Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a new 'aspzincin' metalloendopeptidase from *Grifola frondosa*: implications for the catalytic mechanism and substrate specificity based on several different crystal forms

Tetsuya Hori,[a,b] Takashi Kumasaka,[a]* Masaki Yamamoto,[a] Takashi Nonaka,[c]† Nobuo Tanaka,[b] Yohichi Hashimoto,[d] Tatzuo Ueki[e] and Koji Takio[c]

[a]RIKEN Harima Institute, 1-1-1 Kouto, Mikazuki, Sayo-gun 679-5148, Japan, [b]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8501, Japan, [c]RIKEN, 2-1 Hirosawa, Wako 351-0198, Japan, [d]Department of Biochemistry, Saitama University, Urawa 338-8570, Japan, and [e]Japan Synchrotron Radiation Research Institute (JASRI), 1-1-1 Kouto, Mikazuki, Sayo-gun 679-5198, Japan

† Present address: The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.

Correspondence e-mail: kumasaka@postman.riken.go.jp

Received 23 October 2000
Accepted 5 December 2000

PDB References: $P6_522$ MEP, 1ge6; $P4_3$ MEP, 1ge7; $C2$ MEP, 1g12; $P1$ MEP, 1ge5.

Crystal structures of a peptidyl-Lys metalloendopeptidase (MEP) from the edible mushroom *Grifola frondosa* (*Gf*MEP) were solved in four crystal forms. This represents the first structure of the new family 'aspzincins' with a novel active-site architecture. The active site is composed of two helices and a loop region and includes the **H**E*xx***H** and GT*x***D***xx*YG motifs conserved among aspzincins. His117, His121 and Asp130 coordinate to the catalytic zinc ligands. An electrostatically negative region composed of Asp154 and Glu157 attracts a positively charged Lys side chain of a substrate in a specific manner. A Tyr133 side chain located on the S1′ pocket had different configurations in two crystal forms and was not observed in the other crystal forms. The flexible Tyr133 plays two roles in the enzymatic function of *Gf*MEP. The first is to provide a hydrophobic environment with Phe83 in order to accommodate the alkyl part of the Lys side chain of a substrate and the second is as a 'proton donor' to the oxyanion of the tetrahedral transition state to stabilize the reaction transition state.

1. Introduction

Metalloendopeptidases (MEPs) are widely distributed from bacteria to mammals. Many MEPs have the potential to cause biologically serious events such as tissue destruction (Blundell, 1994) or pathogenic bacterial infections in humans (Hase & Finkelstein, 1993). Recently, many MEP crystal structures have been reported as free enzymes and/or enzyme–inhibitor complexes (Barrett, 1995). Such detailed structural knowledge has provided researchers with an understanding of the catalytic mechanism and primary information needed for structure-based drug design.

Many MEPs have been classified into two major families, 'gluzincins' and 'metzincins', according to the structural topology around the active site (Bode *et al.*, 1993; Hooper, 1994; Rawlings & Barrett, 1995; Fig. 1*a*). Although they commonly possess two His residues as zinc ligands in the conventional **H**E*xx***H** motif on the 'active-site helix', other structural features around the zinc-binding site are quite different in the two families (Fig. 1*a*). In gluzincins, the conserved N**E***xx*SD segment, which is located about 20 amino-acid residues downstream from the **H**E*xx***H** motif, is a part of the helix which is adjacent to the active-site helix and contains the Glu residue as the third zinc ligand. Metzincins have a β-turn with a Met residue (Met-turn) near the zinc ion at the opposite side of the catalytic water molecule, as well as the active-site consensus **H**E*xx***H***xx*G*xx***H** sequence containing another His residue as the third zinc ligand on a loop.

Furthermore, metzincins are divided into four subfamilies based on their overall structures or the source of origin (Bode *et al.*, 1993; Fig. 1*a*).

The peptidyl-Lys metalloendopeptidase from *G. frondosa* (*Gf*MEP) has been purified (Nonaka *et al.*, 1995) and characterized (Nonaka *et al.*, 1997, 1998). It consists of 167 amino-acid residues with one catalytic zinc ion (Fig. 1*b*). The substrate specificity of *Gf*MEP is strict for peptidyl-Lys and longer peptides are favoured as substrates (Nonaka *et al.*, 1998). Because of its unique specificity, the amino-acid sequence of *Gf*MEP has been determined (Yun *et al.*, 1991; Nonaka *et al.*, 1997). *Gf*MEP has a high thermostability, resistance to denaturing agents and binding capability to β-glycans (Nonaka *et al.*, 1995). *Gf*MEP is categorized as a deuterolysin together with several MEPs from fungi based on amino-acid sequence homology (Nonaka *et al.*, 1997; Fig. 1*b*). Based on their unique primary structure, it has been implied that deuterolysins do not share common three-dimensional structural features with either gluzincins or metzincins, with the exception of the common **HE***xx***H** motif (Nonaka *et al.*, 1997). Fushimi *et al.* (1999) proposed a new family name 'aspzincins' based on a site-directed mutagenesis experiment with a deuterolysin, although their assignment of the third zinc ligand of the Zn atom had been an Asp residue corresponding to Asp154 rather than Asp130 of *Gf*MEP. The three-dimensional structure of *Gf*MEP is indispensable for further insight into the catalytic mechanism in MEPs and for the specific recognition of the substrate.

2. Materials and methods

2.1. Preparation of *Gf*MEP and its crystal

*Gf*MEP was isolated from *G. frondosa* as described previously (Nonaka *et al.*, 1995). Four crystal forms, triclinic (space group *P*1), monoclinic (*C*2), tetragonal ($P4_3$) and hexagonal ($P6_522$), were obtained from almost identical crystallization conditions. Briefly, the enzyme was crystallized by the hanging-drop vapour-diffusion method at 293 K by mixing equal volumes of protein solution (about 10 mg ml^{-1} in 10 m*M* acetate buffer pH 5.0) and reservoir solution (0.1–3.0 *M* NaCl in 10 m*M* imidazole buffer pH 6.0–8.0).

