The First Structure of Dipeptidyl-peptidase III Provides Insight into the Catalytic Mechanism and Mode of Substrate Binding

Pravas Kumar Baral†, Nina Jajčanin-Jozić, Sigrid Deller†, Peter Macheroux§, Marija Abramiec, and Karl Gruber‡

From the †Institute of Molecular Biosciences, University of Graz, A-8010 Graz, Austria, the ‡Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, HR-10002 Zagreb, Croatia, and the §Institute of Biochemistry, Graz University of Technology, A-8010 Graz, Austria

Dipeptidyl-peptidases III (DPP III) are zinc-dependent enzymes that specifically cleave the first two amino acids from the N terminus of different length peptides. In mammals, DPP III is associated with important physiological functions and is a potential biomarker for certain types of cancer. Here, we present the 1.95-Å crystal structure of yeast DPP III representing the prototype for the M49 family of metallopeptidases. It shows a novel fold with two domains forming a wide cleft containing the catalytic metal ion. DPP III exhibits no overall similarity to other metallopeptidases, such as thermolysin and neprolysin, but zinc coordination and catalytically important residues are structurally conserved. Substrate recognition is accomplished by a binding site for the N terminus of the peptide at an appropriate distance from the metal center and by a series of conserved arginine residues anchoring the C termini of different length substrates.

A number of bioactive peptides that play crucial roles in the nervous and endocrine systems have been identified. These peptides interact with corresponding receptors and subsequently modulate downstream pathways. The consequences of such processes are terminated by the degradation of those active peptides either by peptidases or through internalization by other mechanisms. A number of peptidases that act on physiologically important peptides have been identified and are classified into different groups according to their mode of action (exo- versus endopeptidases), catalytic signature motifs, or metal content (1). Because of their role in the organism, inhibition of peptidases has therapeutic implications, such as the treatment of hypertension in the case of inhibition of angiotensin I-converting enzyme (2).

Dipeptidyl-peptidase III (DPP III) EC 3.4.14.4 enzymes form a group of aminopeptidases with molecular masses of ~80–85 kDa. They specifically cleave dipeptides from the N termini of their substrates and have restricted specificity on dipeptidyl-arylamine substrates, preferring Arg-Arg-2-naphthylamide (3). The cDNAs for human, rat, and fruit fly DPP III were cloned, and their respective amino acid sequences (737, 738, and 786 amino acids long) were deduced (3–5). Based on complete genome sequence data, orthologs have also been identified in >50 species, including lower eukaryotes (yeasts) and some specific bacteria (6). In rat, DPP III is found to be expressed in various tissues with high levels in brain, liver, small intestine, and kidney (7, 8). It is generally found to be a cytosolic protein, but membrane-associated DPP III in rat brain has been reported as well (9). DPP III from human seminal plasma (10), lens tissue (11), and erythrocytes (12) has been purified and biochemically characterized, as were the enzymes from several other mammalian tissues, including bovine pituitary (13), rat brain (9), and liver (8). Regarding non-mammals, DPP III has been partially purified and characterized from the slime mold Dicyostelium discoideum (14) and Drosophila melanogaster (15).

The exact physiological roles of these enzymes are not yet clear. However, some pharmacological experiments link DPP III with pain regulation mechanisms because low levels of DPP III activity have been observed in the cerebrospinal fluid of individuals suffering from acute pain (16). Similarly, high concentrations of DPP III found in the superficial laminae of rat spinal cord dorsal horn (17), as well as the high in vitro affinity shown by the human enzyme toward important neuropeptides, such as enkephalins and endomorphins (18), also indicate that this hydrolase could play a role in the mammalian pain modulatory system. These findings make DPP III a potential drug target, and efforts toward inhibitor design and synthesis are under way (19).

Recently, DPP III has also obtained much attention because of its overexpression in ovarian malignant tissues, and this
property can be further exploited as a potential biomarker for carcinoma (20). Furthermore, a role for DPP III has been implied in cataractogenesis and endogenous defense mechanisms against oxidative stress (21, 22).

