Acta Cryst. (1997). D53, 464–468

Structure of apo duck ovotransferrin: the structures of the N and C lobes are in the open form

AHMAD RAWAS, HILARY MUIRHEAD* AND JOHN WILLIAMS *at Department of Biochemistry and Molecular Recognition Centre, University of Bristol, Bristol BS8 1TD, England. E-mail: h.muirhead@bristol.ac.uk*

(Received 15 July 1996; accepted 20 January 1997)

Abstract

The structure of apo duck ovotransferrin (APODOT) has been determined at a resolution of 4.0 Å by the molecular replacement method using the structure of duck ovotransferrin (DOT) as the search model. The DOT structure contains two iron binding sites; one in the N-terminal lobe lying between domains N1 and N2 and one in the C-terminal lobe between domains C1 and C2. Both lobes have a closed structure. Models of various forms of both the N and C lobes were used in the search. The final model was refined to give an R factor of 0.22. The comparison of the structure of APODOT with that of DOT shows that both the N and the C lobes are in an open form, where the N2 and C2 domains undergo large rigid-body rotations of 51.6 and 49.9° relative to the N1 and C1 domains, respectively. The interface between the N and C lobes, which is formed by the N1—C1 contact in the core of the molecule does not change significantly. The DOT molecule may be described in terms of three rigid bodies; the N1 and C1 domains as one rigid body forming the static core of the molecule and the N2 and C2 domains as two other rigid bodies which, on the release of iron, move away from the static core of the molecule to form the open structure of APODOT.

1. Abbreviations

APODOT, apo duck ovotransferrin; APOHLT, apo human lactoferrin; DOT, duck ovotransferrin; HLT, human lactoferrin; HOT, hen ovotransferrin; RST, rabbit serum transferrin.

2. Introduction

Proteins of the transferrin family are monomeric glycoproteins of molecular weight 80 kDa. They consist of a single polypeptide chain folded into two homologous N and C lobes. Both the N and the C lobes contain about 330 amino acids and are made up of two domains N1, N2 and C1, C2 respectively, with the iron binding site situated in the interdomain cleft. Diferric transferrins have the ability to bind two Fe^{3+} ions concomitantly with two carbonate CO_3^{2-} anions (Brock, 1985; Crichton, 1991). The crystallographic structure determinations of human lactoferrin (Anderson *et al.*, 1987; Anderson, Baker, Norris, Rice & Baker, 1989), rabbit serum transferrin (Bailey *et al.*, 1988), the N-terminal half molecule of RST (Sarra, Garratt, Gorinsky, Jhoti & Lindley, 1990), the N-terminal half-molecule of hen ovotransferrin (Dewan, Mikami, Hirose & Sacchettini, 1993), HOT (Kurokawa, Mikami & Hirose, 1995), the mutant Asp60Ser of the N-terminal half-molecule of HLT (Faber *et al.*, 1996) and duck ovotransferrin (Rawas, Muirhead & Williams, 1996) have provided structural information regarding the organization of these proteins in the presence of iron, the environment of the iron binding sites, and the interdomain interactions. In the presence of iron these interdomain interactions result in the closed form of these proteins.

In the only single crystal structure of an apo transferrin (human apo lactoferrin, Norris, Anderson & Baker, 1991), the N2 domain moves substantially away from the N1 domain (rotation of 53°) to form an open N lobe, while apparently no conformational change takes place in the C lobe; this form was called 'one open, one closed'.

X-ray solution scattering studies of the intact proteins of human serum transferrin, human lactoferrin and hen ovotransferrin as well as the isolated N and C lobes of hen ovotransferrin (Grossmann *et al.*, 1992) provided an estimation of the opening of the lobes. These studies showed that both the N and the C lobes undergo similar conformational changes when iron is bound or released. The change occurring in the C lobe is significant and not as observed in the crystal structure of human apo lactoferrin (Norris *et al.*, 1991).

This paper describes the determination of the crystal structure of apo duck ovotransferrin with both lobes in the open form.