2.2. Data collection

X-ray diffraction data from a triclinic crystal were collected at room temperature using Cu $K\alpha$ X-rays (λ = 1.5418 Å) from a Rigaku RU-200 rotating-anode X-ray generator and a Rigaku R-AXIS IIc imaging-plate system. Images were processed and merged using *PROCESS* (Rigaku) (Table 1). The data sets of the other crystals were collected at 100 K using cryocrystallographic techniques after soaking them in reservoir solutions with stepwise increments in glycerol concentration up to 30%. The data were collected at the BL45XU beamline (Yamamoto *et al.*, 1998) at SPring-8 using a Rigaku R-AXIS IV detector. A MAD experiment was performed on the hexagonal crystal using three wavelengths, λ_1 (1.0400 Å, remote), λ_2 (1.2829 Å, edge) and λ_3 (1.2823 Å,

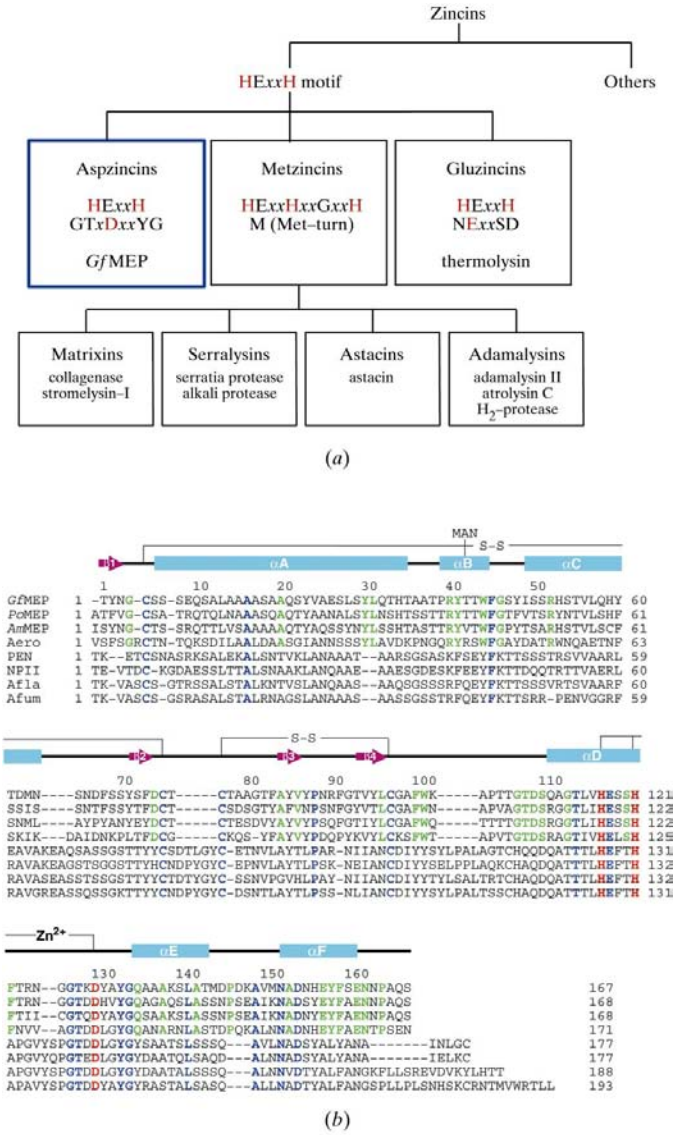

Figure 1
Introduction of the aspzincin family. (*a*) Classification of metalloendopeptidase. The conserved amino-acid sequence at the active site is shown under the family name, with the zinc-ligand residues coloured in red. Some typical MEPs whose structures have been solved are indicated: thermolysin (Holden *et al.*, 1987), collagenase (neutrophil collagenase; Bode *et al.*, 1994; fibroblast collagenase; Lovejoy *et al.*, 1994), stromelysin-1 (Gooley *et al.*, 1994), serratia protease (Baumann, 1994), alkali protease (Baumann *et al.*, 1993), astacin (Bode *et al.*, 1992), adamalysin II (Gomis-Rüth *et al.*, 1994), atrolysin C (Zhang *et al.*, 1994) and H_2-protease (Kumasaka *et al.*, 1996). (*b*) Sequence alignment of MEPs of the aspzincin family. The MEPs from *G. frondosa* (*Gf*MEP; Nonaka *et al.*, 1997), *Pleurotus ostreatus* (*Po*MEP; Nonaka *et al.*, 1997), *Armillaria mellea* (*Am*MEP; Healy *et al.*, 1999 and GenPept database accession CAB42792), *Aeromonas hydrophila* (Aero; Chang *et al.*, 1997), *Penicillium citrinum* (PEN, penicillolysin; Matsumoto *et al.*, 1994), *Aspergillus oryzae* (NPII, neutral protease II; Tatsumi *et al.*, 1991), *Aspergillus flavs* (Afla; Ramesh *et al.*, 1995) and *Aspergillus fumigata* (Afum; Ramesh *et al.*, 1995). Zinc-ligand residues are coloured in red, other entirely conserved residues in blue and those conserved among Lys-specific MEPs in green (Aero has not been proven experimentally). The locations of the α-helices and β-strands are indicated above the alignment. The zinc-binding site, the disulfide bridge and the O-linked mannose site in *Gf*MEP are also shown. The sequences were aligned using *CLUSTAL* (Higgins & Sharp, 1988).

peak). The Bijvoet pairs were collected using inverse-beam geometry. The data sets of the monoclinic and tetragonal crystals were collected with only one wavelength (Table 1). These data sets were then processed and merged using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) (Table 1).

2.3. Structure determination

The hexagonal crystal structure was determined by MAD phasing using the catalytic zinc ion. After scaling the three data sets collected at the three wavelengths with the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), Bijvoet and dispersive difference Patterson maps were calculated. One clear peak was found in an asymmetric unit and indicated that there was one Zn atom and one monomer molecule in an asymmetric unit. The zinc position was determined with the *CCP*4 suite and was refined using *SHARP* (de La Fortelle *et al.,* 1997). An initial MAD phase was calculated for reflections in the resolution range 2.2–25.0 Å (Table 2). Solvent flattening and histogram matching were performed with *DM* (Collaborative Computational Project, Number 4, 1994). The enantiomorph was selected to be $P6_522$ because the electron-density map in $P6_522$ was clearer than that in $P6_122$ and could be used for tracing the polypeptide chain. Finally, one Zn atom and all residues except the four N-terminal residues were traced using *O* (Jones *et al.*, 1991) and then refined with *CNS* (Brunger *et al.*, 1998) (Table 3).

