Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Thermolysin in the absence of substrate has an open conformation

Andrew C. Hausrath† and Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, 1229 University of Oregon, Eugene, OR 97403-1229, USA

† Present address: Department of Biochemistry and Molecular Biophysics, University of Arizona, 1007 E. Lowell, Tucson, AZ 85721, USA.

Correspondence e-mail: brian@uoxray.uoregon.edu

Received 28 November 2001
Accepted 2 April 2002

PDB Reference: tetragonal thermolysin, 1l3f, r1l3fsf.

The bacterial neutral proteases have been proposed to undergo hinge-bending during their catalytic cycle. However, in thermolysin, the prototypical member of the family, no significant conformational change has been observed. The structure of thermolysin has now been determined in a new crystal form that for the first time shows the enzyme in the absence of a ligand bound in the active site. This is shown to be an 'open' form of the enzyme. The relative orientation of the two domains that define the active-site cleft differ by a 5° rotation relative to their positions in the previously studied ligand-bound 'closed' form. Based on structural comparisons, kinetic studies on mutants and molecular-dynamics simulations, Gly78 and Gly135–Gly136 have previously been suggested as two possible hinge regions. Comparison of the 'open' and 'closed' structures suggests that neither of the proposed hinge regions completely accounts for the observed displacement. The concerted movement of a group of side chains suggested to be associated with the hinge-bending motion is, however, confirmed.

1. Introduction

There are a number of examples indicating that enzymes are dynamic and that conformational change can be an essential component of catalytic activity (Bennett & Huber, 1984; Ringe & Petsko, 1985; Dobson, 1993; Wagner, 1995).

Based on comparison of the structures of several members of the bacterial neutral protease family, Holland *et al.* (1992) proposed that hinge-bending might play a role in their catalytic mechanism. Members of the family with substrate analogs bound in their active sites were observed to display 'closed' conformations, while those that had unoccupied active sites were seen to have more 'open' conformations. Thermolysin itself, however, which is the prototypical member of the family, has only been observed in the closed form. Even the structure of the 'native' enzyme (Matthews *et al.*, 1972) was subsequently found to have electron density in the active site which was interpreted to be the bound dipeptide Val-Lys (Holland *et al.*, 1992). This, therefore, was inferred to correspond to the 'closed' form of the enzyme, in common with the many inhibitor complexes that have been described (Matthews, 1988).

It was inferred that if a non-bound form of thermolysin could be obtained, its structure would be in an 'open' conformation. We here describe the structure of the enzyme in a new crystal form in which the active site appears to be unoccupied and, as anticipated, the conformation is open.

2. Materials and methods

100 mg of 3× recrystallized thermolysin (Calbiochem) was dissolved in 1 ml of 50 m*M* 2-*N*-morpholinoethanesulfonic acid (MES) pH 6.0, 45%(*v*/*v*) dimethylsulfoxide (DMSO) by gently rocking at room temperature for 1 h. Insoluble material was removed by centrifugation and the clarified supernatant was used for crystallization trials. All trials were conducted at room temperature. 2 µl of protein solution was mixed with 2 µl of 50 m*M* MES pH 6.0, 0–2.2 *M* NaCl, 0–0.4 *M* zinc acetate and 45% DMSO in sitting drops over a well containing 0.5 ml of 30–50% saturated ammonium sulfate. Large hexagonal crystals (0.6 × 0.6 × 1.0 mm) grew in 1–3 d in drops initially containing 1.0–1.1 *M* NaCl, but no zinc acetate, over 35% ammonium sulfate. Large tetragonal crystals (0.5 × 0.5 × 0.5 mm) grew in 4–5 d in drops initially containing 0.7–0.9 *M* NaCl, 0.4 *M* zinc acetate over wells containing 30% ammonium sulfate.

The hexagonal crystals have space group $P6_122$ and are the same as those used for the original structure determination (Matthews *et al.*, 1972). The tetragonal crystals have unit-cell parameters $a = b = 97.05$, $c = 106.52$ Å. Precession photos established that the space group was $P4_12_12$ or $P4_32_12$.