Biochemical investigations showed DPP III to be a zinc metallopeptidase containing 1 mol of zinc/mol of protein with a dissociation constant of $2.5 \times 10^{-13}$ M at pH 7.4 (4). Commonly, cobalt ions significantly activate DPP III from various origins (3, 23). All known DPP III sequences contain the unique motif HEXXGH, which enabled the recognition of the dipeptidyl-peptidase III family as a distinct evolutionary metallopeptidase family (family M49, MEROPS, the Peptidase Database (1), www.merops.ac.uk). Mutagenesis data on the HEXXGH motif in the rat enzyme imply that the histidine residues coordinate the metal ion and that the glutamic acid acts as a general base in peptide hydrolysis (24). A second conserved linear motif within the DPP III family, EEKR(K)AE(D), is found 22–55 amino acids toward the C terminus from the first one (6, 7), and the involvement of the second glutamic acid residue in the coordination of the active-site zinc was shown (24). Apart from that, little is known about the molecular mechanism of action of this hydropeptidase because no experimental three-dimensional structure of DPP III enzymes is available. Homology modeling was not possible because of the low sequence identity to proteins with known structure, already indicating a novel fold. Because of the more facile accessibility, we chose the ortholog from Saccharomyces cerevisiae, sharing ~40% sequence identity with mamalian DPP III, for our structural studies. The enzyme was cloned and expressed in Escherichia coli, and we present here its high resolution crystal structure and discuss the implications of this first DPP III structure for the catalytic mechanism and substrate recognition.

**EXPERIMENTAL PROCEDURES**

**Cloning of Yeast DPP III for Large-scale Expression in E. coli**—

The open reading frame YOL057W encoding yeast DPP III was amplified by PCR using genomic DNA isolated from S. cerevisiae (BY4741). The PCR product was inserted into the NdeI/Xhol restriction sites of pET21a (Novagen). Cloning into this vector using NdeI/Xhol and deletion of the stop codon allow expression of the protein with a hexahistidine affinity tag at the C terminus. Cysteine-to-serine point mutations in yeast DPP III were produced using the QuikChange XL site-directed mutagenesis kit (Stratagene) with pET21a-DPP II-His$_6$ as template.

**Expression and Purification of S. cerevisiae DPP III-His$_6$**—

A single colony of *E. coli* BL21-CodonPlus(DE3)-RII was grown overnight in 5 ml of Luria broth supplemented with 100 µg/ml ampicillin and 36 µg/ml chloramphenicol, which was then used to inoculate a 500-ml culture. After 3 h at 37°C, expression was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 0.3 mM. Cells were allowed to grow overnight at 20°C, harvested by centrifugation, and subsequently stored at −70°C. For protein purification, 4 liters of *E. coli* cells expressing DPP III-His$_6$ were grown in parallel, and the resulting bacterial paste was resuspended in 50 mM sodium phosphate (pH 8.0) containing 300 mM sodium chloride and 10 mM imidazole (buffer A) as the resuspension buffer. The cells were lysed by five cycles of sonication, 1 min each. Cell debris were removed by centrifugation at 25,000 × g for 20 min at 4°C, and the supernatant was subjected to affinity chromatography on nickel-nitrilotriacetic acid resin essentially according to protocol number 11 provided by the supplier (Qiagen). In particular, buffer A containing 20 mM imidazole was used to wash the column, and the bound protein was eluted in buffer A containing 150 mM imidazole. The progress of DPP III-His$_6$ purification was monitored by 12.5% SDS-PAGE.

All fractions of high purity and an appropriate amount of enzyme (according to SDS-PAGE) were pooled and concentrated by centrifugation with the Amicon CentriPrep system (molecular mass cutoff of 10 kDa). Further analysis of recombinant wild-type DPP III by size-exclusion chromatography indicated that the protein exists in a monomeric and dimeric form. The latter form could be resolved into monomers using dithiothreitol, suggesting formation of intermolecular disulfide bonds between monomers. Crystallization trials with wild-type DPP III were unsuccessful and prompted us to elucidate the nature of the dimerization process by generating five single cysteine-to-serine mutant proteins. This led to the identification of Cys-130 as the cause for dimerization. A biochemical characterization of these five single mutant proteins will be presented elsewhere. Here, we have continued our structural work with the C130S mutant protein, which exhibited comparable enzymatic activity to the wild-type protein.

Selenomethionine-labeled DPP III was expressed from the met$^{-}$ *E. coli* strain B834(DE3). The purification of the selenomethionine-containing DPP III followed the same protocol as for the recombinant C130S mutant protein.