3. Experimental

3.1. *Protein crystallization*

Crystals of apo duck ovotransferrin were grown as described previously (Rawas, Moreton, Muirhead & Williams, 1989). Crystallization was performed either by vapour diffusion or by modification of the traditional batch method using quartz capillary tubes of diameter 1.5 mm. APODOT crystals were obtained at 277 K within a week, the crystals grown were orthorhombic, space group $P2_12_12_1$, with unit-cell dimensions $a = 77.6$, $b = 98.8$, $c = 127.0$ Å. There is one molecule of APODOT in the crystallographic asymmetric unit and the crystals have an approximate solvent content of 60% (Matthews, 1968).

3.2. *Data collection*

Data to 3.5 Å resolution were collected on the CCLRC X-ray synchrotron radiation source at Daresbury and used to solve the molecular replacement rotation and translation functions. Reflections were recorded on photographic film using an Arndt–Wonacott rotation camera. Five crystals were used to collect data to 3.5 Å at a wavelength of 1.488 Å. The films were scanned with an Optronics P1000 microdensitometer and the data processed using the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). The overall R_{sym} on intensities for the 3.5 Å data was 18%. The rapid deterioration of the crystals in the X-ray beam meant that the high-resolution data were poor and only data to 4 Å resolution with an overall R_{sym} of 9% were used (Table 1).

4. Results and discussion

The cross-rotation function was calculated by the real-space Patterson method with the program *X-PLOR* version 3.0

Table 1. *Data-processing statistics and refinement parameters for the final coordinates of APODOT using the program X-PLOR (Brünger, 1990)*

Resolution (Å)	4–10
Total number of observations	28770
Range of R_{sym} between 4 and 10 Å	0.12–0.054
Overall R_{sym}	0.09
Average $I/\sigma(I)$	10.2
Number of unique reflections with $I > \sigma$	7487
Percent completeness with $I > \sigma$ for;	
The total unique data	92.8
Resolution ranges 4.2–10.0 and 4.0–4.2 Å	96.0 and 50.0
Final R factor	0.21
Free R factor	0.32
No. of protein atoms	5300
R.m.s. deviation from ideal values	
Bond lengths (Å)	0.016
Bond angles (°)	3.881
Improper angles (°)	1.742
Dihedral angles (°)	25.40

Average B factor (Å^2) for each domain

Domain	Main-chain atoms	Side-chain atoms
N1	42.9	53.9
N2	38.3	49.3
C1	32.8	47.6
C2	38.8	49.2

(Brünger, 1990, 1992), using as search models the four separate domains N1, N2, C1 and C2, the individual N and C lobes, and the complete DOT molecule (Table 2). Each search model was placed in a large $P1$ cell; Patterson vectors of 30–45 Å, a resolution range of 10–4 Å and a map grid size of 0.3 Å were chosen. Lattman angle grids $\theta_1 + \theta_3$, θ_2 and $\theta_1 - \theta_3$ were searched with step sizes of $\Delta/\cos(\theta_2/2)$, Δ and $\Delta/\sin(\theta_2/2)$ respectively, where $\Delta = 0.25°$ (Lattman, 1985). The 6000 highest peaks of the rotation function were collected into clusters of width 10°, reduced to about 200 peaks and ordered by maximum height. The orientation parameters giving the highest peak in each cluster of the reduced list were refined using Patterson correlation refinement (PC refinement) and 15 cycles of rigid-body refinement were carried out for each selected peak and search model. The individual domains and lobes of the search model (Table 2) were then refined independently for a further 20 cycles.

The first calculation used the complete DOT molecule as the search model (Table 2, search model 1). The rotation search gave one plausible solution with a correlation coefficient of 0.083 (Fig. 1*a*). The translation function (correlation coefficient grid search) of *X-PLOR* (Brünger, 1990) was calculated for the highest peak from the PC refinement with a step of 1 Å. The most favourable solution, with a maximum value of the translation function of 0.141 (5σ above the mean and 1σ above the next highest peak), gave a reasonable overall crystal packing. However, there were a number of clashes between the N and C lobes of symmetry-related molecules and it became evident that the N and/or the C lobe must be in an open conformation.