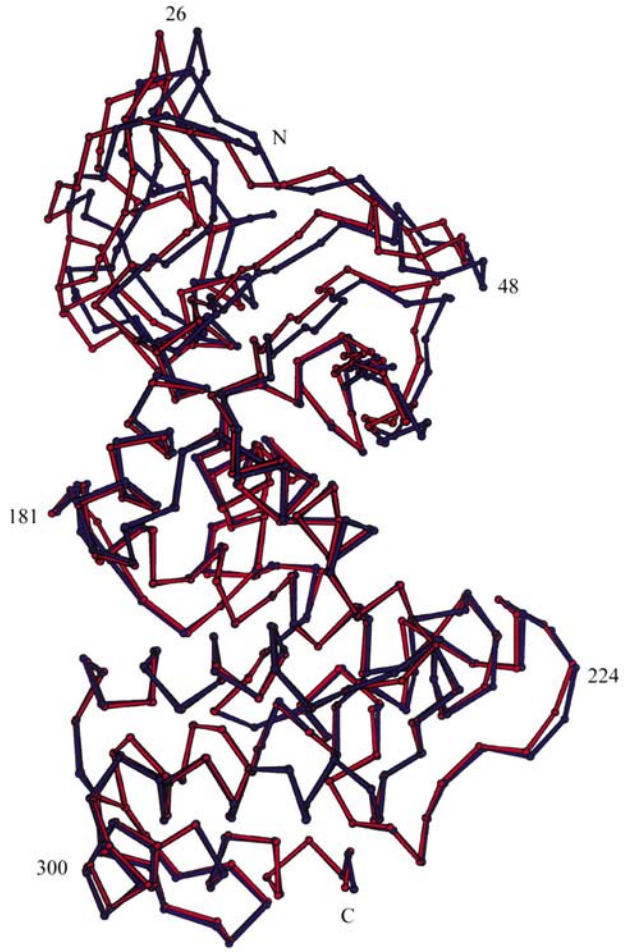

Figure 1
Superposition of the backbone structures of the 'closed' (blue) and 'open' (red) forms of thermolysin. The structures are superimposed based on the C-terminal domain (residues 78–90, 135–194 and 200–316; Holland *et al.*, 1992).

Table 1
X-ray data collection and refinement statistics.

Values in parentheses correspond to the outermost data shell (2.36–2.30 Å).

Resolution range (Å)	25.0–2.3
No. of observations	205247
No. of independent reflections	22538
Completeness (%)	96.7 (79.2)
Multiplicity	9.1 (3.5)
$I/\sigma(I)$	4.4 (1.4)
R_{sym} (%)	9.6 (54)
R factor (%)	20.2
R_{free} (%)	30.2
No. of protein atoms	2427
No. of zinc/calcium ions	3/4
Number of waters	127
Δ_{bonds} (Å)	0.016
$\Delta_{bond\ angles}$ (°)	3.2
$\Delta_{planar\ groups}$ (°)	0.021

Crystals were transferred by pipette to a shallow dish containing mineral oil and a fine glass capillary was used to draw away the liquid from the surface of the crystal. The crystals could then be captured in a nylon loop and rapidly cooled to 103 K in a stream of liquid-nitrogen vapor for data collection.

Data were collected on beamline 9-2 at SSRL using a wavelength of 0.9778 Å and were processed with *MOSFLM* (Leslie, 1991) and *SCALA* (Evans, 1991). 5% of the data were reserved for R_{free} calculations. Data statistics are shown in Table 1. The structure was solved with *AMoRe* (Navaza, 1994) using protein atoms from PDB entry 8tln as the search model. The model was refined using *TNT* (Tronrud, 1997). Rigid-body refinement of the molecular-replacement solution resulted in an R factor of 42.3% and R_{free} of 42%. On separating the model into two rigid domains as defined by Holland *et al.* (1992), the R factor decreased to 35.9% and an R_{free} to 35.5%. The refinement was started with the B values from the hexagonal form (average of 21 Å^2). These values were clearly too low and the values were then allowed to rise as a single unit, preserving the relative values. In the latter stages the B factors were refined with the correlated B-factor option of *TNT* (Tronrud, 1996). The weight was adjusted so that the resulting B-factor discrepancy (6.3 Å^2) was approximately equal to the discrepancy of the B-factor correlation restraint library (6.0 Å^2; Tronrud, 1996). After multiple rounds of model-building and refinement, the R factor converged at 20.2% and R_{free} was 30.2%. Refinement statistics are included in Table 1. We attribute the relatively high value of R_{free} to a combination of factors. Firstly, the solvent content of the crystals is fairly high (65%) and the Wilson B factor for the observed structure amplitudes is also high (64 Å^2). The average B factor for the refined structure is 72.1 Å^2 (71.8 Å^2 for the protein atoms and 77.3 Å^2 for the solvent). These values all suggest that the molecules in this crystal form are not well packed. The atoms that are closest to the hinge axis have average B factors of about 64 Å^2, whereas the average value for those that are furthest away is about 68 Å^2. This