**X-ray Crystallography**—

For crystallization, the C130S mutant of DPP III from *S. cerevisiae* was concentrated to 18 mg/ml (100 mM Tris-HCl (pH 7) and 100 mM NaCl). Initial conditions were identified from the Hampton Index Screen using the microbatch technique. After optimization, diffraction quality plate-like crystals (with a maximum dimension of 0.3 mm) were grown using sitting drop vapor diffusion in 20% (w/v) polyethylene glycol 3350 and 900 mM MgCl$_2$ in 100 mM Tris-HCl (pH 7.0). The crystals were monoclinic, space group *P*$_2_1$, with one DPP III molecule in the asymmetric unit. Crystals of selenomethionine-labeled protein were grown under the same conditions. In parallel, we tried to obtain “classical” heavy atom derivatives. Native PAGE shift experiments (25) indicated that mercury and osmium compounds specifically bound to DPP III. On the basis of these results, we succeeded in preparing a mercury derivative by soaking the crystals for 2 h in a solution containing mercury p-hydroxybenzoate.

All data sets were collected from flash-cooled crystals at 100 K without the use of any additional cryoprotectant. The structure was solved by a combination of MAD and MIR. Selenomethionine MAD data were collected at beamline X12 at the Deutsches Elektronen-Synchrotron/European Molecular Biology Laboratory Hamburg (λ$_{\text{peak}}$ = 0.9779 Å, λ$_{\text{nol}}$ = 0.9977 Å, λ$_{\text{ref}}$ = 0.9180 Å). The mercury derivative data set was collected at the XRD beamline at Elettra Trieste ($\lambda$ = 0.8298 Å), and the high resolution native data set was obtained at beamline 1D23 at the...
Structure of DPP III

TABLE 1

Data collection, phasing, and refinement statistics
SeMet, selenomethionine; r.m.s.d., root mean square deviation.

|                        | Native                  | SeMet                  | Hg      |
|------------------------|-------------------------|------------------------|---------|
| **Data collection**    |                         |                        |         |
| Space group            |                         |                        |         |
| Cell dimensions        |                         |                        |         |
| a, b, c (Å)            | 60.62, 110.12, 67.90    | 60.62, 110.21, 67.73   | 59.35, 107.83, 67.78 |
| α, β, γ                | 90°, 113.4°, 90°        | 90°, 113.4°, 90°       | 90°, 114.2°, 90°    |
| **Wavelength (Å)**     | 0.9537                  | 0.9779                 | 0.9787   |
| **Resolution (Å)**     | 30-1.95 (2.06–1.95)**   | 25-2.20 (2.25–2.20)**  | 0.9180   |
| R_{sym}                | 0.081 (0.414)           | 0.085 (0.324)          | 0.091 (0.437)      |
| l/αl                   | 11.5 (2.5)              | 41.5 (6.4)             | 38.8 (4.7)         |
| Completeness (%)       | 95.5 (99.5)             | 98.0 (80.1)            | 97.6 (75.7)        |
| Redundancy             | 2.9 (2.9)               | 7.0 (4.9)              | 6.9 (4.0)          |
| **Refinement**         |                         |                        |         |
| Resolution (Å)         | 30-1.95                 | 0.1923/0.2283          |         |
| No. reflections        | 57,149                  | 450                    |         |
| No. atoms              | Protein: 5703           | Ligand/ion: 3          |         |
|                        | Water: 450              | Water: 47.0            |         |
| B-factors              | Protein: 47.1           | Ligand/ion: 49.2       |         |
|                        | Water: 47.0             |                        |         |
| r.m.s.d.               | Bond length (Å) 0.015    | Bond angle 0.578°      |         |

*Values in parentheses are for the highest resolution shell.

European Synchrotron Radiation Facility in Grenoble (λ = 0.9537 Å). The data sets were processed using DENZO/SCALEPACK (26) and MOSFLM (27). Heavy atom sites (13 selenium and 1 mercury) were found independently with SOLVE/RESOLVE (28) as well as SHELXD (29). Phases (overall figure of merit of 0.65) were calculated and extended to 1.95-Å resolution using PHENIX (30). A first model was built using PHENIX and ARP/wARP (31), yielding ~600 of the 711 amino acids.

The starting model was completed and refined against the high resolution data set using PHENIX and COOT (32). R_free values (33) were computed from 5% randomly chosen reflections not used for the refinement. In the later stages of refinement, a TLS analysis was performed (34), and four TLS groups were defined (residues 2–310, 311–432, 433–672 plus a zinc ion, and 673–711) for refinement in PHENIX. The first refinement cycle including TLS led to a reduction in both R and R_free of ~2%. The final structure contains amino acid residues 2–711, one zinc ion, two magnesium ions, and 450 water molecules, resulting in R = 0.1923 and R_free = 0.2283. Three residues (Ser-37, His-578, and Ser-680) were modeled with alternative conformations. The stereochemistry of the structure was checked using PROCHECK (35), showing 99.7% of the residues in the core and allowed regions and none in the disallowed regions of the Ramachandran plot (supplemental Fig. 5). Details of the data collection, processing, and structure refinement are summarized in Table 1. The coordinates and structure factors (code 3csk) have been deposited in the Protein Data Bank.

