Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The incorporation of a non-natural amino acid (aza-tryptophan) may help to crystallize a protein and to solve its crystal structure. Application to bacteriophage λ lysozyme.

Christine Evrard,[a,b] Jacques Fastrez[b] and Jean-Paul Declercq[a]*

[a]Laboratoire de Chimie physique et de Cristallographie, Université Catholique de Louvain, Place Louis Pasteur 1, B-1348 Louvain-la-Neuve, Belgium, and [b]Laboratoire de Biochimie physique et des Biopolymères, Université Catholique de Louvain, Place Louis Pasteur 1, B-1348 Louvain-la-Neuve, Belgium

Correspondence e-mail: declercq@cpmc.ucl.ac.be

Received 18 June 1998
Accepted 16 September 1998

PDB Reference: bacteriophage λ lysozyme, 1am7.

Until now, wild-type bacteriophage λ lysozyme had been impossible to crystallize. This difficulty could be overcome by the replacement of the four tryptophan residues by aza-tryptophans. Analysis of the intermolecular and intramolecular contacts in this modification allows understanding of the differences in behaviour between the native and modified molecules. Furthermore, this mutation was very useful for the creation of new heavy-atom binding sites and for the solution of the non-crystallographic symmetry, which is extremely important for phase improvement. This procedure seems to be generally applicable, at least in the search for new possibilities for heavy-atom binding sites.

1. Introduction

In the course of the analysis of a protein crystal structure, some modifications of the primary structure can be extremely helpful. Most modifications are mutations, *i.e.* replacements of one or several natural amino acids in the sequence. We present here a non-exhaustive list of some recent examples of such mutations which have helped to improve at least one step of the analysis. During crystallization experiments on rat procathepsin B (Sivaraman *et al.*, 1996), it was found to be of interest to avoid self-aggregation by mutating exposed Cys residues. In the same reference, a Ser→Ala mutation was introduced to eliminate a glycosylation site. In the crystallization of human H ferritin, a Lys→Gln substitution was engineered (Lawson *et al.*, 1991) to recreate an intermolecular metal bridge that had been described earlier as an important crystal contact in other known ferritin structures. Also with ferritin, it has been shown (Takeda *et al.*, 1995) that appropriate alteration in the surface charge distribution induces new molecular alignment and hence controls the crystal form. Mutations are also widely used to create heavy-atom binding sites for multiple isomorphous replacement (MIR). The best-known examples are cysteine mutations providing binding sites for Hg atoms (Martinezhackert *et al.*, 1996) or for small dense metal clusters at selected locations (Thygesen *et al.*, 1996). Another well known modification concerns the resolution of the phase problem by the multiple anomalous dispersion (MAD) method and includes the replacement of methionine residues by selenomethionines (Hendrickson, 1991).

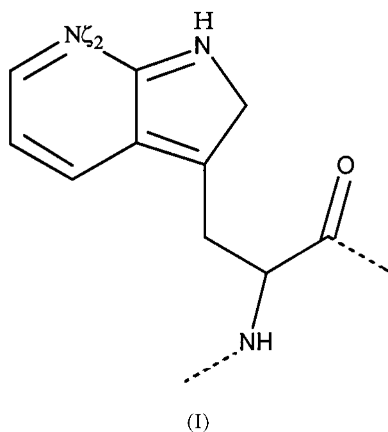
Here, we report a new kind of modification applied to the crystal structure analysis of bacteriophage λ lysozyme (λL). This structure has recently been solved (Evrard *et al.*, 1998) by the multiple isomorphous replacement method and the application of non-crystallographic symmetry. The replacement of the tryptophan (Trp) residues by aza-tryptophans

© 1999 International Union of Crystallography
Printed in Denmark – all rights reserved

(Trn) was essential for crystallization, preparation of heavy-atom derivatives and solution of the non-crystallographic symmetry. The crystal structure confirms that λL is a prototype of a new lysozyme family, evolutionarily related to the type C (chicken) and V (viral) lysozymes, and that the application of the molecular-replacement method (for a review, see Rossmann, 1990) would not have been possible. The originality of this new modification consists mainly of its usefulness in different steps of the analysis and in its general applicability, at least for some aspects.