suggests that there may be some hinge-bending motion within this crystal form. This would not be directly accounted for in the refinement and might also contribute to the observed value of R_{free}.

98.6% of the residues are in the allowed region of the Ramachandran map, 1.1% are in the generously allowed region and 0.4% are in the disallowed region (Laskowski *et al.*, 1993). The single disallowed residue is Thr26, which forms a γ-turn (Matthews, 1972) as in the hexagonal structure.

3. Results

3.1. Hinge-bending

The primary difference between the previously reported structure of thermolysin in space group $P6_122$ and the present study is that the N- and C-terminal domains have rotated relative to one another by 5.1°. Fig. 1 shows a superposition, based on the C-terminal domain, of the closed structure (from PDB entry 8tln) and the open structure described in this work. The structural change is nearly a pure rotation, with a translational component of only 0.08 Å.

The result is an opening of the active-site cleft. The backbones on opposite sides of the cleft move apart by 0.7 Å. Residues farthest from the rotation axis move approximately 3 Å.

3.2. Active site

The map showing the electron density within the active site of the new crystal form is shown in Fig. 2. There is a very strong feature of height over 12σ that appears to correspond to an additional zinc ion. It lies 3.2 Å from the catalytic zinc and is liganded by His231. This is the same position as observed in a previous study of thermolysin in the presence of excess zinc (Holland *et al.*, 1992).

No electron density is seen at the position corresponding to the dipeptide Val-Lys that is observed in the closed form of the enzyme, suggesting that it is not bound in the present crystal structure.

There is negative density superimposed on the side chain of Tyr157. No corresponding positive density is seen nearby to suggest an alternative position for this side chain. The simplest interpretation is that this side chain has become more mobile or disordered in the open crystal form.

3.3. Cascade of side chains

Accompanying the hinge-bending motion is the rearrangement of a series of residues reaching from near the axis of rotation to the active site (Fig. 3). The side chains of Leu144 and Met120 change to alternative rotamers and the

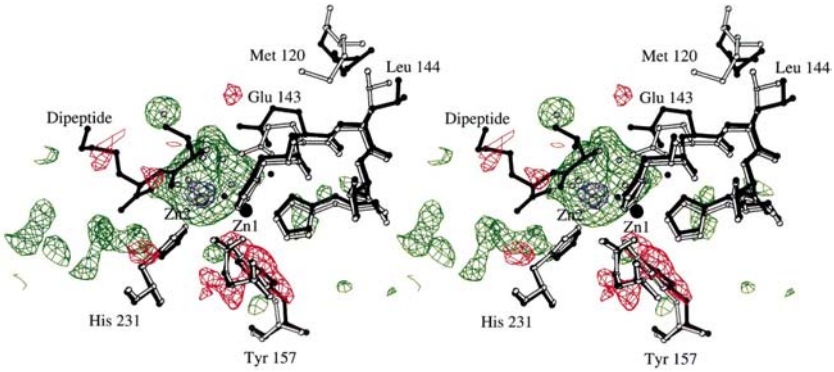

Figure 2
Omit map showing the electron density within the active site attributed to a second zinc ion plus associated water molecules. The coefficients are $(F_o - F_c)$, where F_o is the observed structure amplitude and F_c is the calculated structure amplitude, and the phases were calculated from the refined model with the catalytic zinc (Zn1), the second zinc (Zn2) and the nearby water molecules removed. The resolution is 2.3 Å. Positive density is contoured at 3σ (green) and 12σ (blue), while negative density is contoured at -3σ (red). The model of the closed form of thermolysin plus the dipeptide bound in the active site is drawn in black. The model for the open form is colored gray. The models are superimposed based on the C-terminal domains of the respective proteins.