The structures of the other crystal systems were solved by the molecular-replacement method with *AMoRe* (Collaborative Computational Project, Number 4, 1994) using the hexagonal structure as a search model and were refined with *CNS* (Brunger *et al.*, 1998) (Table 3). In the tetragonal crystal there are two molecules in an asymmetric unit. Finally, the water molecules and an *O*-linked mannose were included in the models. In the $|F_o| - |F_c|$ Fourier map, no significant residual electron density was observed. Stereochemistry was verified using *PROCHECK* (Collaborative Computational Project, Number 4, 1994). All residues lie within the allowed regions of the Ramachandran plot, with over 90% of the residues in the most favourable regions in all four crystals (Table 3). Pro164 was in the *cis* peptide-bond conformation.

2.4. The model building of the proposed transition state

A transition state of the peptide substrates was modelled based on the monoclinic and hexagonal crystal structures and a transition-state analogue inhibitor of thermolysin (Holden *et al.*, 1987; PDB code 4tmn). The P atom of the phosphonamidate group was replaced by a C atom and its sequence of carbobenzoxy(CB)(P2 site)-Phe(P1)-Leu(P1′)-Ala(P2′) was also replaced by CB(P2)-Ala(P1)-Lys(P1′)-Ala(P2′) using *O*. Firstly, the substrate-complex model with the monoclinic crystal structure was constructed using *O* under the following conditions. (i) The main chain of the substrate was aligned with the strand $\beta3$ in an antiparallel manner. Two hydrogen bonds can be formed between the backbone N atom of Val86 and the carbonyl O atom of the P2 CB group (2.9 Å) and between the carbonyl O atom of Ala84 and the main-chain N atom of P1′ Lys (2.8 Å). (ii) The amino group of the replaced P1′ Lys residue was placed so as to make a salt bridge with Glu157 located at the bottom of the S1′ pocket at a distance of 2.8 Å. (iii) The tetrahedral 'water-attacked carbonyl groups' as a transition intermediate were coordinated to the zinc ion (2.5 Å in both interactions). One water-attacked carbonyl O atom interacts with the carboxyl group of catalytic residue Glu118 (2.7 and 3.0 Å). Another water-attacked carbonyl O atom makes a hydrogen bond with the OH group of the putative proton donor Tyr133 (3.3 Å). (iv) The torsion angles of the P2 CB group and P2′ Ala were rotated to fit in *Gf*MEP

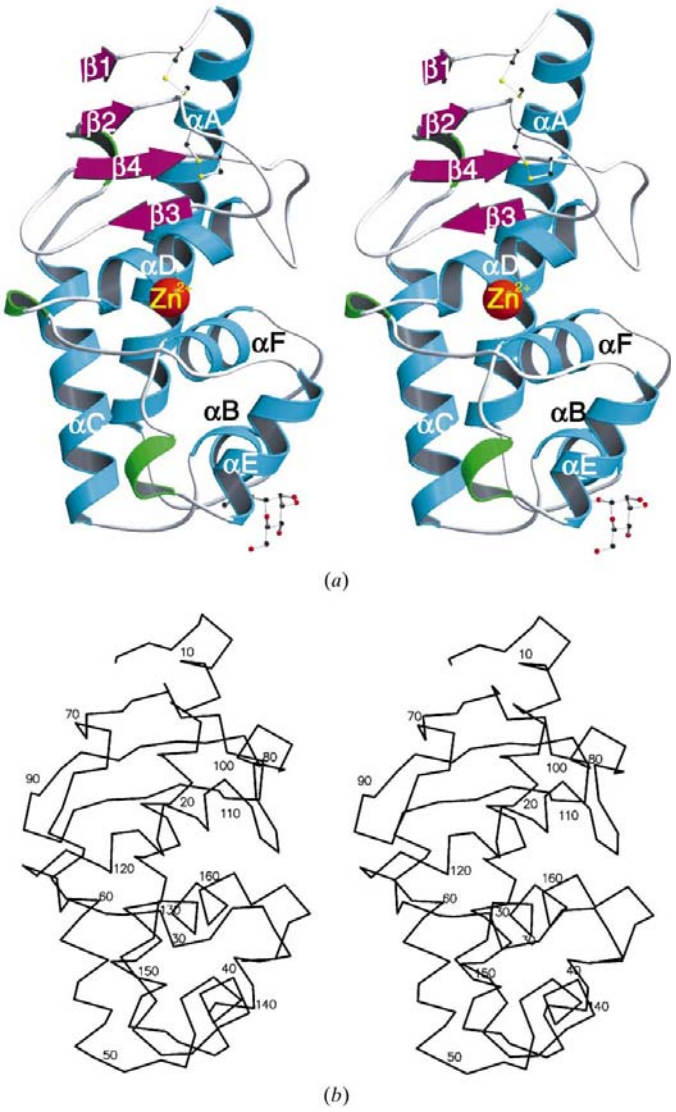

Figure 2
The structure of *Gf*MEP. (*a*) Ribbon representation of the overall structure of *Gf*MEP. The α-helices are shown in cyan, the β-strands in magenta, the β-turn in green and the loops in grey. The catalytic zinc ion, two disulfide bridges and Thr42 with O-linked mannose are also shown as ball-and-stick models. Atoms are coloured with carbon in black, oxygen in red, sulfur in yellow and zinc in orange. (*b*) Stereoview of the C^{α} backbone representation of *Gf*MEP. Every tenth residue is labelled. These figures were prepared with *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Bacon, 1997).

Table 1
Data-collection statistics.