2. Experimental

After many years of effort directed towards crystallization of the native enzyme, no crystal suitable for even preliminary analysis could be obtained. A number of variants, obtained either by genetic engineering or incorporation of non-natural amino acids in the protein, were used for screening the crystallization conditions. The best results were obtained with a molecule in which the four tryptophan residues present in the sequence were replaced by aza-tryptophans [$C^{\zeta 2}$ replaced by $N^{\zeta 2}$; (I)]. This non-natural enzyme was initially prepared for a thermodynamic study of the protein stability (Soumillion *et al.*, 1995). This modification reduces the hydrophobic character of the side chain and may allow new contacts by means of hydrogen bonds.



(I)

The expression and purification of this molecule have been described (Soumillion & Fastrez, 1992; Soumillion *et al.*, 1995). Crystals were grown under microgravity conditions during flight STS-67 of the Space Shuttle by sitting-drop vapour diffusion of the protein solution (20 mg ml^{-1} in 50 m*M* NaH_2PO_4/Na_2HPO_4 pH 6.7, 0.02% NaN_3) mixed 50:50 with a well solution of 20%(*w*/*v*) PEG 4000, 15%(*v*/*v*) 2-propanol, 0.1 *M* sodium citrate pH 5.3, as previously reported (Evrard *et al.*, 1997). The crystals are orthorhombic and belong to space group $P2_12_12_1$ with unit-cell dimensions *a* = 73.01, *b* = 78.80, *c* = 82.31 Å. Synchrotron data at 2.3 Å were collected on beamline X31 at EMBL c/o DESY (Evrard *et al.*, 1997). Solution of the crystal structure (Evrard *et al.*, 1998) revealed the presence of three independent molecules in the asymmetric unit. The molecules are essentially composed of two domains connected by a long helix (H3). The principal catalytic residue (Glu19) is located in the cleft between the two domains, at the C-terminal end of helix H1.

All figures were produced using the program *O* (Jones *et al.*, 1991) and the coordinates of λL are available from the Protein Data Bank (Bernstein *et al.*, 1977).

3. Results and discussion

3.1. Why does the modified molecule give better crystallization results than the wild enzyme?

Knowledge of the crystal structure allows us to look specifically at the intramolecular or intermolecular contacts involving the modified tryptophan residues. For this analysis, two preliminary determinations have to be made: firstly, the packing of the molecules and, secondly, whether the tryptophan side chains are exposed.

3.1.1. The packing of the molecules. As shown in Fig. 1(*a*), the three independent molecules of the asymmetric unit are nearly perfectly stacked on top of each other in the direction of the *b* crystallographic axis. A succession of three independent molecules in this direction corresponds exactly to the length of one unit cell. An infinite column of molecules (–*A*–*B*–*C*–*A*–*B*–*C*–) thus appears in that direction, involving close intermolecular contacts between the crystallographically independent molecules.

In the other two directions, the *ac* face can be separated into four equivalent rectangles (almost squares) and one column of molecules can be projected onto each of these rectangles (Fig. 1*b*). There are large solvent channels between the four columns. As a consequence, the intermolecular contacts between the columns are much looser than those inside a column.

