A new variant of the Ntn hydrolase fold revealed by the crystal structure of L-aminopeptidase D-Ala-esterase/amidase from *Ochrobactrum anthropi*

Coralie Bompard-Gilles[1], Vincent Villeret[1], Gideon J Davies[2], Laurence Fanuel[3], Bernard Joris[3], Jean-Marie Frère[3] and Jozef Van Beeumen[1]*

Background: The L-aminopeptidase D-Ala-esterase/amidase from *Ochrobactrum anthropi* (DmpA) releases the N-terminal L and/or D-Ala residues from peptide substrates. This is the only known enzyme to liberate N-terminal amino acids with both D and L stereospecificity. The DmpA active form is an αβ heterodimer, which results from a putative autocatalytic cleavage of an inactive precursor polypeptide.

Results: The crystal structure of the enzyme has been determined to 1.82 Å resolution using the multiple isomorphous replacement method. The heterodimer folds into a single domain organised as an αββα sandwich in which two mixed β sheets are flanked on both sides by two α helices.

Conclusions: DmpA shows no similarity to other known aminopeptidases in either fold or catalytic mechanism, and thus represents the first example of a novel family of aminopeptidases. The protein fold of DmpA does, however, show structural homology to members of the N-terminal nucleophile (Ntn) hydrolase superfamily. DmpA presents functionally equivalent residues in the catalytic centre when compared with other Ntn hydrolases, and is therefore likely to use the same catalytic mechanism. In spite of this homology, the direction and connectivity of the secondary structure elements differ significantly from the consensus Ntn hydrolase topology. The DmpA structure thus characterises a new subfamily, but supports the common catalytic mechanism for these enzymes suggesting an evolutionary relationship.

Addresses: [1]Laboratorium voor Eiwitbiochemie en Eiwitengineering, Rijksuniversiteit-Gent, K L Ledeganckstraat, 35, B-9000 Gent, Belgium, [2]Department of Chemistry, University of York, Heslington, York Y01 5DD, UK and [3]Laboratoire d'Enzymologie et Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, B-4000 Sart-Tilman, Belgium.

*Corresponding author.
E-mail: jozef.vanbeeumen@rug.ac.be

Key words: autocatalysis, crystal structure, Ntn amidohydrolase, serine aminopeptidase, stereospecificity

Received: **29 September 1999**
Revisions requested: **2 November 1999**
Revisions received: **3 December 1999**
Accepted: **10 December 1999**

Published: **28 January 2000**

Structure 2000, **8**:153–162

0969-2126/00/$ – see front matter

Introduction

Aminopeptidases (APs) are exopeptidases that selectively release N-terminal amino acid residues from polypeptides and proteins. These enzymes are found to be widely distributed amongst both prokaryotic and eukaryotic organisms. The many proposed functions of APs include protein maturation and N-terminal degradation, hormone level regulation and cell-cycle control [1,2]. Some bacterial peptidase systems are of considerable fundamental or agro-industrial interest [3]. Different classification systems for aminopeptidases have been proposed [2,3]. Bacterial APs have been classified into 14 families using different criteria, such as substrate specificity, peptide sequence similarity, physicochemical and enzymatic properties [3]. APs have also been subdivided into three groups on the basis of their catalytic mechanism: metallo-aminopeptidases, the activities of which are regulated by the presence of divalent metallic cations, and cysteine and serine aminopeptidases, classified on the basis of their sensitivities to various types of inhibitors. The metallo-aminopeptidases constitute the largest group of APs and Zn^{2+} appears to be the most frequently associated cation. To date, only structures of metallo-aminopeptidases have been determined by X-ray crystallography [4–10]. In contrast, the cysteine and serine APs have no ionic cofactor [2]. Catalysis requires a highly reactive cysteine or serine residue the nucleophilicity of which must be enhanced by the local environment. Reactive serine residues have been detected for enzymes of the proline iminopeptidase and X-proline dipeptidyl aminopeptidase families (families 12 and 13 according to [3]).

Three serine aminopeptidases have been detected in *Ochrobactrum anthropi* [11–13]. The first two APs, D-aminopeptidase (DAP) and D-aminopeptidase B (DmpB), isolated from strains SCRC C1-38 and LMG7991, respectively, are homologous and display strict D stereospecificity [11–13]. Their catalytic activity seems to rely on the tetrad Ser/Lys/Ser/Glu, with the first two residues found in an SXXK motif (single-letter amino acid code) [12–14]. The lysine residue is believed to act as a Brønsted base. The biological function of DAP and DmpB is still unknown, but these enzymes display approximately 25% sequence identity with a *Streptomyces* R61 DD-carboxypeptidase. DAP and

DmpB are inhibited by β-lactam compounds which suggests that they are both members of the family of 'penicillin-recognising enzymes', as originally proposed for DAP by Asano and coworkers [12]. The third serine aminopeptidase from *O. anthropi*, DmpA, isolated from strain LMG7991 [13], is an L-aminopeptidase which also shows the unique ability to hydrolyse both D-amides and D-esters.

The *dmpA* gene has been cloned and overexpressed in *Escherichia coli* [13]. It has been completely sequenced on both strands and the deduced amino acid sequence does not exhibit significant similarity with known APs. It does show varying degrees of similarity with sequences corresponding to several open reading frames found in the genomes of other bacteria for which translation products have not yet been characterised [13]. DmpA is an AP that liberates the N-terminal residues from peptide substrates and the efficiency of the enzyme increases with peptide length. To allow recognition, the N-terminal residue must be in the L configuration and its α-amino group must be free. The enzyme's affinity profile for residues at the first N-terminal position in dipeptides is (Arg, Lys) > Phe > aliphatic amino acids (Leu, Gly, Ala) > hydroxylated amino acid (Ser). An acidic residue at the first or second position prevents hydrolysis, suggesting the presence of a negative charge in the substrate-binding pocket [14]. Although DmpA is an L-aminopeptidase it also shows D-amidasic and D-esterasic activities on D-alanine derivatives [13]. To our knowledge, this enzyme is so far the only one that can liberate both N-terminal D and L amino acids.

DmpA is synthesised as a single polypeptide precursor. The active form consists of two different peptides resulting from the unique cleavage of the Gly249–Ser250 peptide bond of the precursor. Site-directed mutagenesis studies revealed that both residues are essential for protein maturation and catalysis [15]. The fact that the cleavage site is recognised both in *O. anthropi* and *E. coli*, as well as sequence comparison of this site with those of enzymes of the N-terminal nucleophile (Ntn) hydrolase family, led us to propose that DmpA may be the prototype of a new Ntn hydrolase family [15].