Crystal	Hexagonal			Tetragonal	Monoclinic	Triclinic
X-ray source	SPring-8	SPring-8	SPring-8	SPring-8	SPring-8	RU-200
Wavelength (Å)	1.0400 (remote)	1.2829 (edge)	1.2823 (peak)	1.0100	1.0400	1.5418
Temperature (K)	100	100	100	100	100	293
Resolution (Å)	2.2	2.2	2.8	2.0	1.6	2.0
Space group	$P6_522$			$P4_3$	$C2$	$P1$
Unit-cell parameters†						
a (Å)	58.7 (0.08)			30.2 (0.02)	43.6 (0.09)	30.8 (0.05)
b (Å)	58.7 (0.08)			30.2 (0.02)	41.9 (0.08)	40.5 (0.09)
c (Å)	145.2 (0.13)			308.0 (0.17)	76.9 (0.17)	30.5 (0.06)
α (°)	90			90	90	105.9 (0.08)
β (°)	90			90	95.5 (0.08)	91.7 (0.15)
γ (°)	120			90	90	103.5 (0.20)
No. of molecules‡	1			2	1	1
No. of reflections						
Observed	50233	47504	23765	137781	96644	9943§
Unique	8115	8056	4047	18300	15467	6440§
Completeness¶ (%)	99.5 (99.7)	99.8 (95.8)	99.8 (97.6)	98.5 (98.1)	84.3 (40.9)	64.8 (43.4)§
I/σ¶	13.2 (11.9)	14.9 (9.5)	24.8 (24.1)	9.6 (6.1)	21.4 (10.8)	5.4 (2.4)§††
R_{merge}¶‡‡ (%)	7.5 (18.4)	7.0 (19.1)	4.6 (7.2)	9.7 (34.1)	6.9 (16.5)	13.5 (19.9)§

† Numbers in parentheses represent the standard deviations of the unit-cell parameters. ‡ The number of molecules in the asymmetric unit. § $F/\sigma > 1$. ¶ Numbers in parentheses represent values in the highest resolution shell. †† F/σ. ‡‡ $R_{merge} = \sum|I_i - \langle I_i\rangle|/\sum I_i$ in $P6_522$, $P4_3$ and $C2$; $R_{merge} = \sum|F_i - \langle F_i\rangle|/F_i$ in $P1$.

Table 2
MAD phasing ($P6_522$ crystal).

	Remote		Edge		Peak	
Data set	Acentric Iso/ano	Centric	Acentric Iso/ano	Centric	Acentric Iso/ano	Centric
Phasing power†	—/1.7		4.4/1.7	2.8	3.7/1.7	2.8
R_{cullis}‡	—/0.8		0.4/0.8	0.4	0.5/0.8	0.5
FOM (MAD)§	0.56	0.52				
FOM (DM)§	0.78					

† Phasing power = $\langle|F_{hc}|/E\rangle$. ‡ $R_{cullis} = \langle E\rangle/\langle|F_{PH}| - |F_P|\rangle$. § Mean overall figure of merit.

without colliding. Moreover, the location of the Tyr133 side chain was adjusted with χ_1 and χ_2 angles to keep P2′ Ala and Tyr133 out of close contact. The substrate molecule was also superimposed onto the hexagonal crystal structure in the same way. The interactions between the substrate molecule and the hexagonal structure, including the hydrogen-bond distances, were almost the same except one for Tyr133 which moved to the cleft region. As a result, the P1′ Lys side chain was surrounded by the aromatic rings of Tyr133 and Phe83 of the hexagonal structure.

3. Results and discussion

3.1. Overall structure

*Gf*MEP has a monomeric structure that is made up of upper and lower domains (Figs. 2*a* and 2*b*) and the catalytic zinc ion is located in the active-site cleft between the two domains. The upper domain displays an α/β fold which consists of a β-sheet and two helices, αD and the N-terminal half of αA. The short β-sheet contains four β-strands with two disulfide bridges, Cys5–Cys75 and Cys77–Cys97, in its loop regions. The lower domain is composed of five helices, αB, αC, αE, αF and the C-terminal half of the long αA. An O-linked mannose is located at Thr42 on helix αB and is exposed to the solvent.

The upper domains of most MEP structures have a similar architecture, an α/β fold, whereas the number of its components differ among the three MEP families. *Gf*MEP has two helices and a four-stranded β-sheet as mentioned above (Fig. 3*a*). In gluzincins, two α-helices and a five-stranded β-sheet follow an additional β-sheet (Fig. 3*c*). In metzincins, the β-sheet is composed of five β-strands and there are two or four helices (Fig. 3*b*).

In contrast, the lower domains are quite unique. The lower domain of *Gf*MEP is made up of several α-helices (Fig. 3*a*). The lower domain of the metzincins are fundamentally composed of only one α-helix and an irregularly folded long loop region (Fig. 3*b*), although the structural folds of the lower domain are not well conserved. The lower domain of the gluzincins has an α/β structure that is made up of a two-stranded β-sheet and five α-helices (Fig. 3*c*).

The solved structures in the four crystal forms were coincident except for the side-chain conformation of Tyr133 and four N-terminal residues. The overall r.m.s. deviations in the main chains among any pair of structures was less than 0.78 Å. The side-chain conformations of Tyr133 were different in the hexagonal and monoclinic crystals and the electron density corresponding to the side chain could not be observed in the tetragonal and triclinic crystals. In the hexagonal crystal form, the four N-terminal residues were not constructed because of their poor electron density.

3.2. The active cleft

The structure of the catalytic cleft is unique in *Gf*MEP. It is composed of helix αD, half of the domain-connecting loop αD–αE and helix αF (Fig. 3*a*). Because it contains the common **H**E*xx***H** motif at the C-terminal end, the helix αD is called the active-site helix (Bode *et al.*, 1993). It surrounds the catalytic zinc ion together with loop αD–αE.

The active-site structure around the zinc-binding site in *Gf*MEP resembles those of metzincins, which have a helix–loop structure (Figs. 3*a* and 3*b*), although the amino-acid sequences are distinct (Nonaka *et al.*, 1997). Moreover, the helix αF forms a part of the putative S1′ pocket which determines the substrate specificity of *Gf*MEP (Fig. 3*a*). The arrangement of the two helices αD and αF somewhat resembles that in gluzincins (Figs. 3*a* and 3*c*).

Table 3
Structural refinement.

Space group	$P6_522$	$P4_3$	$C2$	$P1$
Resolution (Å)	27.2–2.2	10.0–2.0	38.3–1.6	29.8–2.0
No. of heavy atoms	1	2	1	1
No. of waters	145	297	201	131
R (%)†	19.8	17.9	21.2	18.4
R_{free} (%)†	23.3	21.8	22.9	22.2
R.m.s. deviation				
Bonds (Å)	0.006	0.005	0.005	0.006
Angles (°)	1.12	1.08	1.06	1.16
Average B factors (Å)				
Protein	12.7	18.8	14.5	17.6
Waters	31.3	34.9	34.0	28.3
Ramachandran plot				
Most favoured (%)	91.7	92.9	93.9	93.2
Allowed (%)	8.3	7.1	6.1	6.8

† $R = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$. 95 and 5% of reflections were used for the working and test sets, respectively.