A 'one open, one closed' molecular structure of DOT was constructed where the N2 domain of the DOT structure was rotated relative to the N1 domain in order to mimic the open N lobe of APOHLT (Norris *et al.*, 1991). The rotation and the translation searches were repeated using this model with (search model 2) and without (search model 3) the coordinates of domain C2. Both these search models gave the same solution (rotation and translation parameters) as search model 1. However, the correlation coefficient of 0.112 was higher (Fig. 1*b*) and the correlation coefficient grid search yielded a significant peak, with a maximum value of the translation function of 0.353, that is 17σ above the mean and 9σ above the next highest peak. Inspection of the crystal packing showed that some of the intermolecular clashes had been removed.

Since X-ray scattering studies in solution (Grossmann *et al.*, 1992) suggested that both lobes should be open, while the single crystal structure of APOHLT (Norris *et al.*, 1991) had the N lobe open and the C lobe closed, several search models were constructed with the N lobe open and different partial degrees of opening of the C lobe. These models gave the same molecular replacement solution without removing the clashes between the N1 domain and the C2 domain of the symmetry-related molecule at $(-\frac{1}{2} + x, \frac{1}{2} - y, 1 - z)$. In the final model a molecule was constructed with both the N and C lobes in the open form as found in the N lobe of APOHLT (search model 4).

The results of these molecular replacement calculations confirmed that, for the following reasons, both the N and C lobes must be open in APODOT and that N1/C1 forms a relatively rigid core with the second domains N2 and C2 hinging on this core.

Table 2. *Domains and lobes of DOT and APODOT and their superposition*

Domains and lobes of DOT and APODOT as used in the search models. Domains N1 and C1 represent the static core of APODOT and DOT.

	Domain	Residues
Lobe N	N1	1–91, 248–332
	N2	94–244
Lobe C	C1	345–431, 589–672
	C2	433–586

Search model (using DOT coordinates)	Number	Correlation coefficients	Figure
Whole molecule of DOT	1	0.083	1 (*a*)
N lobe open and C lobe closed	2	0.112	1 (*b*)
N lobe open with domain C1 only	3	0.112	1 (*b*)
Both N and C lobes open	4	0.138	1 (*c*)
N lobe open (C lobe omitted)	5	0.073	1 (*d*)
C lobe open (N lobe omitted)	6	0.080	1 (*d*)

Superposition of domains of APODOT onto the DOT structure using the *LSQKAB* program of the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

Rot1(°), Rot2(°) and Rot3(°) are the additional rotations required to superpose each domain and the core N1–C1 of APODOT on to the corresponding structure of DOT, after each of N1, C1, and N1–C1 of APODOT were superposed in turn on N1, C1 and N1–C1 of the DOT structure. rms1, rms2, and rms3 are the resulting r.m.s. deviations in Ca positions between equivalent structures in APODOT and DOT.

	DOT	N1		C1		N1–C1	
APODOT	Ca atoms	Rot1 (°)	rms1 (Å)	Rot2 (°)	rms2 (Å)	Rot3 (°)	rms3 (Å)
N1	175	—	—	5.3	0.26	2.6	0.26
N2	151	51.6	0.11	51.4	0.15	51.9	0.15
C1	171	5.9	0.27	—	—	2.8	0.15
C2	154	53.3	0.13	49.9	0.22	51.4	0.26
N1–C1	346	2.9	0.68	2.8	0.68	—	—

(1) Higher peak-to-background ratios were obtained in the molecular replacement solution when using the complete structure in the open form (Fig. 1*c*).

(2) When the cross-rotation and rigid-body PC refinements were performed using the isolated open N and C lobes as search models 5 and 6 (half the molecule of APODOT in each search) a single significant peak was present for each model (Fig. 1*d*). Searching with either the N or the C lobe in the closed form did not give any significant peaks.

(3) A plausible packing with no clashes in the atomic coordinates had been achieved.

In order to provide a further confirmation of the conclusion that both lobes were open an omit electron-density map of $|F_o| - |F_c|$ was calculated omitting the coordinates of domain C2. This omit electron-density map revealed the correct position of a number of the regular secondary-structure elements in the C2 domain (*e.g.* helix H5 in domain C2).