A number of structures of Ntn hydrolases have been determined: penicillin acylase (PA) [16,17], the proteasome subunits (PRO) [18,19], glycosyl asparaginase (AGA) [20,21], and the glutamine 5-phosphoribosyl-1-pyrophosphate amidotransferases (GATs) from *Bacillus subtilis* and *E. coli* [22–24]. The functions, modes of activation and folds of these enzymes have been reviewed [25–27]. Ntn hydrolase enzymes are amidohydrolases characterised by their unusual utilisation of an N-terminal nucleophile (threonine, serine or cysteine) that appears to use its own α-amino group as a general base in the catalytic mechanism. This catalytic N-terminal residue is generated by a self-catalysed protein splicing process [27] and is situated at the extremity of a β sheet. Ntn hydrolases share a common fold, consisting of a core of two stacked antiparallel β sheets flanked on both sides by helices, which results in the capacity for nucleophilic attack as well as the possibility of autocatalytic processing [26].

Here we report the three-dimensional structure of DmpA at 1.82 Å resolution and present a comparison with known structures of APs and Ntn hydrolases. The structure confirmed that DmpA belongs to the Ntn hydrolase family and allowed us to propose a catalytic mechanism for this enzyme. Its novel connectivity between secondary structure elements, however, suggests that the consensus Ntn superfamily may have to be re-evaluated.

Results and discussion

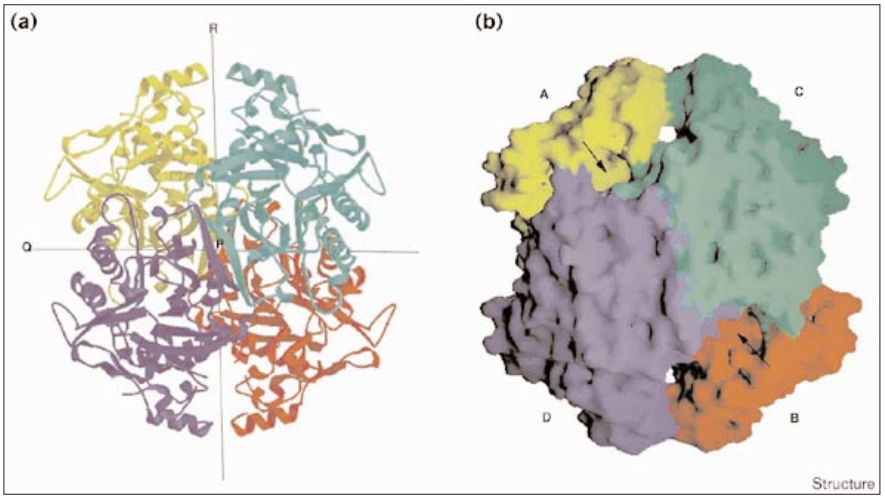
Quaternary structure

The observed contacts between molecules in the asymmetric unit initially suggested that DmpA was organised as a homotetramer. Gel-filtration analysis has confirmed this hypothesis (C Goffin, unpublished results). The four molecules of the tetramer were called A, B, C and D; when specifying an amino acid residue, we attach the suffix A, B, C or D accordingly. The DmpA homotetramer is a doughnut-shaped molecule (Figure 1). The four subunits are related mutually by perpendicular twofold axes, labelled P, Q and R (Figure 1a). Contacts along the molecular Q axis (2195 $Å^2$) are much more extensive than those along the P and R axes (1338 $Å^2$ and 1305 $Å^2$, respectively). In total, 31.5% of the total surface area of a subunit is involved in tetramer formation. The substrate-binding site of a monomer is located at the interface with its Q- and R-related subunits (Figure 1b).

Overall structure of the monomer

The processed DmpA molecule consists of two chains (α and β) containing the residues 1–249 and 250–375, respectively. It folds into a single domain and has an ellipsoidal shape of approximate dimensions 66 Å × 50 Å × 30 Å (Figures 2a,b). The central motif consists of nine β strands (S_1 to S_9) and four α helices (H_2 to H_5) organised in an αββα sandwich fold in which the two stacked mixed β sheets (sheets I and II) are flanked on both sides by two α helices (Figures 2b,c). Two extra helices (H_1 and H_6) and one strand (S_{10}) are involved in the crystal packing. The topology of the β strands may be described as (–5X, +1, +1, +1X) and (+1, +1, +1X) for sheets I and II, respectively (nomenclature according to [28]). In each sheet there are two parallel strands connected to each other via an α helix (Figure 2c). Sheets I and II have a left-handed twist of about 50° and 30°, respectively, and one is rotated relative to the other through a positive dihedral angle of about +30°. The β chain of the molecule (residues 250–375) is situated only in the 'upper' part of the structure (Figures 2b,c) forming the two parallels β strands (S_8 and S_9) of sheet I, the two upper helices (H_4 and H_5), and

Figure 1



The DmpA homotetramer. **(a)** Ribbon diagram of the DmpA homotetramer viewed along the P axis. The perpendicular Q and R axes are shown. Individual subunits are coloured separately: molecules A, B, C and D are coloured yellow, red, green and blue, respectively. **(b)** Surface view of the DmpA homotetramer; the arrows indicate the position of the substrate-binding pockets of subunits A and B. (The figure was produced using the programs MOLSCRIPT [45], RASTER3D [46] and GRASP [47].)

two long loops ($L_{S8–H5}$ and $L_{H5–H6}$) wrapping the β sandwich on both sides. The β sandwich is open on the side containing the substrate-binding pocket and the C-terminal region of the α chain. The opposite side is enclosed by the long $L_{S8–H5}$ loop and two short loops, $L_{S2–S3}$ and $L_{H1–S4}$. The β sandwich is flanked by the α helices H4 and H5 at the top, and by the helices H2 and H3 at the bottom; both pairs of helices are parallel and tilted by 40° relative to each other. Two 3_{10} helices (D_1 and D_2) are inserted between both sheets, causing the sheets to move away from each other on the front side of the molecule.