3.1.2. The accessibility of the tryptophan residues. The sequence of λL contains four tryptophan residues, in positions 17, 73, 74 and 124. It has been shown (Evrard *et al.*, 1998) that the independent molecules *A* and *C* are very similar, while the conformation of molecule *B* is quite different from the two others in the region of the loops guarding the entrance to the active cleft. This modification involves, for molecule *B*, the destruction of a *β* ladder and the unwinding of one complete helix turn at the lower and the upper lip of this cleft, respectively. For this reason, the accessibilities of the four aza-tryptophan residues were computed independently for the three molecules. The accessible surfaces were computed using the program *X-PLOR* (Brünger, 1992), applying the algorithm of Lee & Richards (1971) and using a probe radius of 1.6 Å. The results are presented in Table 1. The values concerning isolated molecules (Table 1) clearly show that while the side chains of residues 17 are only partially accessible (12–25%), their $N^{\zeta 2}$ atoms are well exposed; the side chains of residue 73 are more accessible while residues 74 and 124 are completely buried. The location of the four residues in molecule *A* is illustrated in Fig. 2: the four Trn side chains have been added to a schematic ribbon diagram showing the folding of the molecule. The same diagram also shows the position of the

catalytic side chain (Glu19) in the active cleft between the two domains, at the C-terminal end of helix H1.

Trn17 is located in helix H1; in the three molecules, the side chain is clearly directed towards the surface of the molecule. The molecules in one column are oriented in such a way that this side chain is always facing in the direction of the large solvent channels between the columns. In spite of this favourable exposure to the solvent, no ordered water molecule connected to the $N^{\zeta 2}$ atom of the Trn17 residues has been identified. As a consequence, Trn17 does not take part in the improvement of the crystallization by specific contacts, although the lowering of its hydrophobicity compared with a natural tryptophan may be a favourable factor for a partially exposed residue.

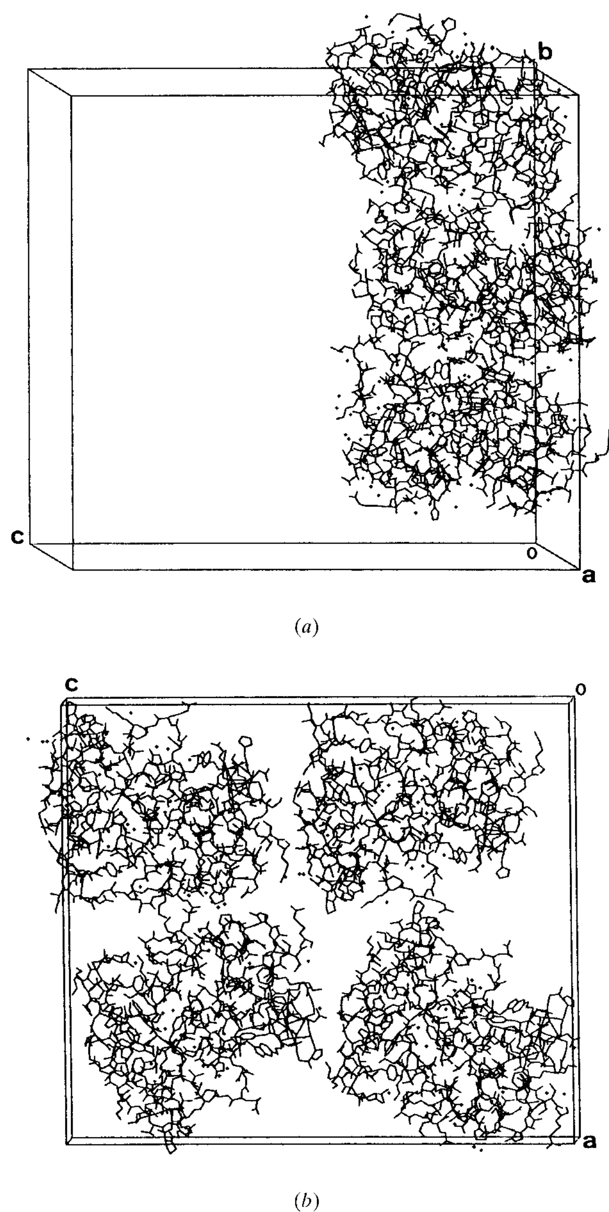


Figure 1
Packing of the molecules in the unit cell. (*a*) The disposition of the three independent molecules generates a column in the direction of the *b* axis. Close intermolecular contacts occur between the independent molecules in this direction. Molecule *A* is at the top, molecule *C* in the middle and molecule *B* at the bottom. (*b*) View along *b* of four molecules lying at approximately the same altitude. The packing is much looser and large solvent channels appear.