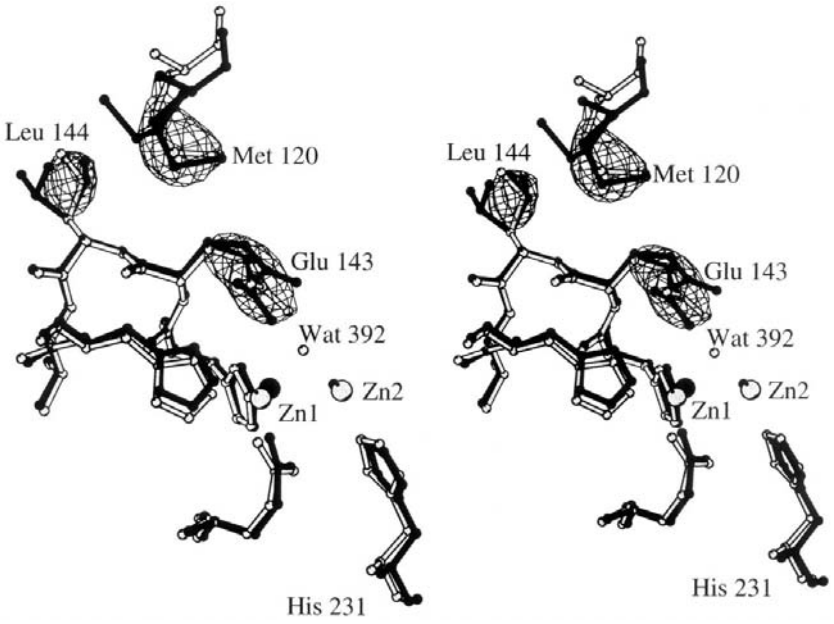

Figure 3
Omit map showing the electron density for the side chains of Met120, Glu143 and Leu144. These side chains move in going from the closed to the open form. They were also observed to occupy two alternative positions in a previous study of 'closed' thermolysin substituted with cadmium (Holland *et al.*, 1995). The coefficients for the map are $(F_o - F_c)$, where F_o is the observed structure amplitude and F_c is the calculated structure amplitude, and the phases are calculated from the refined model with the side chains of Met120, Glu143 and Leu144 deleted. The map is to 2.3 Å resolution and contoured at 5σ. The refined model in the open form is shaded gray. The refined model for the closed form, in black, is for the cadmium-substituted enzyme with the alternative positions shown for the side chains of Met120, Glu143 and Leu144.

catalytic Glu143 moves deeper into the active site. Movement in these residues was observed in a previous study of cadmium- and manganese-substituted thermolysins (see Holland *et al.*, 1995).

4. Discussion

4.1. Open and closed forms

Based on comparisons of three members of the bacterial neutral protease family, thermolysin from *Bacillus thermoproteoliticus* (Matthews *et al.*, 1972), the neutral protease from *B. cereus* (Pauptit *et al.*, 1988; Stark *et al.*, 1992) and the elastase from *Pseudomonas aeruginosa* (Thayer *et al.*, 1991; McKay *et al.*, 1992), Holland *et al.* (1992) suggested that these family members undergo hinge-bending motion during catalysis. At that time the structure of thermolysin in a non-ligand-bound form was not known, although it was presumed to be 'open'. The present work supports these earlier proposals.

In the case of the elastase from *P. aeruginosa* both the open and closed forms of the enzyme have been characterized and shown to differ by a change in hinge-bending angle of 16° (Holland *et al.*, 1992). The neutral protease from *B. cereus* is known only in the open form. Its amino-acid sequence is closely related to that of thermolysin (73% sequence identity) and the respective structures are presumed to be very similar. It was noted by Holland *et al.* (1992) that the change in the hinge-bending angle between the closed form of thermolysin and the open form of the neutral protease was 6°. This is very close to the 5° rotation now seen between the closed and open forms of thermolysin itself.