Substrate-binding pocket

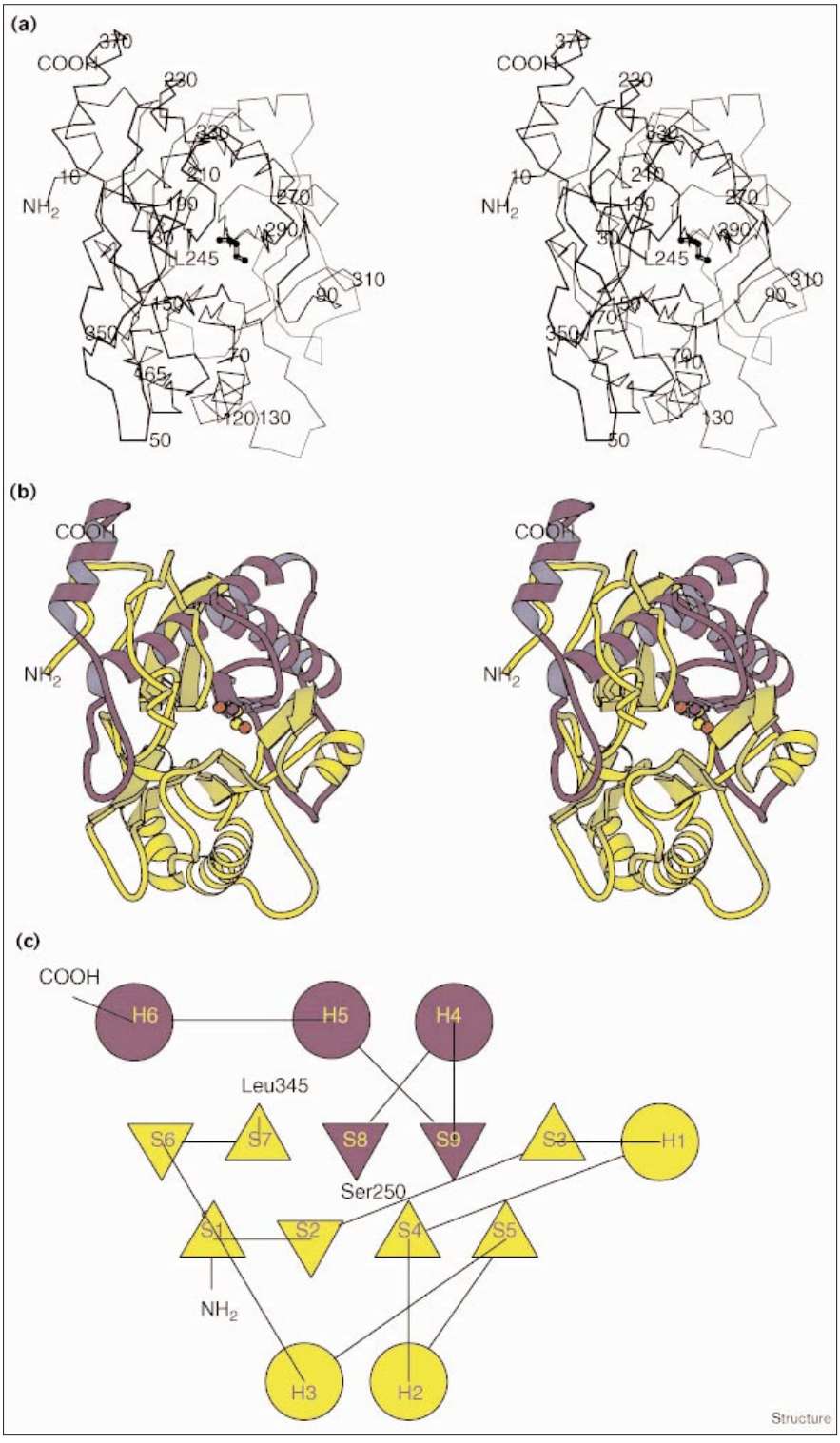
Several protease inhibitors were tested on DmpA without success [15]. All attempts to obtain complex structures of DmpA with potential inhibitors or reaction products by soaking crystals and cocrystallization trials were unsuccessful. In order to identify residues potentially involved in the substrate-binding site we manually introduced D-Ala-L-Ala and L-Ala-L-Ala dipeptides into the active site of DmpA on the basis of the known structure of the complexes between AGA and aspartic acid [21] and between PA and phenylmethylsulfonyl fluoride (PMSF) [16] (Figure 3a).

The substrate-binding site consists of a pocket, fairly open to the solvent, which is situated at the interface of an enzyme molecule and its Q- and R-related subunits (molecules labelled respectively D and C if the subunit A is chosen as the reference; Figure 1b).

The enzymatic properties of DmpA [15] suggest the presence of acidic residues in the substrate-binding site that would be responsible for the known affinity for basic sidechains, as well as for the stabilisation of the N-terminal α-amino group of the substrate. Glu144A, the only negatively charged residue in the vicinity of the active site (Figures 3a,b), is in an appropriate position to make a salt bridge with the free α-amino group of an L-amino acid which might also be hydrogen bonded to the γ-oxygen of Thr108A (Figure 3a). On the other side of the substrate-binding site there is an empty space allowing the binding of the sidechain of the N-terminal L-amino acid. The positive charge of a D-Ala N-terminal amino acid cannot be stabilised by the negative charge of Glu144A, as it is not well positioned to do so. However, the α-amino group can hydrogen bond to the sidechain γ-oxygen of Thr145. The steric hindrance due to the presence of both the sidechains of Glu144A and Thr108A precludes the binding of D-residues with a sidechain longer than that of D-Ala (Figure 3a). This observation may explain why the recognition of D-Xaa residues is limited to D-Ala, and also why the hydrolysis of L-isomers is much more efficient than that of D isomers. The Glu144A residue is the only negatively charged residue in the substrate-binding pocket, close to the active site, which can be responsible for the aminopeptidasic activity. Its presence is probably also responsible for the high affinity of the enzyme for basic residues and explains the lack of affinity for negatively charged residues situated at the first or second position in the peptide substrate.

A hydrophobic cluster has been identified close to the active site. The cluster, comprising residues Tyr146A, Phe135D, Leu136D and Trp137D, is located at the interface of molecules A and D and is in a good position to make hydrophobic interactions with the sidechain of a residue located downstream from the peptide bond to be cleaved (Figure 3a). This is consistent with kinetic studies which show that DmpA activity towards dipeptide substrates is increased by the presence of a phenylalanine residue at the second position from the N terminus [15].

Figure 2

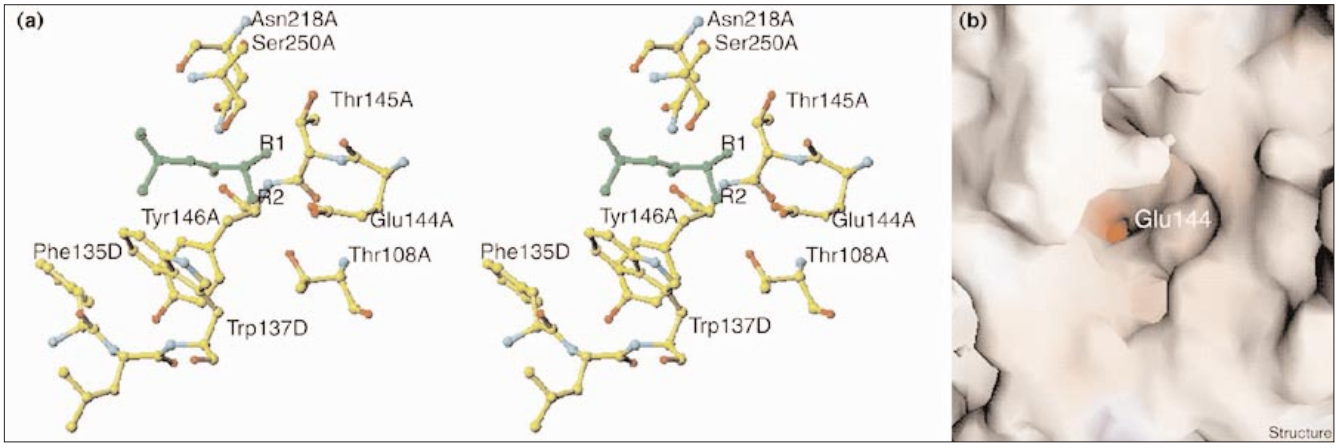


The DmpA heterodimer. **(a)** Stereoview Cα trace of the DmpA heterodimer, the catalytic serine is shown in ball-and-stick representation. **(b)** Stereoview ribbon drawing of DmpA. The α and β chains are coloured yellow and blue, respectively. The catalytic serine is shown in ball-and-stick representation and coloured by atom type. **(c)** Topological diagram of the DmpA fold; α helices (circles) and β strands (triangles) are labelled H1 to H6 and S1 to S9, respectively. (The figure was produced using the programs MOLSCRIPT [45] and TOPS [48].)