Trn73 and Trn74 are located in helix H2, close to the entrance to the catalytic cleft. Helix H2 makes the connection between the β-sheet structure in domain I and helix H3 which connects the two domains. The side chain of Trn73 is oriented outwards while that of Trn74 is completely buried, as shown in Table 1. In the interaction between molecules *A* and *C* in one column, the two clefts interpenetrate each other in a very symmetrical way, as previously described (Evrard *et al.*, 1998). Since Trn73 is located close to the entrance to this cleft, it can be expected that the packing between the two molecules will considerably modify the accessibility of this side chain in molecules *A* and *C*. Indeed, Table 1 shows that when the two molecules are considered together, the accessible surface of the $N^{\zeta 2}$ atoms is practically reduced to zero and the total accessibility of the residues is considerably lowered. However, we do not observe close intermolecular contacts (<3.5 Å) involving Trn73 $N^{\zeta 2}$ of molecules *A* or *C*. Nevertheless, a hydrogen bond (3.03 Å) exists between Trn73 $N^{\varepsilon 1}$ of molecule *A* and Ala130 O of molecule *C*. In the interaction between molecules *B* and *C*, a hydrogen bond (2.72 Å) is observed between Trn73 $N^{\zeta 2}$ of molecule *B* (acting as the acceptor) and Lys9 N^{ζ} of molecule *C* (acting as the donor) (Fig. 3). Again, the accessible surface of the Trn residue is lowered if molecules *B* and *C* are considered together (Table 1). Although Trn74 is completely buried in the three molecules and cannot take part in intermolecular contacts, it is interesting to note that in the three cases there is an intramolecular hydrogen bond between the $N^{\zeta 2}$ atom of this residue and Arg78 N^{ε} (Fig. 4). As a result, the polar head of Arg78 ($N^{\eta 1}$ and $N^{\eta 2}$) covers the aromatic side chain of the Trn residues and creates a shield between the $N^{\zeta 2}$ atoms and the solvent molecules. With a normal Trp residue, a less favourable situation would occur, since this hydrogen-bonded shield could not exist and the hydrophobic $C^{\zeta 2}$ atoms might become more exposed, as shown in Table 1.

Trn124 is located in the loop between helices H5 and H6. It is buried deep in the active-site cleft, close to the essential catalytic residue (Glu19). However, in the three independent molecules, there is a hydrogen bond with an ordered water molecule which seems to be held captive in the cavity. This water molecule is also hydrogen bonded to the side chain of Gln98 of the same molecule, in the helix H3 connecting the two domains. This intramolecular interaction, through a water molecule, connects sequentially distant residues belonging to different secondary-structural elements and may help to stabilize the conformation of the enzyme.

The analysis of the crystal structure thus reveals that the $N^{\zeta 2}$ atoms of the Trn residues are involved in different kinds of favourable contacts, which may play a role in the packing and conformation of the molecules and hence in the crystallization process. As might be expected, exposure to the solvent channels and close intermolecular contacts with adjacent molecules involve the residues exposed at the surface. More surprisingly, the buried residues take part in hydrophilic

Table 1
Accessible surface (Å^2) of the aza-tryptophan residues (Trn) for the three independent molecules in the absence of ordered solvent molecules.

The surfaces were computed by the program *X-PLOR* (Brünger, 1992) according to the method of Lee & Richards (1971), using a probe with a radius of 1.6 Å. For each molecule, the '$N^{\zeta 2}$' column shows the accessible surface of the $N^{\zeta 2}$ atom while the 'Side chain' column shows the sum of the accessible surfaces of all the atoms belonging to the side chain of the residue. The same quantities were computed for an isolated tripeptide with sequence Gly–Trn–Gly. The accessible surfaces are 12.7 Å^2 for $N^{\zeta 2}$ and 241.6 Å^2 for the Trn side chain.