The edges of the cleft are composed of strand β3 and the C-terminal half of loop αD–αE in *Gf*MEP (Fig. 3*a*). A corresponding β-strand to β3 also exists in other MEPs. This arrangement is similar to catalytic zinc and is conventionally called the 'edge strands' (Holden *et al.*, 1987; Grams *et al.*, 1996). It is believed to interact with the peptide backbone of the substrate in the mode of the antiparallel β-sheet.

3.3. Catalytic zinc environment

The active-site helix, helix αD, contains two zinc ligands, His117 and His121, and the catalytic carboxylic residue Glu118 as components of the **H**E*xx***H** motif (Fig. 4*a*). The loop αD–αE includes the GT*x***D***xx*YG segment, which is strictly conserved among aspzincins (Fig. 1*b*) and contains Asp130 as the third zinc ligand (Fig. 4*a*). Asp130 makes a unique bifurcated interaction with the zinc ion (Figs. 4*a* and 4*b*). Such a bifurcated interaction has not been observed in other MEPs whose third zinc ligand is Asp or Glu (Banbula *et al.*, 1998; Holden *et al.*, 1987; Kurisu *et al.*, 1997).

The fourth ligand is the catalytic water molecule, which also makes hydrogen bonds with both Glu118 carboxyl O atoms (about 3.0 Å distance in the four crystal forms; Figs. 4*a* and 4*b*). The water molecule exists in all MEPs and would act as a catalytic base in the enzymatic action. Unlike most MEPs, the coordination of the four ligands in *Gf*MEP is not tetrahedral but pyramidal. The pyramidal base plane is formed by His117 N^{ε}, both carboxyl O atoms of Asp130 and the catalytic water molecule. His121 N^{ε} is located at the vertex.

There is another water molecule which acts as the fifth zinc ligand at the opposite side of the bipyramidal vertex to His121 and forms a hydrogen bond with the phenolic hydroxyl group of Tyr133 only in the monoclinic crystal (Figs. 4*a* and 4*b*). The electron density of the fifth water molecule was weaker than in the fourth ligand water and the B factor of the fifth water is higher than the fourth one by about 10 $Å^2$ in each structure. Since a zinc ion could form complexes with coordination numbers 4, 5 and 6 and the coordination geometry of the zinc ion is very flexible with low energy differences, a variety of zinc coordinations are also reported in MEPs: tetrahedral, pyramidal and bipyramidal (Lipscomb & Strater, 1996). Bipyramidal coordination has also been observed in two other MEPs (Banbula *et al.*, 1998; Bode *et al.*, 1992). One is a gluzincin from *Staphylococcus aureus*, although the fifth water molecule was not well defined and had a high B-factor value (Banbula *et al.*, 1998). The other is astacin, a subfamily of metzincins, but its fifth ligand was the phenolic hydroxyl group of Tyr (Bode *et al.*, 1992), not a water molecule. In the transition state, a scissile peptide bond will form a tetrahedral water-attacked carbonyl group (Holden *et al.*, 1987; Grams *et al.*, 1996); the locations of the two ligand water molecules in *Gf*MEP are almost coincident with those of the two O atoms of the water-attacked carbonyl group.

Another important region near the zinc ion is the opposite side of the catalytic water molecule. In *Gf*MEP, the conserved Ala153 occupies this region. On the other hand, in metzincins there is a 'Met-turn' in the region, which is a sharp β-turn including an invariant Met residue. The Met residue is thought to contribute to the stabilization of the hydrophobic environment around the zinc ion. It remains to be clarified whether Ala153 plays a similar role in *Gf*MEP to the Met in metzincins.

The interactions between the zinc ligands and the surrounding residues are critical in the construction of the chelating architecture. The N^{δ} atoms of His117 and His121 interact with the carboxyl group of Asp154 and the hydroxyl group of Thr128, respectively (Figs. 4*a* and 4*b*). These residues are conserved in the aspzincins (Fig. 1*b*). Moreover, the carbonyl group of Asp154 interacts with the main-chain N atom of Gly134 (Fig. 4*b*). In Asp130, the carboxyl group interacts with the water molecule which is observed in all four crystal forms of *Gf*MEP and also interacts with the N^{δ} atom of Asn152, the main-chain O atom of Ala132 and the main-chain N atom of Asp154 (Fig. 4*b*). Gln152 is also conserved in aspzincins (Fig. 1*b*). In a site-directed mutational study of an aspzincin from *Aspergillus oryzae* (Fushimi *et al.*, 1999), the mutation of the corresponding Asp residue from Asp154 to Asn drastically reduced its zinc-binding ability; it was concluded that the Asp was the third zinc ligand. The crystal structure of *Gf*MEP clearly indicates that the mutation indirectly caused the loss of zinc-binding ability. These observations indicate that the non-bonding interactions between Asp154 and His117 and between Asp154 and Gly134 are both important for the formation of the zinc coordination. The interaction between His121 and Thr128, and between Asp130 and an invariant water molecule also play an important role in sustaining the coordination. A similar network has been found in many other MEPs. In the case of gluzincins, Asp and Asn residues act as hydrogen acceptors for hydrogen bonds to the His residues of zinc ligands; in metzincins, two main-chain carbonyl O atoms are also involved as hydrogen acceptors. The necessity of the interaction network between zinc ligands and other residues has also been discussed in the mutational studies and the crystal structures of carbonic anhydrase II (Lesburg & Christianson, 1995).

3.4. S1′ pocket and Lys specificity

In the electrostatic potential surface of *Gf*MEP, a highly negative region is found at the bottom of the cleft which is surrounded by a hydrophobic bank (Fig. 5). The region, the putative S1′ pocket, specifically accommodates a positively charged amine of the Lys side chain. Among Lys-specific aspzincins, several acidic or polar residues are conserved on the helices αD, αF and the loop β4–αD (Fig. 1*b*) and two negatively charged residues, Asp154 and Glu157, are located at the bottom of the S1′ pocket (Fig. 4*a*). Glu157 is invariable only among Lys-specific aspzincins, whereas Asp154 is conserved among all aspzincins (Fig. 1*b*). In this respect, Glu157 is the prime residue in Lys-specificity determinants. The positively charged ε-amino group of a Lys residue in a substrate could directly interact with the Glu157 side chain as proposed in the substrate-binding model (Fig. 6). On the other hand, the binding of a hydrophobic P1′ residue would need to pay a higher dehydration cost when removing the several water molecules around the S1′ pocket; this would be unfavourable.