Comparison with known structures of aminopeptidases

The previously published three-dimensional structures of APs fall into a number of different structural families. The first determined structure of an AP was that of the bovine lens leucine aminopeptidase (LAP) [4,5]. LAP is a homo-hexameric enzyme and each subunit contains two zinc ions that are essential for catalytic activity [29]. Both metal ions participate in substrate binding and activation, and

Figure 3



The DmpA substrate-binding site. **(a)** Stereoview of the residues of DmpA likely to be involved in substrate binding (coloured by atom type). The natural substrate has been modelled and is represented in green. R1 and R2 represent either the sidechain and the α-amino group if the N-terminal residue is L-Ala, or the the α-amino group and the sidechain if the N-terminal residue is D-Ala. **(b)** The substrate-binding site of DmpA represented by its electrostatic surface potential: red and blue indicate negative and positive electrostatic potential, respectively. The Glu144 sidechain is the only negative charge in the vicinity of the active site. (The figure was produced using the programs TURBO-FRODO [43] and GRASP [47].)

have a possible role in the activation of the nucleophile [30,31]. The second AP structure to be determined was that of a methionine aminopeptidase (MAP) from *E. coli* [8]. Recently, the structure of eukaryotic MAPs have also been solved [9,10]. MAPs contain two cobalt ions in the active site and are unrelated in structure and in sequence to LAP. The most recent AP structures to be determined were those from *Aeromonas proteolytica* (AAP) [6] and *Streptomyces griseus* (SGAP) [7]. Both AAP and SGAP require two zinc ions for activity. Despite their low level of sequence identity, they share a similar topology and a similar zinc coordination suggesting that the two enzymes are likely to have similar catalytic mechanisms [6,7]. These two microbial enzymes differ from LAP in both overall structure and in coordination of the two zinc ions. This led the authors to classify AAP and SGAP separately from LAP. The three-dimensional structure of a proline iminopeptidase from *Xanthomonas campestris* pv. *citri* has been solved at 2.7 Å resolution [32]. The protein is folded into two contiguous domains with the active site located at the interface. The enzyme has a Ser-Asp-His catalytic triad and displays an overall topology similar to that of yeast serine carboxypeptidase [33].

Clearly, DmpA is structurally and mechanistically distinct from any of the previously described APs: it has no ionic

Figure 4

A comparison of the topology of DmpA and the Ntn hydrolase fold. **(a)** Topological diagram of the Ntn hydrolase consensus fold [25]. Open circles (helices) and triangles (strands) represent secondary structure elements occurring in the same position and sequence in PA, GAT, AGA and PRO. Elements shown in grey are structurally conserved but not always in the same sequence order. **(b)** Topological diagram of the fold of DmpA. The circles and triangles represent secondary structure elements occurring in the same position as in other Ntn amidohydrolases. Elements of the β chain are surrounded by a thick line and elements of the α chain by a thin line. Secondary structure elements are numbered according to [25].

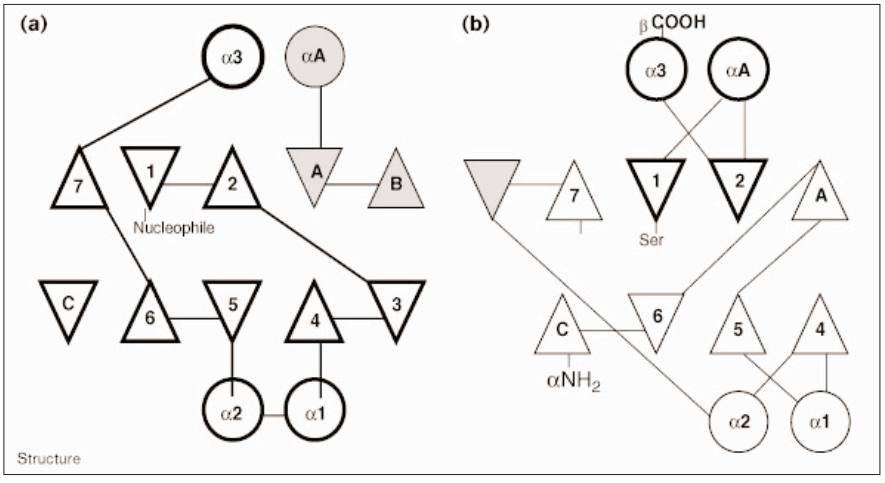
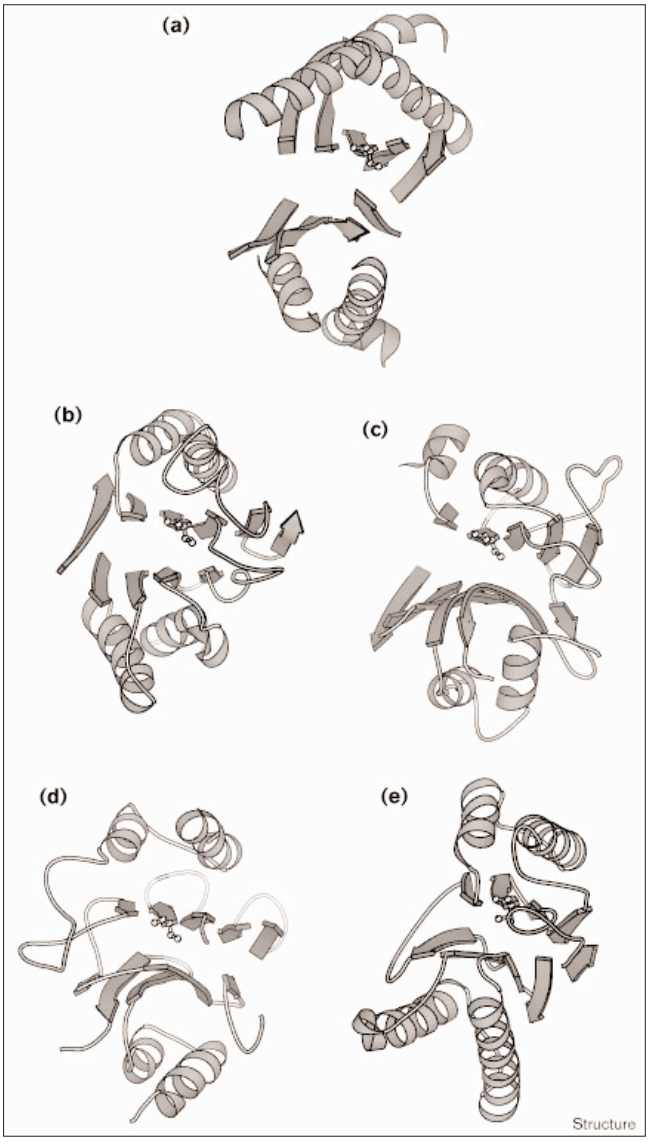