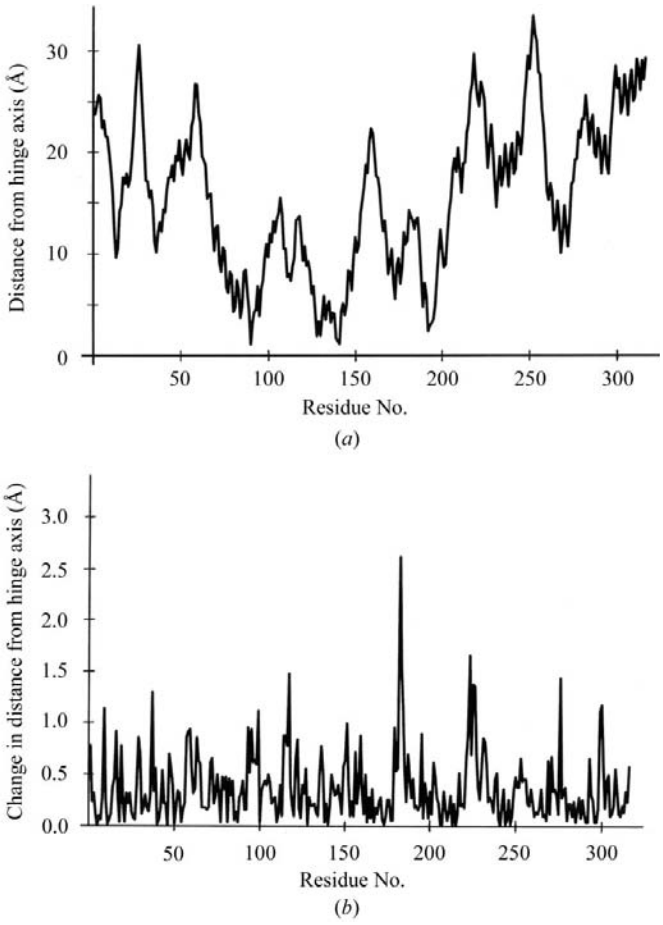

Figure 4
(*a*) Plot showing the distance of each C^{α} atom in the closed form of thermolysin from the hinge-bending axis. The axis was defined by superimposing the C-terminal domains of the open and closed forms of the enzyme (as in Fig. 1) and then using *EDPDB* (Zhang & Matthews, 1995) to determine the axis of rotation that relates the N-terminal domains. In this calculation, the N-terminal domain was taken to include residues 1–77, 92–134 and 195–199, while the C-terminal domain was comprised of residues 78–90, 135–194 and 200–316. (*b*) Magnitude of the difference in distances from the hinge-bending axis between the open and closed forms of thermolysin.

4.2. Nature of the hinge-bending motion

If the open and closed forms of thermolysin are superimposed, the root-mean-square difference between the 316 C^{α} atoms is 0.62 Å. For the 126 C^{α} atoms in the N-terminal domain considered in isolation, the discrepancy is 0.43 Å. For the 190 C^{α} atoms in the C-terminal domain the discrepancy is 0.37 Å. The estimated uncertainty in the C^{α} coordinates for the closed form of thermolysin determined to 1.6 Å resolution is about 0.15 Å. For the open form, however, determined to 2.3 Å resolution, the uncertainty is about 0.3 Å. This means that care has to be taken not to overinterpret small differences between the respective structures.

Fig. 4(*a*) shows the distance of each residue in thermolysin from the hinge-bending axis. It is reasonable to expect that the locations at which the hinge-bending occurs are likely to be located close to the hinge-bending axis. On this basis, the most likely regions are residues 80–95, 125–145 and 190–195. Several groups have suggested on the basis of structure comparisons (Stark *et al.*, 1992), inhibitor-induced structural changes (Thayer *et al.*, 1991) and molecular dynamics (van Aalten *et al.*, 1995) that the most likely location for hinge-bending is at sites 135 and 136, which are relatively conserved glycine residues. These residues are located in an α-helix (residues 133–151) that is in the waist of the molecule, at the bottom of the active-site cleft, and contains several of the catalytically important residues. Veltman *et al.* (1998) substituted Gly135 and Gly136 in a thermolysin-like protease with alanine and in each case observed a substantial reduction in catalytic efficiency.