The structure of APODOT with both the N and the C lobes in the open form was refined using *X-PLOR* rigid-body refinement and treating the individual domains N1, N2, C1, C2, as four rigid bodies. This refinement reduced the R factor from 0.40 to 0.37 with a corresponding drop in R_{free} (falling from 0.37 to 0.35 for a 10% excluded data set; R_{free} reflections were selected using the *CCP*4 program *FREERFLAGR*). Further energy minimizations of 40 steps of positional (*xyz*) refinement followed by ten steps of grouped B-factor refinement (two B-factor groups for each residue, one for backbone and one for side-chain atoms) led to a drop in the R factor to 0.21 and R_{free} to 0.32. Table 1 lists the final geometrical parameters and B factors after the refinement. The refinement of the APODOT structure was performed using 10–4 Å data and according to the example given in the *X-PLOR* V3.0 manual (Brünger, 1992).

A superposition of the individual domains of APODOT on to the DOT structure (Table 2) shows that the main differences between the two structures lie in the orientations of domains N2 and C2 relative to the N1 and C1 domains, respectively (Fig. 2). The relative orientation of domains N1 and C1 in APODOT and the interlobe interactions are almost the same as in the holo form of DOT. These domains represent the static core of the molecule; this explains why a correct solution was found in the molecular replacement search when using the whole molecule of DOT since the relative orientation of the N1 and C1 domains (half the molecule) is almost the same in APODOT and DOT. The search models using either individual domains, or individual lobes in the closed form, contained only 25% of the asymmetric unit-cell content in the correct orientation for any one rotation. This was not sufficient information to determine the correct solution.