Figure 5



Comparison of the folds of the catalytic domains of **(a)** DmpA, **(b)** AGA, **(c)** GAT, **(d)** PA and **(e)** PRO. The common folding elements and connections and the catalytic N-terminal residues of these enzymes are shown. (The figure was produced with the program MOLSCRIPT [45].)

cofactor but possesses the Ntn hydrolase fold (see below), which provides both the capacity for nucleophilic attack and the possibility of autocatalytic processing [26]. The three-dimensional structure of DmpA identifies this protein as the first example of a novel structural family of APs.

Structural comparison with other Ntn hydrolases

As described previously, all the enzymes belonging to the Ntn hydrolase family have been shown to share highly homologous folds and functions [25,26]. A consensus Ntn fold has been proposed and is shown in Figure 4. Comparison of the structures of PA, PRO, GAT, AGA and DmpA shows that the newly determined structure fits well into the Ntn superfamily (Figure 5).

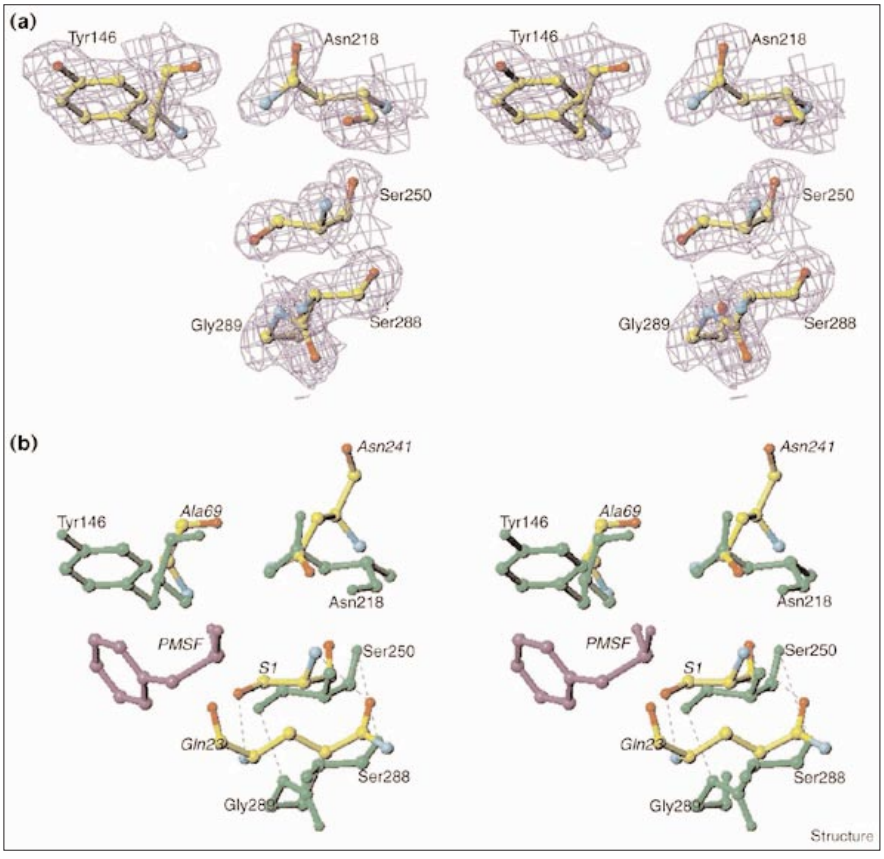
Several catalytic mechanisms have been proposed for Ntn hydrolases on the basis of analogy with the serine proteases. The catalytic N-terminal nucleophile uses its own α-amino group as a general base with an oxyanion hole stabilising the presumed tetrahedral intermediate. The key elements in catalysis are the oxygen or sulphur nucleophile provided by a threonine, serine or cysteine residue, the Brønsted base provided by the α-amino group of the nucleophile, and the oxyanion hole. The oxyanion hole may be composed of a mainchain amide group plus either a hydroxyl group of a polar sidechain or the Nδ2 group of an asparagine residue. Additional noncovalent interactions in the vicinity of the catalytic residues contribute to catalysis [34].

A structural alignment of DmpA with Ntn hydrolases using the conserved β strands reveals that all the catalytic centres overlap. This allowed us to identify active-site residues in DmpA (Figure 6). The enzyme possesses a catalytic serine residue, Ser250, which corresponds to the known catalytic residues of the other enzymes (PA, SerB1; GAT, Cys1; AGA, Thr206; PRO, Thr1B) and its α-amino group is the only candidate in the active site that might enhance the nucleophilic character of the hydroxyl group. The backbone NH group of Tyr146 is in a good position to play the role of the first element of the oxyanion hole (corresponding to PA, AlaB69; GAT, Gly103; AGA, Gly235; PRO, Gly47B). In all previously described Ntn hydrolase structures, this function has been situated at the end of strand β4. In DmpA, however, Tyr146 is located at the end of strand β5 of sheet II. The role of the second contributor to the oxyanion hole (PA, AsnB241 Nδ2; GAT, Asn102 Nδ2; AGA, Thr234 Oγ; PRO, Ser129 Oγ) is most likely to be played by Nδ2 of Asn218. The α-amino group of Ser250 hydrogen bonds with the Oγ of Ser288 and the backbone NH group of Gly289 hydrogen bonds to the nucleophilic hydroxyl group (Figure 6b). It is likely that, by analogy with other Ntn hydrolases, these residues could be indirectly involved in the catalytic mechanism, affecting the properties of the catalytic residue. On the basis of these homologies, and according to the widely accepted mechanism for other Ntn hydrolases, we propose a catalytic mechanism for DmpA (Figure 7). Several acidic amino acid residues are present on the surface of the substrate-binding site. However, only Glu144 is located in the vicinity of the catalytic centre and is probably involved in the stabilisation of the N-terminal α-amino group of the substrate.