	Molecule *A*		Molecule *B*		Molecule *C*	
	$N^{\zeta 2}$	Side chain	$N^{\zeta 2}$	Side chain	$N^{\zeta 2}$	Side chain
Isolated molecules						
Trn17	9.6	59.6	11.2	30.3	9.0	59.7
Trn73	6.8	107.0	10.8	44.6	10.6	119.9
Trn74	0.0	0.0	0.0	0.0	0.0	0.0
Trn124	0.0	0.7	0.0	0.6	0.0	0.5
Considering molecules *A* and *C* together†						
Trn73	0.0	72.5	—	—	1.3	71.9
Considering molecules *B* and *C* together†						
Trn73	—	—	0.9	18.0	—	—
Isolated molecules but removing Arg78 side chains						
Trn74	2.5	6.6	2.4	6.4	2.2	6.4

† Only the values affected by the intermolecular interaction are shown.

intramolecular contacts that could not occur with natural Trp residues.

The effect of replacement of tryptophans by aza-tryptophans on the pH dependence of the fluorescence intensity of the enzyme and on the stability towards denaturation has been investigated previously (Soumillion *et al.*, 1995). This led to an estimate of the number of inaccessible tryptophans of three (with a possibility of partial accessibility for one of them), in agreement with the present determination. The fact that the Trn-containing enzyme was nearly as stable as the wild type was considered surprising on the basis of the following argument. On folding, the $N^{\zeta 2}$ atoms lose a hydrogen bond to water and may not gain a similar compensating interaction in the folded state. In fact, as described above, hydrogen bonds are also formed in the folded state. This is probably made possible through local conformational adjustments of the Trn-containing enzyme compared with the wild type. These possible minor adjustments do not affect function, as the specific activity is unaffected. The observation that all the poorly accessible Trn residues can find suitable hydrogen-bonding partners suggests that protein destabilization on Trn incorporation, a potential limitation of the method, may not be a serious problem.

3.2. The modified TRP residue helps to solve the crystal structure

The structure of bacteriophage λ lysozyme is rather different from the other lysozymes for which three-dimensional coordinates are available (Evrard *et al.*, 1998). For this reason, it was not possible to solve the problem by molecular replacement, and the multiple isomorphous replacement (MIR) method had to be applied combined with averaging over the non-crystallographic symmetry (NCS).

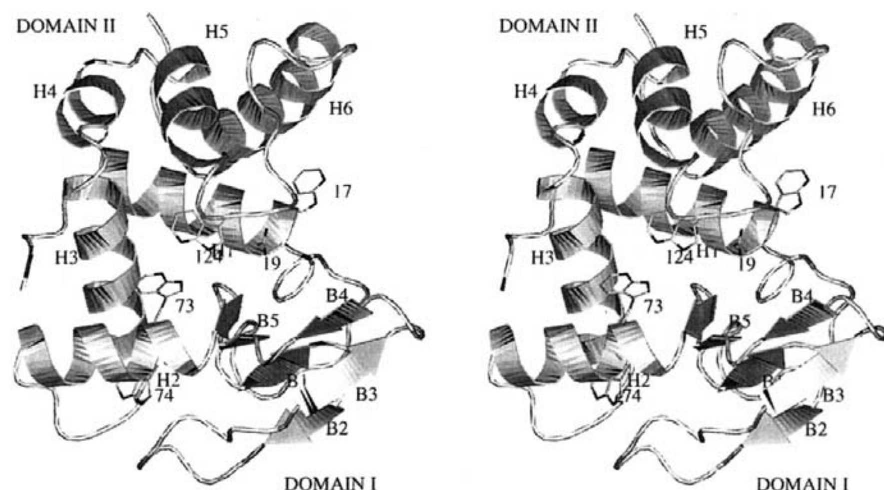


Figure 2
Stereoscopic view of a schematic diagram of molecule *A*. The side chains of the four aza-tryptophan residues (Trn) are represented as well as that of the essential catalytic residue (Glu19). The secondary elements are labelled (H, helix; B, β-strand). The labels which do not begin with a letter are residue numbers.