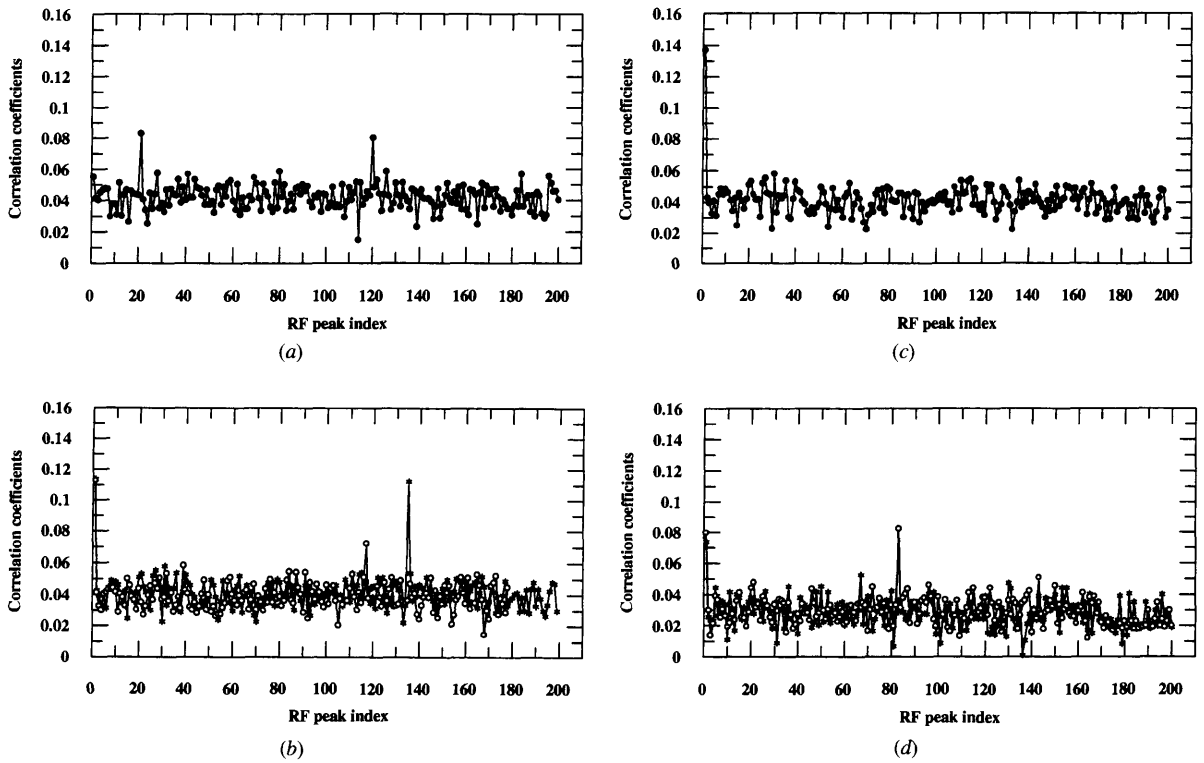

Fig. 1. Rigid-body Patterson correlatin refinement for the selected rotation-function peaks, using the structure of DOT with the following search models. (*a*) Complete molecule with both N and C lobes in the closed form (as in DOT); the two highest peaks are at equivalent positions (search model 1). (*b*) the N lobe is open as in APOHLT and the C lobe is close as in DOT (search model 2, indicated by ○); the N lobe is in the open form plus domain C1 only (search model 3, indicated by *), the two highest peaks 1 and 135 are related by 180°. (*c*) The complete molecule of APODOT, the N and the C lobes are in the open form (search model 4). (*d*) The isolated N lobe in the open form (search model 5, indicated by *) and the isolated C lobe in the open form (search model 6, indicated by ○). Peaks 1 and 83 for the C lobe are related by 180°.

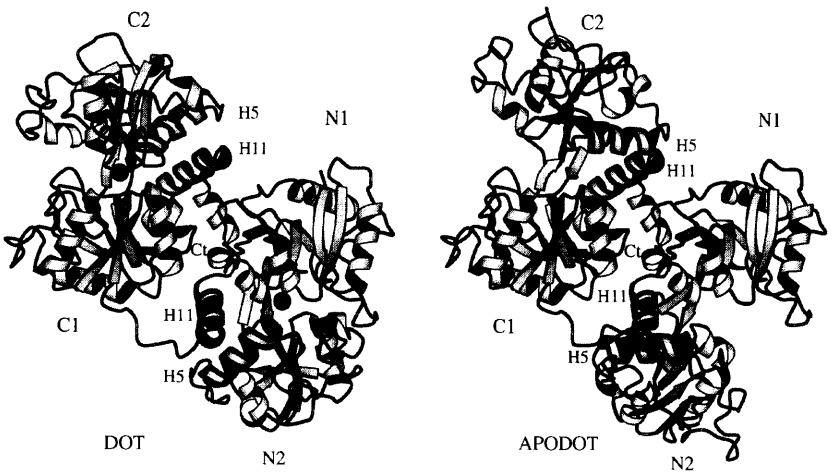

Fig. 2. Ribbon representations of DOT (left) an APODOT (right) drawn with the program *MOLSCRIPT* (Kraulis, 1991). The positions of the two Fe^{3+} ions in DOT are indicatd by surface-filled circles in the cleft between the two domains N1, N2 in the N lobe and C1, C2 in the C lobe; Ct indicates the carboxyl terminus. In APODOT, in the absence of bound Fe^{3+} ions, both lobes are in the open form. Helices 5 (N2 and C2) and 11 (N1 and C1) are labelled.

The similar large changes in the orientation of the N2 and C2 domains of 51.6 and 49.9°, respectively (Table 2) confirm the conclusions of the X-ray solution scattering (Grossmann *et al.*, 1992). This indicated that both the N and the C lobes are open in the apo structure and that the opening of the C lobe is equivalent to 75% of the opening of the N lobe but did not rule out a wider opening of the C lobe in solution similar to that observed for the N lobe in the crystal structure of APOHLT. We have shown that in APODOT both lobes are fully open. In the DOT structure (Rawas *et al.*, 1996) helices 5 (residues 122–135 in domain N2, residues 462–475 in domain C2) and 11 (residues 322–332 in domain N1, residues 659–671 in domain C1) are not in contact. In the APODOT structure, helix 5 moves with domain N2 (or domain C2), while helix 11, remains fixed with the N1 domain (C1 domain); a similar movement was observed in APOHLT (Gerstein *et al.*, 1993). As a result of this movement hydrogen bonds and van der Waals interactions are formed between helix 5 and helix 11 in both the N and C lobes in APODOT. The interactions on the interface between domains N1 and N2 (and C1 and C2) in DOT are lost in the APODOT structure; the protein ligands for Fe^{3+} and CO_3^{2-} Tyr191 (Tyr524) and Arg121 (Arg460) moving away with domain N2 (C2) from the iron binding site by about 9 Å in the APODOT structure. The interaction between NZ atoms of Lys209 and Lys301, which are only 2.62 Å apart in the N lobe of DOT (Rawas *et al.*, 1996), is lost and the distance between the NZ atoms increased to 8.14 Å.