The DmpA structure reveals several significant differences from the consensus Ntn fold. These differences concern the direction and connectivity of the secondary

Figure 6

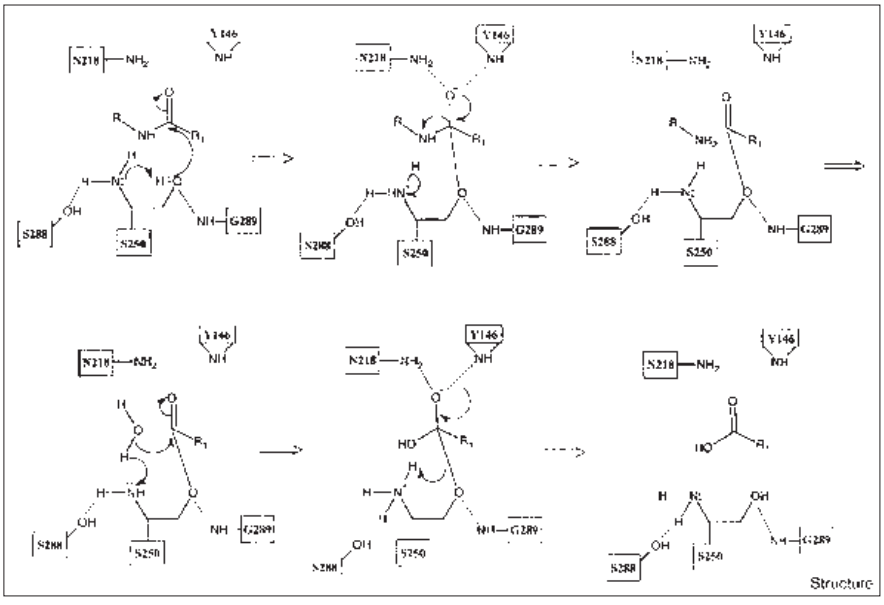
The catalytic site of DmpA. **(a)** Stereoview electron-density map around residues of the active site of DmpA. **(b)** Stereoview superposition of residues involved in the catalytic mechanism of DmpA (coloured green) with those of PA (coloured by atom type) in complex with PMSF (coloured violet). (The figure was produced using the program TURBO-FRODO [43].)



structure elements. The characteristic four-layer αββα topology observed in the Ntn fold, composed of two antiparallel β sheets flanked on both sides by two antiparallel α helices, is replaced here by two mixed β sheets

Figure 7

The catalytic mechanism of DmpA. The key elements are the catalytic residue Ser250, which plays the role of both nucleophile (the Ser250 hydroxyl group) and the base (the Ser250 α-amino group), and the two elements of the oxyanion hole (the mainchain nitrogen atom of Tyr146 and the Nδ2 group of Asn218). The Ser250 α-amino group acts as a base, increasing the nucleophilicity of the Ser250 hydroxyl group which attacks the carbonyl carbon of the substrate. This results in the formation of a covalent enzyme–substrate transition state that is stabilised by the oxyanion hole. This complex is disrupted and the products liberated through the addition of water. Ser288 and Gly289 modify the properties of Ser250.



flanked by two parallel α helices (Figure 4). In DmpA, each mixed β sheet contains two parallel strands connected by a single α helix, S_1,H_4,S_2 in sheet I and S_4,H_1,S_5 in sheet II. The two helices situated on both sides of the stacked β sheets lie parallel to each other. In the consensus Ntn fold, the equivalent strands are antiparallel and are connected by two antiparallel helices. DmpA also shows additional differences in connectivity compared to other Ntn hydrolases and has no counterparts to strands 3 and B (Ntn nomenclature).

Despite the fact that these differences constitute a significant deviation from the consensus Ntn fold proposed by Artymiuk [25], the four helices, strands 1, 2, 4, 5, 6, A and C, and the catalytic N-terminal residue are structurally equivalent. All the residues involved in the catalytic mechanism are in similar locations in the active sites and the enzymes are likely to employ the same catalytic mechanism. Furthermore, DmpA is an amidohydrolase activated by the cleavage of the peptide bond preceding the catalytic residue (Ser 250) at the beginning of a β strand, as is the case with all other Ntn hydrolases. This post-translational processing is probably self-catalysed by the enzyme [15]. All these observations suggest that DmpA should be considered as a member of the Ntn hydrolase structural superfamily, the definition of which must now be enlarged to take into account the variations observed in the DmpA structure.

Biological implications

Ochrobactrum anthropi **is an emerging nosocomial pathogen. This bacterium can be responsible for meningitis or septicemia in an immunocompromised patient. *O. anthropi* is a highly proteolytic bacterial species that extracts amino acid residues from peptides present in the medium using a pool of peptidases. Among these peptidases, DmpA has the unusual property of recognising both D-Ala and L amino acids situated at the N-terminal ends of peptides. This wide specificity profile is likely to contribute to the hydrolysis of a vast number of small peptides transported into the cell. On the basis of its biochemical and enzymatic properties, DmpA has been proposed as a new member of the N-terminal nucleophile hydrolase (Ntn) family. The enzymes of this family share similar functions, modes of activation, and polypeptide folds, but exhibit no sequence similarity. All Ntn hydrolase enzymes are amidohydrolases characterised by their unusual use of an N-terminal nucleophile (threonine, serine or cysteine) as the catalytic residue. The autocatalytic cleavage of the polypeptide chain of the proenzyme immediately upstream to this catalytic residue constitutes the activation step of Ntn hydrolases.**

The crystal structure of DmpA at 1.82 Å resolution reported here shows that the polypeptide folds into a single domain organised as an αββα sandwich. This structure shows a clear similarity to the fold of the Ntn hydrolases of known structure. A structural alignment, using the β sheets as reference, showed the catalytic region of these enzymes to overlap and thus allowed the determination of residues involved in the enzymatic activity of DmpA. On this basis a catalytic mechanism for the aminopeptidase was proposed.

The *O. anthropi* DmpA structure determination demonstrates that this enzyme belongs to the Ntn hydrolase enzyme family and that it probably uses the same mechanism to hydrolyse the peptide bond situated at the N-terminal extremity of the substrate. However, the direction and connectivity of the secondary structure elements of DmpA are significantly different from the consensus Ntn hydrolase topology previously described. The DmpA fold thus characterises a new subfamily of Ntn hydrolase enzymes.

Materials and methods

Crystal data

DmpA was purified and crystallised as described previously [13,35]. The crystals belong to the orthorhombic space group $P2_12_12$ and contain six DmpA molecules per asymmetric unit. The cell dimensions are a = 156.97 Å, b = 96.22 Å and c = 154.41 Å. Heavy-atom derivatives for the enzyme were prepared at 21°C by soaking crystals in heavy-atom solutions (see Table 1 for statistics on native and heavy-atom derivatives). One native and several derivative X-ray diffraction data sets were collected on a MacScience DIP2030 image plate system, using CuKα radiation produced by a Nonius FR591 rotating-anode generator equipped with a double mirror X-ray optical system and running at 100 mA, 45 kV. The diffraction data were auto-indexed, processed, scaled and merged using DENZO and SCALEPACK from the HKL package [36]. Further native data were collected, at 4°C, from a single crystal using a MAR-Research Mar345 imaging plate detector on beamline DW32 (λ = 0.97 Å) at the LURE Orsay synchrotron outstation (Table 1).