If the hinge-bending motion in thermolysin consists of a strict rigid-body rotation, then the distance of any given residue from the axis of rotation should remain constant. To search for possible departures from such rigid-body movement we plotted, for each residue, the difference between the residue-to-hinge distance in the open and closed forms (Fig. 4*b*). Within experimental error the motion does seem to be essentially rigid-body, with the possible exception of the region near residue 180. This region is at the end of a second α-helix (residues 160–180) that also extends across the waist of

the thermolysin molecule. It is close to one of the putative hinge regions suggested in Fig. 4(*a*).

Another way to try to identify putative hinge-bending sites is to compare the 'open' and 'closed' molecules with a sliding window of, for example, 15 residues (Remington & Matthews, 1980; Birktoft *et al.*, 1989; Dixon *et al.*, 1992). This is shown in Fig. 5. There is a rather pronounced peak near residues 175–185, *i.e.* in the same region as seen in Fig. 4(*b*). This can be taken as further evidence for hinge-bending near residue 180.

At the same time, however, it should be noted that the open crystal form of thermolysin has a crystal contact at residues 183–184. Perhaps this could cause a local change in conformation of the backbone that would contribute to the peak near residue 180 seen in Figs. 4(*b*) and 5.

It should also be emphasized that the lack of an obvious peak near Gly135–Gly136 in Figs. 4(*b*) and 5 does not preclude this region being a site of hinge motion. Likewise, Gly78, which was also suggested as a possible site of hinge-bending (Holland *et al.*, 1992), is not precluded. If hinge motion occurred at one or other of these sites and consisted of a simple rotation about one or two backbone peptide bonds it would not necessarily cause a change in conformation that could be seen in either Fig. 4(*b*) or Fig. 5. Thus, while the structural analysis shows clearly that thermolysin does undergo a hinge-bending motion, the specific location(s) at which the hinge motion occur remain somewhat uncertain.

4.3. Side-chain movement associated with hinge-bending

Although, as noted above, the hinge-bending movement is essentially rigid-body in character, there are three side chains in the hinge region that undergo distinct changes in conformation in going from the closed to the open form (Fig. 3). These same residues, Met120, Glu143 and Leu144, were previously seen to undergo related changes in conformation when thermolysin was exposed to cadmium, manganese or excess zinc (Holland *et al.*, 1995). In the hexagonal crystal form of thermolysin these three side chains appear to have rotomer angles that correspond to the 'closed' form of the enzyme. In the present (tetragonal) crystal form they appear to have strictly the 'open' conformations. In the zinc-bound hexagonal crystals the side chains were observed to have both the 'open' and 'closed' conformations. It does not appear that the binding of excess zinc causes *per se* the changes in these side chains, since similar changes were seen in the cadmium-substituted enzyme which does not have a second metal ion at the active site (Holland *et al.*, 1995). By the same token, since the cadmium-substituted structure was not in the open form, but did show the side-chain arrangements to some extent, it appears that these rearrangements are not caused by the crystal contacts in the open form.

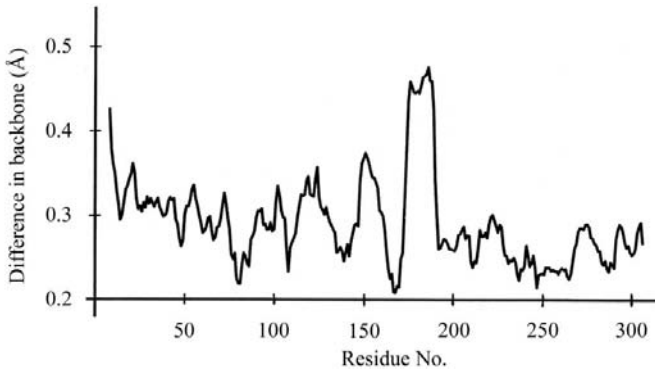

Figure 5
'Sliding window' comparison of the backbones of the open and closed forms of thermolysin. For each residue the value plotted is the root-mean-square difference between the open and closed forms calculated for a 15-residue segment of backbone centered on the residue in question.