Four heavy-atom derivatives were used to solve the structure: two Pt derivatives (Pt1, sodium tetrachloroplatinate; Pt2, platinum terpyridine chloride) and two Hg derivatives (Hg1, mersalyl acid; Hg2, methyl mercury chloride). It has been shown in the previous section that the side chain of Trn17 is exposed in the large solvent channels between the columns of molecules. In both Pt compounds, heavy atoms are observed bound to the $N^{\zeta 2}$ atom of this residue, in spite of the presence of citrate buffer which is well known to be a chelating agent inhibiting the binding of the heavy

atom to the protein (Blundell & Johnson, 1976). In the Pt1 compound, there are a total of seven heavy-atom sites in the asymmetric unit and three of them, one in each independent molecule, correspond to the binding to this $N^{\zeta 2}$ atom. Compound Pt2 is more bulky; there are only three sites in total, one of which is the same $N^{\zeta 2}$ atom, but in molecule *A* only. Therefore, it appears that the replacement of Trp residues by Trn residues may be an efficient way of creating new heavy-atom binding sites. This property may be applicable to other proteins if exposed Trp residues are available.

Moreover, it appears that one of these new binding sites present in the three molecules was essential for solving the non-crystallographic symmetry.

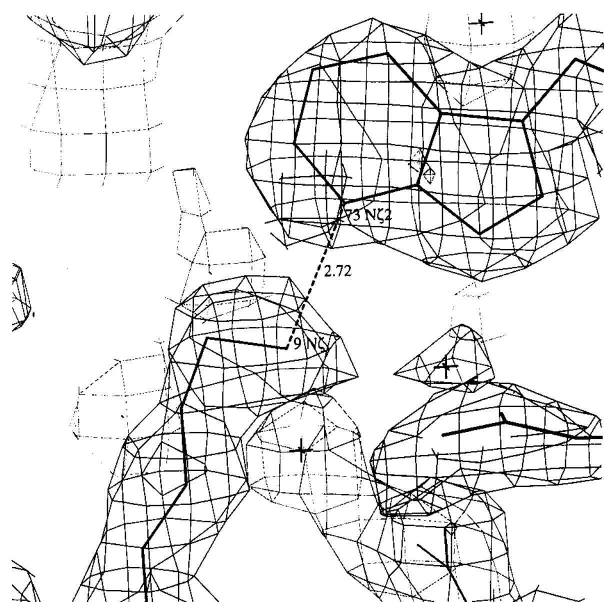


Figure 3
Close contacts between Trn73 $N^{\zeta 2}$ of molecule *B* and Lys9 N^{ζ} of molecule *C*, represented in the electron-density ($2F_o - F_c$) map contoured at a level of 1.0σ.

