to localize the position of the $HN^{\delta 1}$ atom of the catalytic His57 (Fig. 2*b*). It has been hypothesized that this atom forms a so-called LBHB (low-barrier hydrogen bond) near the transition state of the peptide-cleavage reaction, thus strengthening the basicity of the His57 $N^{\varepsilon 2}$ atom so that it can more easily accept a proton from Ser195. Although this would explain the observed enhanced nucleophilicity of Ser195, experimental verification of this hypothesis has led to controversial results in recent years (Frey *et al.*, 1994; Ash *et al.*, 1997).

To observe the position of the $HN^{\delta 1}$ atom in native elastase, an OMIT map was calculated after removing this atom from the model. To avoid covering the difference electron density by anisotropic refinement of the $N^{\delta 1}$ atom, the *B* factors of this atom were refined isotropically. After 20 CGLS cycles, the maps shown in Fig. 2(*b*) were calculated. An $F_o - F_c$ peak at the 2.5σ level which would correspond to a histidine H atom can be observed. The distance of the maximum of this peak to the $N^{\delta 1}$ atom is 0.82 Å, whereas the distance between the His57 $N^{\delta 1}$ and the Asp102 $O^{\delta 2}$ atom is 2.70 $\pm$ 0.04 Å. In an LBHB, the distance between the donor and the acceptor atom is shorter and the H atom is further away from the donor (Frey *et al.*, 1994). The $HN^{\delta 1}$ atom therefore shows conventional hydrogen-bonding geometry in our structure, as expected for native PPE. Calculation of an OMIT map with anisotropic *B* factors for $N^{\delta 1}$ and only five CGLS cycles led to the same result (not shown).

To validate both methods we calculated OMIT maps of the H atoms of 36 main-chain N atoms. In about 40% of all cases, we were able to find $F_o - F_c$ peaks that corresponded to the expected H atoms. The results obtained by refining the N atoms isotropically or anisotropically (only five CGLS cycles) were roughly the same. This indicates that the $F_o - F_c$ electron-density peak related to the $HN^{\delta 1}$ atom in our OMIT calculations is significant.

We were not able to calculate OMIT maps that led either to the unequivocal identification of a significant $F_o - F_c$ electron-density peak that could be related to His57 $HN^{\varepsilon 2}$ or Ser195 HO. This observation was also made in other atomic resolution structures of lytic enzymes [cutinase at pH 7 (Longhi *et al.*, 1997) and subtilisin at pH 5.9 (Kuhn *et al.*, 1998)].

Fig. 2(*c*) shows the thermal ellipsoid representation of the atoms near the catalytic triad of PPE. Probable hydrogen bonds are shown as dotted lines. A sulfate ion is localized in the 'oxyanion hole'. This structure represents PPE in its relaxed native state at atomic resolution and is thus a starting point for further atomic resolution studies concerning the enzymatic mechanism of this and other serine proteases.

The authors thank Dr Ehmke Pohl, Dr Victor Lamzin and the EMBL Hamburg Outstation staff for assigning us synchrotron beam time as well as for instructing us how to obtain atomic resolution. We additionally thank Dr Jacqueline Aÿ for help with the manuscript.

References

Ash, E. L., Sudmeier, J. L., De Fabo, E. C. & Bachovchin, W. W. (1997). *Science*, **278**, 1128–1132.
Bode, W., Meyer, E. & Powers, J. C. (1989). *Biochemistry*, **28**, 1951–1963.
Burnett, M. N. & Johnson, C. K. (1996). *ORTEP III: Oak Ridge Thermal Ellipsoid Plot Program for Crystal Structure Illustrations*. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.
Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.
Frey, P. A., Whitt, S. A. & Tobin, J. B. (1994). *Science*, **264**, 1927–1930.
Hooft, R. W. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kuhn, P., Knapp, M., Soltis, S. M., Ganshaw, G., Thoene, M. & Bott, R. (1998). *Biochemistry*, **37**, 13446–13452.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. & Cambillau, C. (1997). *J. Mol. Biol.* **268**, 779–799.
McArdle, P. (1995). *J. Appl. Cryst.* **28**, 65.
McRee, D. (1993). *Practical Protein Crystallography*. San Diego: Academic Press.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.
Meyer, E., Cole, G., Radhakrishnan, R. & Epp, O. (1988). *Acta Cryst.* B**44**, 26–38.
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.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Trueblood, K. N., Bürgi, H.-B., Burylaff, H. H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Cryst.* A**52**, 770–781.
Watson, H. C., Shotton, D. M., Cox, J. M. & Muirhead, H. (1970). *Nature (London)*, **225**, 806–811.
Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D. W., Clifton, I. J., Pritchard, G. J., Aplin, R. T. & Schofield, C. J. (1998). *Biochemistry*, **37**, 17506–17513.