Structure determination and refinement

The structure was solved by the multiple isomorphous replacement (MIR) method utilising noncrystallographic symmetry (NCS) averaging. Heavy-atom positions were located by difference Patterson maps and difference Fourier maps using the CCP4 package [37]. Heavy-atom positions, occupancies and isotropic B factors were refined with the program MLPHARE [37] to 2.3 Å resolution yielding a mean figure of merit of 0.3. The initial MIR electron-density map displayed clear molecular boundaries, but was otherwise uninterpretable. At this stage, NCS averaging using DM [37] was employed. NCS matrices were calculated by manual inspection of all the heavy atom interatomic vectors, calculated using DISTANG [37]. Four of the six molecule positions could be readily identified. Fourfold averaging was performed using DM, with automatic mask determination and combined with iterative real-space improvement of the NCS matrices. The electron-density map calculated using these phases was of excellent quality and allowed us to build about 80% of the structure, using the program O [38]. The last two molecule positions in the asymmetric unit were then found by molecular replacement using the program suite AMoRe [39] with the partial DmpA structure as search model.

Five percent of the observations were set aside for cross-validation analysis [40] and the behaviour of R_{free} and the likelihood free figure of merit was used to monitor various refinement strategies. Atomic models were refined applying NCS restraints using a maximum-likeli-

Table 1

Data collection and structure determination statistics.

Data set	Native I	Native II	PHMSA*	$HgCl_2$	K_2PtCl_4	PHMSA* + K_2PtCl_4
Soaking conditions			10 mM, 10 weeks	10 mM, 10 weeks	20 mM, 2 h	5 mM + 5 mM, days
Data collection						
Resolution range (Å)	20–2.3	12.5–1.82	20–2.4	20–2.4	20–2.7	20–3.3
(last shell)	(2.4–2.3)	(1.9–1.82)	(2.5–2.4)	(2.5–2.4)	(2.8–2.7)	(3.4–3.3)
Completeness (%)	90.9 (83.8)	95.8 (91.9)	97.0 (92.3)	88.4 (84.8)	88.6 (82.5)	86.5 (83.0)
Multiplicity	4.5	3.5	4.7	3.4	3.4	3.3
R_{merge}† (%)	8.3 (25.5)	5.8 (28.8)	9.5 (27.3)	8.1 (25.0)	9.2 (30.1)	15.7 (29.4)
$\langle I \rangle / \langle \sigma I \rangle$	16.4 (3.5)	26.03 (6.9)	15 (2.9)	11.1 (2.3)	13.3 (2.5)	10.2 (2.7)
MIR phasing (20–2.7 Å)						
Number of sites (occupancy)			6 (1.3–1.7)	6 (0.7–2.1)	9 (0.35–0.8)	15 (0.35–2.2)
R_{Cullis}‡			0.67	0.60	0.70	0.80
Phasing power§			1.83	1.87	1.52	1.2
Refinement (12.5–1.82 Å)						
Reflections used ($F > 1\sigma$)		189,926				
No. protein atoms (molecule/AU#)		2747/16,482				
No. water molecules per AU#		1331				
R_{factor} (180,391 reflections) (%)		16.9				
R_{free} (9535 reflections) (%)		20.6				
Rms deviation from ideality						
bonds (Å)		0.013				
angles (Å)		0.034				
Average B factors (Å^2)						
mainchains/sidechains		18.4/23.8				
solvent		31.15				

*PHMSA, p-hydroxymercuriphenyl-sulfonic acid. †$R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the measured intensity and $\langle I \rangle$ is the average intensity obtained from multiple measurements of symmetry-related reflections. ‡R_{Cullis} = rms lack of closure/rms isomorphous difference. §Phasing power = $\langle |FH| \rangle$ / rms lack of closure for acentric reflections. #AU, asymmetric unit.

hood-based refinement strategy employing the program REFMAC [41], followed by manual fitting into SIGMAA-weighted [42] electron-density maps with TURBO-FRODO [43]. As all observed data from 19.0 to 2.3 Å were used, a 2-Gaussian bulk-solvent correction was applied. After the refinement converged, the 1.8 Å data set was substituted for the 2.3 Å set and the refinement was cycled to convergence at 1.8 Å. The model was refined first with tight NCS restraints, which were gradually relaxed as appropriate. The final model contains 363 residues for each independent molecule (Table 1). Residues Met1 to Arg8 and Gln246 to Gly249 were not visible in the electron density and were therefore not included in the model, although their presence has been demonstrated by N-terminal [13] and C-terminal sequence analysis (data not shown). A summary of the refinement statistics is shown in Table 1. A Ramachandran plot of the mainchain torsion angle pairs, calculated with the program PROCHECK [44], shows all the residues to occur in allowed regions, with 91.4% of the residues being located in the most favoured regions.

Structural alignments of Ntn hydrolases structures were performed with the 'RIGID' option of TURBO-FRODO using both conserved β sheets and catalytic residues of the enzymes.

Accession numbers

The refined coordinates and structure factors for DmpA have been deposited in the Brookhaven Protein Data Bank and have the accession codes 1b65 and r1b65sf, respectively.

Acknowledgements

We thank Eleanor Dodson for crucial assistance with the map averaging procedure, Thierry Prangé for assistance in data collection at the LURE (Orsay, France) and Bart Samyn for the C-terminal sequence analysis. This work was supported by a project of the Fund for Scientific Research Flanders (No. 3G006896) and by the Belgian Program of Interuniversity Poles of Attraction (PAI No. P4/03). The University of York was funded by the provision of a Structural Biology Centre award from the Biotechnology and Biological Sciences Research Council. GJD is a Royal Society University Research Fellow.

References

1. Taylor, A. (1993). Aminopeptidases: towards a mechanism of action. *Trends Biochem. Sci.* **18**, 167-172.
2. Taylor, A. (1993). Aminopeptidases: structure and function. *FASEB J.* **7**, 290-298.
3. Gonzales, T. & Robert-Baudouy, J. (1996). Bacterial aminopeptidases: properties and functions. *FEMS Microbiol. Rev.* **18**, 319-344.
4. Burley, S.K., David, P.R., Taylor, A. & Lipscomb, W.N. (1990). Molecular structure of leucine aminopeptidase at 2.7 Å resolution. *Proc. Natl Acad. Sci. USA* **87**, 6878-6882.
5. Burley, S.K., David, P.R., Sweet, R., Taylor, A. & Lipscomb, W.N. (1992). Structure determination and refinement of bovine lens leucine aminopeptidase and its complex with bestatin. *J. Mol. Biol.* **224**, 113-140.
6. Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.-M., Moras, D. & Tarnus, C. (1994). Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure* **2**, 283-291.
7. Greenblatt, H.M., *et al.,* & Shoham, G. (1997). *Streptomyces griseus* aminopeptidase: X-ray crystallographic structure at 1.75 Å resolution. *J. Mol. Biol.* **265**, 620-636.
8. Roderick, S.L. & Matthews, B.W. (1993). Structure of the cobalt-dependent methionine aminopeptidase from *Escherichia coli*: a new type of proteolytic enzyme. *Biochemistry* **32**, 3907-3912.
9. Liu, S., Widom, J., Kemp, C.W., Crews, C.M. & Clardy, J. (1998). Structure of human methionine aminopeptidase-2 complexed with fumagillin. *Science* **282**, 1324-1327.
10. Tahirov, T.H., *et al.*, & Kato, I. (1998). Crystal structure of methionine aminopeptidase from hyperthermophile, *Pyrococcus furiosus. J. Mol.*