The depth of the S1′ pocket of *Gf*MEP may contribute to the strict Lys specificity. Although *S*-(2-aminoethyl)cysteine-containing peptides are hydrolyzed

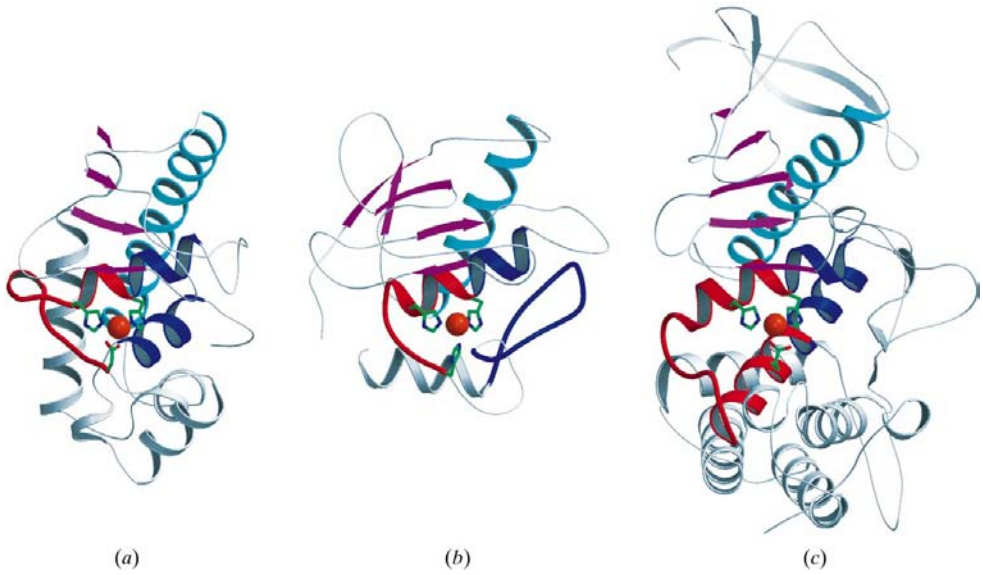

Figure 3
Representative structures of the respective MEP families. (*a*) *Gf*MEP, (*b*) neutrophil collagenase (Bode *et al.*, 1994; metzincin; PDB code 1jap) and (*c*) thermolysin (Holden *et al.*, 1987; gluzincin; PDB code 4tmn). The S1′ pocket region is coloured in blue and the zinc-binding site in red. The α-helices and β-strands which compose the common folding topology among the three families are coloured cyan and magenta, respectively. The catalytic zinc ion and its ligand residues are also presented. Atoms are coloured with carbon in green, oxygen in red, nitrogen in blue and zinc in orange. All molecules are aligned in the same orientation with regard to the two ligand His residues and the zinc ion.

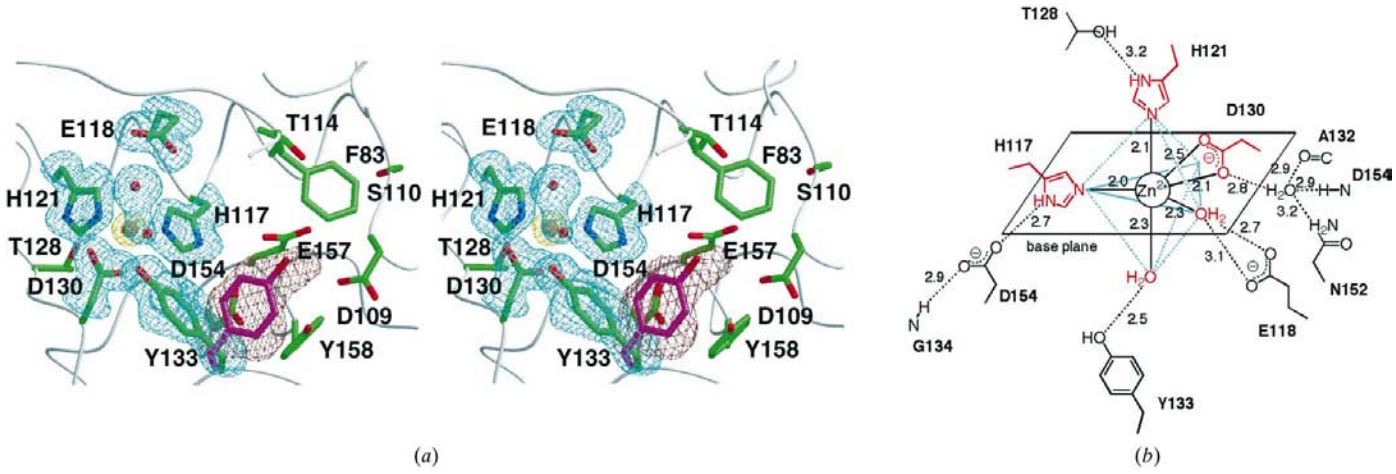

Figure 4
The zinc-binding site and the cleft region. (*a*) A stereoview with omitted electron density around the active site. Two structures (the monoclinic and hexagonal crystal) are superimposed based on the zinc-ligand residues. In monoclinic crystal, all residues related to the catalytic reaction are presented with C atoms coloured green, O atoms red and N atoms blue. Tyr133 of the hexagonal crystal is also drawn with C atoms in magenta and O atoms in red. The cyan-coloured map was calculated after omitting the side chains of His117, Glu118, His121, Asp130, Tyr133 and two bound waters in the monoclinic crystal form. The yellow map is the omitted map of a zinc ion in a monoclinic crystal. The red map is the omitted map of the side chain of Tyr133 in a hexagonal crystal. The cyan and red maps are contoured at 3.0σ and the yellow map at 16.0σ. These figures were prepared with *Bobscript* (Esnouf, 1997). (*b*) Schematic drawing of the zinc environment. The zinc ligands are coloured red. The interaction between the zinc ion and its ligands is represented by black lines and other interactions by broken black lines. Cyan lines indicate the pyramidal formation of the zinc coordination. His117, Asp130 and a water molecule together with the zinc ion form the pyramidal base plane. His121 and another water molecule are located at its vertices. Numbers by the black lines indicate the distances (Å) in the monoclinic structure.