The equilibrium between the open and closed forms of the C lobe in the apo structure seems to be influenced by the crystallization conditions and the crystal packing forces. The solvent contents are 48% in DOT and 60% in APODOT (Rawas *et al.*, 1989) and 54% in both HLT and APOHLT (Baker & Rumball, 1977; Norris, Baker & Baker, 1989). The tighter packing in APOHLT is sufficient to prevent the opening of both lobes. This suggests that the energy barrier between the open and closed forms of the apo structure must be small and the equilibrium between states is influenced by crystal packing forces. The high solvent content of 60 and 63% of the other two crystal forms of APOHLT (Norris *et al.*, 1989) and the low resolution of the diffraction patterns may indicate that both the N and the C lobes are in the open form in these crystals. In the diferric form the equilibrium is strongly in favour of the closed form and is uninfluenced by crystal packing.*

We thank E. N. Baker and his colleagues at Massey University, New Zealand, for providing the atomic coordinates of apo lactoferrin. We are grateful to the CCLRC for funding and for synchrotron time at the SRS, Daresbury, England.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1AOV, R1AOVSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0661).

References

Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M., & Baker, E. N. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 1768–1774.

Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. E. & Baker, E. N. (1989). *J. Mol. Biol.* **209**, 711–734.

Bailey, S., Evans, R. W., Garratt, R. C., Gorinsky, B., Hasnain, S., Horsburgh, C., Jhoti, H., Lindley, P. F., Mydin, A., Sarra, R. & Watson, L. (1988). *Biochemistry*, **27**, 5804–5812.

Baker, E. N. & Rumball, S. V. (1977). *J. Mol. Biol.* **111**, 207–210.

Brock, J. H. (1985). *Metalloproteins*, Part II, edited by P. Harrison, pp. 183–262. London: Macmillan.

Brünger, A. T. (1990). *Acta Cryst.* A**46**, 46–57

Brünger, A. T. (1992). *X-PLOR Manual.* Version 3.0. Yale University, New Haven, Connecticut, USA.

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

Crichton, R. R. (1991). *Chemistry and Biology of the Transferrins.* In *Inorganic Biochemistry of Iron Metabolism*, pp. 101–119, Ellis Horwood Series in Inorganic Chemistry. Chichester: Ellis Horwood.

Dewan, J. C., Mikami, B., Hirose, M. & Sacchettini, J. C. (1993). *Biochemistry*, **32**, 11963–11968.
Faber, H. R., Bland, T., Day, C. L., Norris, G. E., Tweedie, J. W. & Baker, E. N. (1996). *J. Mol. Biol.* **256**, 352–363.
Gerstein, M., Anderson, B. F., Norris, G. E., Baker, E. N., Lesk, A. M. & Chothia, C. (1993). *J. Mol. Biol.* **234**, 357–372.
Grossmann, J. G., Neu, M., Pantos, E., Schwab, F. J., Evans, R. W., Townes-Andrews, E., Lindley, P. F., Appel, H., Thies, W.-G. & Hasnain, S. (1992). *J. Mol. Biol.* **225**, 811–819.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kurokawa, H., Mikami, B. & Hirose, M. (1995). *J. Mol. Biol.* **254**, 196–207.
Lattman, E. E. (1985). *Methods Enzymol.* **115**, 55–77.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Norris, G. E., Anderson, B. F. & Baker, E. N. (1991). *Acta Cryst.* B**47**, 988–1004.
Norris, G. E., Baker, H. M. & Baker, E. N. (1989). *J. Mol. Biol.* **209**, 329–331.
Rawas, A., Moreton, K., Muirhead, H. & Williams, J. (1989). *J. Mol. Biol.* **208**, 213–214
Rawas, A., Muirhead, H. & Williams, J. (1996). *Acta Cryst.* D**52**, 631–640.
Sarra, R., Garratt, R. C., Gorinsky, B., Jhoti, H. & Lindley, P. F. (1990) *Acta Cryst.* B**46**, 763–771.