Based on comparisons of thermolysin with the related structures of *B. cereus* neutral protease (Pauptit *et al.*, 1988; Stark *et al.*, 1992) and *P. aeruginosa* elastase (Thayer *et al.*, 1991; McKay *et al.*, 1992), it was suggested that the cadmium-substituted or related forms of thermolysin represented, at least to some degree, the 'open' substrate-free form of the enzyme (Holland *et al.*, 1995). It was also suggested that hinge-bending in thermolysin would be accompanied by conformational changes in Met120, Glu143 and Leu144. Both these predictions are supported by the present structure determination. It might be noted that the concerted movement of a group of side chains close to the hinge-bending axis of a protein is not without precedent. A very similar situation was observed in variants of T4 lysozyme with different hinge-bending angles (Zhang *et al.*, 1995).

We thank the staff at the Stanford Synchrotron Research Laboratory for help with data collection and Dr Dale Tronrud for helpful discussions. This work was supported in part by NIH grant GM20066 to BWM.

References

Aalten, D. M. van, Amadei, A., Linssen, A. B., Eijsink, V. G., Vriend, G. & Berendsen, H. J. (1995). *Proteins*, **22**, 45–54.
Bennett, W. S. & Huber, R. (1984). *Crit. Rev. Biochem.* **15**, 291–384.
Birktoft, J. J., Rhodes, G. & Banaszak, L. L. (1989). *Biochemistry*, **28**, 6065–6081.
Dixon, M. M., Nicholson, H., Shewchuk, L., Baase, W. A. & Matthews, B. W. (1992). *J. Mol. Biol.* **227**, 917–933.
Dobson, C. M. (1993). *Curr. Biol.* **3**, 530–532.
Evans, P. R. (1991). *Crystallographic Computing 5. From Chemistry to Biology*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 136–144. Oxford: IUCr/Oxford University Press.
Holland, D. R., Hausrath, A. C., Juers, D. & Matthews, B. W. (1995). *Protein Sci.* **4**, 1955–1965.
Holland, D. R., Tronrud, D. E., Pley, H. W., Flaherty, K. M., Stark, W., Jansonius, J. N., McKay, D. B. & Matthews, B. W. (1992). *Biochemistry*, **31**, 11310–11316.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (1991). *Crystallographic Computing 5. From Chemistry to Biology*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 50–61. Oxford: IUCr/Oxford University Press.
McKay, D. B., Thayer, M. M., Flaherty, K. M., Pley, H. & Benvegnu, D. (1992). *Matrix Metalloproteinases and Inhibitors*, edited by H. Birkedal-Hansen, Z. Werb, H. Welgus & H. Van Wart, p. 112. New York: Gustav Fischer Verlag.
Matthews, B. W. (1972). *Macromolecules*, **5**, 818–819.
Matthews, B. W. (1988). *Acc. Chem. Res.* **21**, 333–340.

Matthews, B. W., Colman, P. M., Jansonius, J. N., Titani, K., Walsh, K. A. & Neurath, H. (1972). *Nature New Biol.* **238**, 41–43.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Pauptit, R. A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A. S. & Jansonius, J. N. (1988). *J. Mol. Biol.* **199**, 525–537.
Remington, S. J. & Matthews, B. W. (1980). *J. Mol. Biol.* **140**, 77–99.
Ringe, D. & Petsko, G. A. (1985). *Prog. Biophys. Mol. Biol.* **45**, 197–235.
Stark, W., Pauptit, R. A., Wilson, K. & Jansonius, J. N. (1992). *Eur. J. Biochem.* **207**, 781–791.
Thayer, M. M., Flaherty, K. M. & McKay, D. B. (1991). *J. Biol. Chem.* **266**, 2864–2871.
Tronrud, D. E. (1996). *J. Appl. Cryst.* **29**, 100–104.
Tronrud, D. E. (1997). *Methods Enzymol.* **277**, 306–319.
Veltman, O. R., Eijsink, V. G. H., Vriend, G., de Kreij, A., Venema, G. & Van Den Berg, B. (1998). *Biochemistry*, **37**, 5305–5311.
Wagner, G. (1995). *Nature Struct. Biol.* **2**, 255–257.
Zhang, X.-J. & Matthews, B. W. (1995). *J. Appl. Cryst.* **28**, 624–630.
Zhang, X.-J., Wozniak, J. A. & Matthews, B. W. (1995). *J. Mol. Biol.* **250**, 527–552.