slowly (Mochizuki & Hashimoto, unpublished results), those containing Arg or ornithine (Orn) residues are not substrates; rather, they act as competitive inhibitors (Nonaka *et al.*, 1998). Thus, either an Arg or an Orn residue in a peptide also has the ability to bind to the S1′ pocket. The complex model of peptidyl-Arg with *Gf*MEP suggests that the carbonyl C atom of the scissile peptide bond is distant from the catalytic water by 3.0 Å when the guanidino group of the Arg residue makes a salt bridge with the Glu157 carboxyl group in the same manner as that of the peptidyl-Lys. This would not allow the nucleophilic attack by the water molecule and hydrolytic action would not take place at all. In other words, the S1′ pocket of *Gf*MEP is not deep enough to accommodate a peptidyl-Arg side chain for hydrolysis owing to it having a longer side chain than that of Lys. On the contrary, trypsin, which is known to be a basic residue-specific serine protease, hydrolyzes substrates with either Lys or Arg as the P1 residues (Bode *et al.*, 1984; Marquart *et al.*, 1983). Arg as a P1 residue reaches the Asp residue of trypsin at the bottom of the S1 pocket directly and the side-chain amino group of P1 Lys forms a hydrogen bond with the carboxyl group of the Asp at the S1 pocket by way of a water molecule (Bode *et al.*, 1984; Marquart *et al.*, 1983). It would be interesting to determine whether the E157D mutant of *Gf*MEP mimics trypsin substrate specificity.

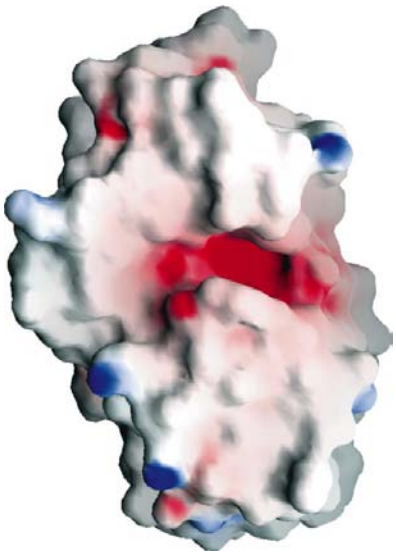

Figure 5
The molecular-surface representation of the cleft region in the hexagonal crystals coloured according to their electrostatic potential: red for the electrostatically negative region and blue for the positive region. The molecular orientation is the same as in Fig. 2. The molecular surface and electrostatic potential were calculated with *GRASP* (Nicholls *et al.*, 1991).

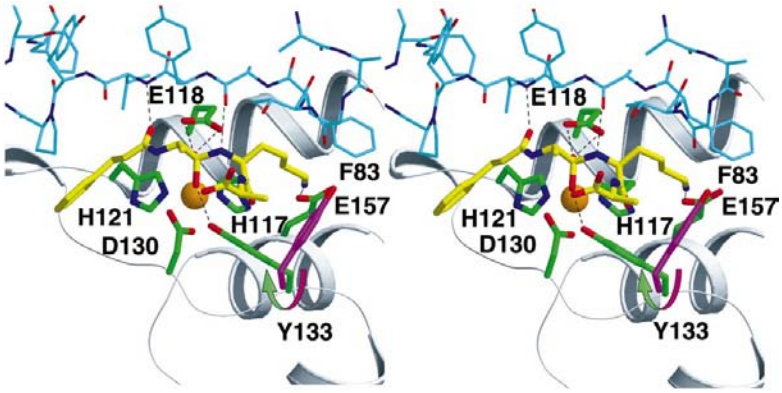

Figure 6
A stereoview of the *Gf*MEP tetrahedral transition-state model. The two structures (monoclinic and hexagonal crystal forms) are superimposed based on the zinc-ligand residues. Residues related to catalytic reaction (His117, Glu118, His121, Asp130 and Tyr133) in the monoclinic crystal are presented with C atoms coloured green, O atoms red and N atoms blue. Strand β3 is presented as a thin stick model with C atoms coloured cyan. In the hexagonal crystal, only Tyr133 is presented, with C atoms in magenta. The modelled substrate whose sequence is a carbobenzoxy (CB)(P2 site)-Ala(P1)-Lys(P1′)-Ala(P2′) is presented with C atoms in yellow. Putative hydrogen bonds and the salt bridge between the substrate and *Gf*MEP are drawn in broken lines. The expected rotation of the side chain of Tyr133 is shown by an arrow.

3.5. The role of Tyr133

Tyr133 is invariant among aspzincins (Fig. 1*b*) and is located near the catalytic zinc ion (Fig. 4*a*). Tyr133 may play several important roles in the catalytic action. The Tyr133 residues in the hexagonal and monoclinic crystals show different conformations and the electron density corresponding to the side chain is disordered in the tetragonal and triclinic crystals. The differences in the conformations of Tyr133 are mainly a consequence of a change in the torsion angle χ_1 (Fig. 4*a*).

Tyr133 is located over the cleft region away from the zinc ion in the hexagonal crystal structure (Fig. 4*a*). In this conformation, Tyr133 obstructs the active-site cleft and appears to prevent the S1′ pocket from the Lys side chain of a substrate from interacting with Glu157. One possible role of Tyr133 is that the Tyr133 aromatic ring together with the Phe83 aromatic ring stabilize the hydrophobic alkyl chain of Lys, contributing to the Lys specificity (Fig. 6). The *Gf*MEP S1′ pocket is shaped like a long trench rather than a hole (Fig. 5), so Tyr133 covers the S1′ pocket and forms a hydrophobic bank as seen in the hexagonal crystal (Fig. 4*a*). In the case of trypsin, the S1 pocket is composed of several polar and hydrophobic residues and an Asp residue is the only source of the counter charge at the bottom of the hole (Bode *et al.*, 1984). This suggests that the lined hydrophobic residues at the surface of the specificity pocket of trypsin also contribute to the strict recognition of the Lys or Arg side chain and Tyr133 of *Gf*MEP would play the same role as such hydrophobic residues.