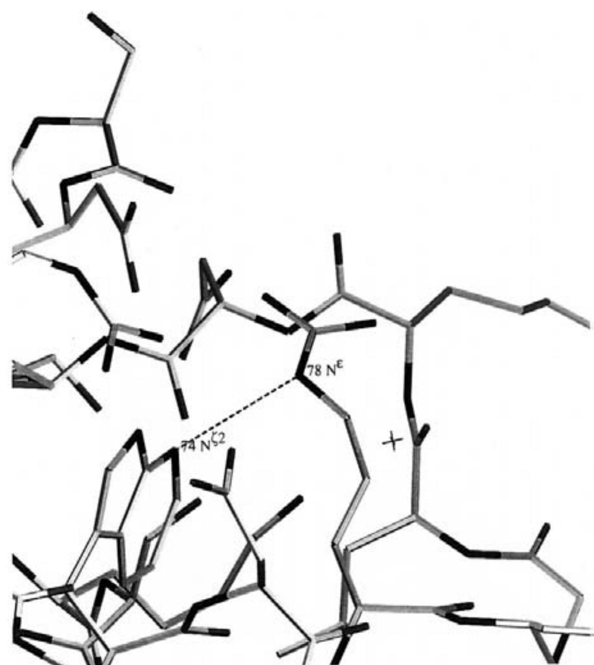


Figure 4
The buried side chain of the Trn74 residues (molecule *B* shown here) is protected from the solvent by an intramolecular hydrogen bond involving its $N^{\zeta 2}$ atom and Arg78 N^{ε} of the same molecule.

4. Concluding remarks

In the case of bacteriophage λ lysozyme, a threefold benefit has resulted from the replacement of the tryptophan residues by aza-tryptophan: firstly, it was possible to crystallize the modified enzyme, unlike the native one; secondly, new heavy-atom binding sites appeared, involving the new N atom of the aza-tryptophan side chain and, finally, one of these new binding sites was critical for solving the non-crystallographic symmetry and, as a consequence, for solving the crystallographic phase problem.

Although some of these benefits could be specific to this molecule, some other properties, such as the possibility of binding Pt derivatives, present the possibility of more general applicability.

The authors are indebted to the Fonds National de la Recherche Scientifique (Belgium) and to the Fonds de Développement Scientifique (Université Catholique de Louvain) for financial support. CE thanks the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture for a fellowship. The use of the EMBL X31 beamline at the DORIS storage ring, DESY, Hamburg is gratefully acknowledged. We thank the European Union for support of the work at EMBL Hamburg through the HCMP to Large Installations Project, contract No. CHGE-CT93-0040. We thank NASA and in particular Dr Daniel Carter for providing the opportunity of growing crystals in microgravity conditions.

References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. New York/London/San Francisco: Academic Press.
Brünger, A. T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT, USA.
Evrard, C., Declercq, J.-P. & Fastrez, J. (1997). *Acta Cryst.* D**53**, 217–219.
Evrard, C., Fastrez, J. & Declercq, J.-P. (1998). *J. Mol. Biol.* **276**, 151–164.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M. A., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G., Thomas, C. D., Shaw, W. V. & Harrisson, P. M. (1991). *Nature (London)*, **349**, 541–544.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

Martinezhackert, E., Harlocker, S., Inouye, M., Berman, H. M. & Stock, A. M. (1996). *Protein Sci.* **5**, 1429–1433.

Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 73–82.

Sivaraman, J., Coloumbe, R., Magny, M. C., Mason, P., Mort, J. S. & Cygler, M. (1996). *Acta Cryst.* D**52**, 874–875.

Soumillion, P. & Fastrez, J. (1992). *Biochem. J.* **286**, 187–191.

Soumillion, P., Jespers, L., Vervoort, J. & Fastrez, J. (1995). *Protein Eng.* **5**, 451–456.

Takeda, S., Yoshimura, H., Endo, S., Takahashi, T. & Nagayama, K. (1995). *Proteins Struct. Funct. Gen.* **23**, 548–556.

Thygesen, J., Krumbholz, S., Levin, I., Zaytzevbashan, A., Harms, J., Bartels, H., Schlunzen, F., Hansen, H. A. S., Bennett, W. S., Volkman, N., Agmon, I., Eisenstein, M., Dribin, A., Maltz, E., Sagi, I., Morlang, S., Fua, M., Franceschi, F., Weinstein, S., Boddekker, N., Sharon, R., Anagnostopoulos, K., Peretz, M., Geva, M., Berkovitchyellin, Z. & Yonath, A. (1996). *J. Cryst. Growth*, **168**, 308–323.