Biol. **284**, 101-124.
11. Asano, Y., Nakazawa, A., Kato, Y. & Kondo, K. (1989). Properties of a novel D-stereospecific aminopeptidase from *Ochrobactrum anthropi*. *J. Biol. Chem.* **264**, 14233-14239.
12. Asano, Y., Kato, Y., Yamada, A. & Kondo, K. (1992). Structural similarity of D-aminopeptidase to carboxypeptidase DD and β-lactamases. *Biochemistry* **31**, 2316-2328.
13. Fanuel, L., *et al.,* & Frere, J.M. (1999). Two new aminopeptidases from *Ochrobactrum anthropi* active on D-alanyl-*p*-nitroanilide. *Cell. Mol. Life Sci.* **55**, 812-818.
14. Rawlings, N.D. & Barrett, A.J. (1993). Evolutionary families of peptidases. *Biochem. J.* **290**, 205-218.
15. Fanuel, L., *et al.,* & Frere, J.M. (1999). The DmpA aminopeptidase from *Ochrobactrum anthropi* LMG7991 is the prototype of a new terminal nucleophile hydrolase family. *Biochem. J.* **341**, 147-155.
16. Duggleby, H.J., Shirley, P.T., Christopher, P.H., Dodson, E.J., Dodson, G. & Moody, P.C.E. (1995). Penicillin acylase has a single-amino-acid catalytic center. *Nature* **373**, 264-268.
17. Suresh, C.G., *et al.,* & Dodson, G.G. (1999). Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. *Nat. Struct. Biol.* **6**, 414-416.
18. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). Crystal structure of the 20S proteasome from the Archeaon *T. acidophilum* at 3.4 Å resolution. *Science* **268**, 533-539.
19. Groll, M., *et al.,* & Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**, 463-471.
20. Guan, C., *et al.,* & Comb, D. (1996). Activation of glycosylasparaginase. Formation of active N-terminal threonine by intramolecular autoproteolysis. *J. Biol. Chem.* **271**, 1732-1737.
21. Oinonen, C., Tikkanen, R., Rouvinen, J. & Peltonen, L. (1995). Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nat. Struct. Biol.* **2**, 1102-1108.
22. Smith, J.L., *et al.,* & Satow, Y. (1994). Structure of the allosteric regulatory enzyme of purine biosynthesis. *Science* **264**, 1427-1433.
23. Isupov, M.N., *et al.,* & Teplyakov, A. (1996). Substrate binding is required for assembly of the active conformation of the catalytic site in Ntn amidotransferase: evidence from the 1.8 Å crystal structure of the glutaminase domain of glucosamine 6-phosphate synthase. *Structure* **4**, 801-810.
24. Muchmore, C.R.A, Krahn, J.M., Hyun Kim, J., Zalkin, H. & Smith, J.L. (1998). Crystal structure of glutamine phosphoribosylpyrophosphate amidotransferase from *Escherichia coli*. *Protein Sci.* **7**, 39-51.
25. Artymiuk, P.J. (1995). A sting in the (N-terminal) tail. *Nat. Struct. Biol.* **2**, 1035-1037.
26. Brannigan, J.A., *et al.,* & Murzin, A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self activation. *Nature* **378**, 416-419.
27. Shao, Y. & Kent, S.B.H. (1997). Protein splicing: occurrence, mechanisms and related phenomena. *Chem. Biol.* **4**, 187-194.
28. Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **34**, 167-339.
29. Kim, H. & Lipscomb, W.N. (1994). Structure and mechanism of bovine lens leucine aminopeptidase. *Adv. Enzymol.* **68**, 153-213.
30. Sträter, N. & Lipscomb, W.N. (1995). Transition state analogue L-leucinephosphonic acid bound to bovine lens leucine aminopeptidase: X-ray structure at 1.65 Å resolution in a new crystal form. *Biochemistry* **34**, 9200-9210.
31. Sträter, N. & Lipscomb, W.N. (1995). Two metal ion mechanism of bovine lens leucine aminopeptidase: active site solvent structure and binding mode of L-leucinal, a gem-diolate transition state analogue, by X-ray crystallography. *Biochemistry* **34**, 14792-14800.
32. Medrano, F.J., Alonso, J., Garcia, J.L., Romero, A., Bode, W. & Gomis-Rüth, F.X. (1998). Structure of proline iminopeptidase from *Xanthomonas compestris* pv. *Ctri*: a prototype for the prolyl oligopeptidase family. *EMBO J.* **17**, 1-9.
33. Endrizzi, J.A., Breddam, K. & Remington, S.J. (1994). 2.8 Å Structure of yeast serine carboxypeptidase. *Biochemistry* **33**, 11106-11120.
34. Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, J. & Peltonen, L. (1996). Functional analyses of active site residues of human lysosomal aspartylglucosaminidase: implications for catalytic mechanism and autocatalytic activation. *EMBO J.* **15**, 2954-2960.
35. Bompard-Gilles, C., Villeret, V., Fanuel, L., Joris, B., Frère, J.M. & Van Beeumen, J. (1999). Crystallization and preliminary X-ray analysis of a new L-aminopeptidase-D-amidase/D-esterase activated by a Gly-Ser peptide bond hydrolysis. *Acta Crystallogr. D* **55**, 699-701.
36. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
37. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763.
38. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density and the location of errors in these models. *Acta Crystallogr. A* **47**, 110-119.
39. Navazza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157-163.
40. Brünger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.
41. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53**, 240-255.
42. Read, R.J. (1986). Improved Fourier coefficients for maps using phase from partial structure with errors. *Acta Crystallogr. A* **42**, 140-149.
43. Roussel, A. & Cambillau, C. (1992). TURBO-FRODO. In *Silicon Graphics Geometry Directory Vol.86.* Silicon Graphics, Mountain View, CA.
44. Laskowski, R.A., McArthur, M.W., Moss, D.S. & Thornton, J. (1993). PROCHECK: a program to check the quality of protein structures. *J. Appl. Crystallogr.* **26**, 282-291.
45. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950.
46. Merrit, E.A. & Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869-873.
47. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association: insight from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.
48. Westhead, D.R., Slidel, T.W., Flores, T.P. & Thornton, J.M. (1999). Protein structural topology: automated analysis and diagrammatic representation. *Protein Sci.* **8**, 897-904.