As described above, Tyr133 interacts with the zinc ion *via* the water molecule in the monoclinic crystal structure (Figs. 4*a* and 4*b*). The Tyr133 side chain is able to rotate by about 300° in torsion angle χ_1 from the configuration which covers the cleft region (Fig. 6). The phenolic hydroxyl group of Tyr133 occupies the proper position to form a hydrogen bond with the water-attacked carbonyl group in the proposed transition-state model (Fig. 6). Thus, Tyr133 acts

as a 'proton donor' to stabilize the tetrahedral transition state *via* the hydrogen bond to the oxyanion, resulting in the reduction of the transition-state energy. The proton donor residues are also found in gluzincins (Holden *et al.*, 1987), astacin (Grams *et al.*, 1996), serralysin (Baumann *et al.*, 1993) and carboxypeptidase A (Kim & Lipscomb, 1990) and correspond to the oxyanion hole in serine proteases (Bode *et al.*, 1984; Marquart *et al.*, 1983).

We are grateful to M. Miyano for stimulating and fruitful discussions and to T. Ishikawa, Y. Inoue and the RIKEN SR Structural Biology Research Group for their many helpful suggestions and support. This work was also supported by a Grant-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan.

References

Banbula, A., Potempa, J., Travis, J., Fernandez-Catalan, C., Mann, K., Huber, R., Bode, W. & Medrano, F. J. (1998). *Structure*, **6**, 1185–1193.
Barrett, A. J. (1995). *Methods Enzymol.* **248**, 183–187.
Baumann, U. (1994). *J. Mol. Biol.* **242**, 244–251.
Baumann, U., Wu, S., Flaherty, K. M. & McKay, D. B. (1993). *EMBO J.* **12**, 3357–3364.
Blundell, T. L. (1994). *Nature Struct. Biol.* **1**, 73–75.
Bode, W., Gomis-Rüth, F. X., Huber, R., Zwilling, R. & Stöcker, W. (1992). *Nature (London)*, **358**, 164–167.
Bode, W., Gomis-Rüth, F. X. & Stöcker, W. (1993). *FEBS Lett.* **331**, 134–140.
Bode, W., Reinemer, P., Huber, R., Kleine, T., Schnierer, S. & Tschesche, H. (1994). *EMBO J.* **13**, 1263–1269.
Bode, W., Walter, J., Huber, R., Wenzel, H. R. & Tschesche, H. (1984). *Eur. J. Biochem.* **144**, 185–190.
Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chang, T. M., Liu, C. C. & Chang, M. C. (1997). *Gene*, **199**, 225–229.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Fushimi, N., Ee, C. E., Nakajima, T. & Ichishima, E. (1999). *J. Biol. Chem.* **274**, 24195–24201.
Gomis-Rüth, F. X., Kress, L. F., Kellermann, J., Mayr, I., Lee, X., Huber, R. & Bode, W. (1994). *J. Mol. Biol.* **239**, 513–544.
Gooley, P. R., O'Connell, J. F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P. & Johnson, B. A. (1994). *Nature Struct. Biol.* **1**, 111–118.
Grams, F., Dive, V., Yiotakis, A., Yiallouros, I., Vassiliou, S., Zwilling, R., Bode, W. & Stöcker, W. (1996). *Nature Struct. Biol.* **3**, 671–675.
Hase, C. C. & Finkelstein, R. A. (1993). *Microbiol. Rev.* **57**, 823–837.
Healy, V. O., O'Connell, J., McCarthy, T. V. & Doonan, S. (1999). *Biochem. Biophys. Res. Commun.* **262**, 60–63.
Higgins, D. G. & Sharp, P. M. (1988). *Gene*, **73**, 237–244.
Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H. & Matthews, B. W. (1987). *Biochemistry*, **26**, 8542–8553.
Hooper, N. H. (1994). *FEBS Lett.* **354**, 1–6.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kim, H. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 5546–5555.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kumasaka, T., Yamamoto, M., Moriyama, H., Tanaka, N., Sato, M., Katsube, Y., Yamakawa, Y., Omori-Satoh, T., Iwanaga, S. & Ueki, T. (1996). *J. Biochem. (Tokyo)*, **119**, 49–57.
Kurisu, G., Kinoshita, T., Sugimoto, A., Nagara, A., Kai, Y., Kasai, N. & Harada, S. (1997). *J. Biochem. (Tokyo)*, **121**, 304–308.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lesburg, C. A. & Christianson, D. W. (1995). *J. Am. Chem. Soc.* **117**, 6838–6844.
Lipscomb, W. N. & Strater, N. (1996). *Chem. Rev.* **96,** 2375–2433.
Lovejoy, B., Cleasby, A., Hassell, A. M., Longley, K., Luther, M. A., Weigl, D., McGeehan, G., McElroy, A. B., Drewry, D., Lambert, M. H. & Jordan, S. R. (1994). *Science*, **263**, 375–377.
Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. (1983). *Acta Cryst.* B**39**, 480–489.
Matsumoto, K., Yamaguchi, M. & Ichishima, E. (1994). *Biochim. Biophys. Acta*, **1218**, 469–472.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Nonaka, T., Dohmae, N., Hashimoto, Y. & Takio, K. (1997). *J. Biol. Chem.* **272**, 30032–30039.
Nonaka, T., Hashimoto, Y. & Takio, K. (1998). *J. Biochem. (Tokyo)*, **124**, 157–162.
Nonaka, T., Ishikawa, H., Tsumuraya, Y., Hashimoto, Y. & Takio, K. (1995). *J. Biochem. (Tokyo)*, **118**, 1014–1020.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ramesh, M. V., Sirakova, T. D. & Kolattukudy, P. E. (1995). *Gene*, **165**, 121–125.
Rawlings, N. D. & Barrett, A. J. (1995). *Methods Enzymol.* **248**, 183–228.
Tatsumi, H., Ikegaya, K., Murakami, S., Kawabe, H., Nakano, E. & Motai, H. (1991). *Mol. Gen. Genet.* **228**, 97–103.
Yamamoto, M., Kumasaka, T., Fujisawa, T. & Ueki, T. (1998). *J. Synchrotron Rad.* **5**, 222–225.
Yun, S.-J., Hiraoka, Y., Nishizawa, M., Takio, K., Titani, K., Nogi, Y. & Fukasawa, T. (1991). *J. Biol. Chem.* **266,** 693–697.
Zhang, D., Botos, I., Gomis-Rüth, F. X., Doll, R., Blood, C., Njoroge, F. G., Fox, J. W., Bode, W. & Meyer, E. F. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 8447–8451.