RESULTS

Overview of the DPP III Structure—We determined the crystal structure of the C130S mutant of DPP III from S. cerevisiae at 1.95-Å resolution by a combination of selenomethionine MAD and MIR. (For a representative portion of the electron density map, see supplemental Fig. 1.) This specific variant was chosen because the mutation gave rise to homogeneous monomeric protein samples compared with the monomer/dimer mixtures usually obtained for the native enzyme.3 The mutation site turned out to be on the surface of the protein and far away (56 Å) from the metal center. Not surprisingly, this particular amino acid exchange also did not significantly alter the catalytic activity of the enzyme.

DPP III is an elongated molecule with overall dimensions of 75 × 60 × 50 Å. Two lobes can be identified that are separated by a cleft, ~40 Å wide and 25 Å high. This cleft divides the whole enzyme into two domains: an upper mostly helical domain (residues 429 – 670) and a lower domain with mixed α- and β-secondary structures (Fig. 1). The lower domain has an extension (residues 346 – 383) that forms an α-helix followed by a long loop and embraces the upper domain. A five-stranded β-structure forms the core of the lower domain (β7, β8, β9, β10, and β11) (magenta in Fig. 1; see supplemental Fig. 2 for the numbering of secondary structure elements). On one side, this structure is flanked by five helices (α1, α8, α9, α10, and α12), whereas the other side is accessible and forms part of the “floor” of the cleft opposite to the catalytic metal center (Fig. 1). The C-terminal tail of the protein (residues 671–711) folds back and interacts with the lower domain mainly by forming a short, parallel, two-stranded β-sheet with residues 9–18 and by contacts of the C-terminal α-helix. The catalytic zinc ion is bound to the upper domain at the “roof” of the cleft. An extensive search for similar folds was carried out using MSDssm (36), DALI (37), VAST (38), and DEJAVU (39) but did not yield any significant matches.
Zinc Binding—The zinc-binding site is part of the upper domain and is built up by His-460, His-465, and Glu-517, which coordinate the metal ion through the N-ε atoms of the imidazole rings and the carboxylate group, respectively (Fig. 2A). The two His ligands belong to the conserved HEXXH signature motif (24) of the M49 family of metallopeptidases (1), which is part of helix $\alpha_{13}$ (Fig. 1 and supplemental Fig. 2). The third ligand (Glu-517) is part of the second conserved sequence motif of the DPP III family (516EECRAE521) (6, 7) and is situated on helix $\alpha_{16}$. The two other glutamate residues in this motif (Glu-516 and Glu-521) are hydrogen-bonded to the zinc-coordinating histidine residues (His-465 and His-460, respectively). A water molecule completes the tetrahedral coordination of the zinc ion. This water molecule is also hydrogen-bonded to Glu-461 (Fig. 2B).

Although the DPP III structure represents a novel fold, the mode of zinc binding closely resembles that observed in other structurally unrelated metallopeptidases, such as neprilysin (40) and thermolysin (41) (supplemental Fig. 3). These two enzymes also coordinate the metal ion through two histidines and one glutamate, and a structural superposition of the zinc-binding residues reveals a very similar coordination geometry (Fig. 2C). In all three cases, the two histidine residues are part of a conserved motif, which, in the case of DPP III and its homologs, contains an additional amino acid residue (HEXXGH compared with the conventional HEXXH motif). In neprilysin and thermolysin, this metal-binding motif is part of a helical structure in which the terminal histidine residues are located on the same side of the helix. Despite the insertion of an additional residue, an equivalent coordination geometry is feasible in DPP III because of a slight widening of the last turn of the helix (Fig. 2D and supplemental Fig. 4) and a concomitant distortion of the helical hydrogen bonding pattern ($i$ to $i+5$ instead of $i$ to $i+4$). In line with this observation is experimental evidence that the deletion of one leucine residue from the HELLGH motif in rat DPP III did not completely abolish its activity (24), indicating that this truncated motif is still able to build up a functional zinc coordination site.

Active-site Cleft—The central cleft region extending on both sides of the zinc ion is the most plausible area for substrate binding. This region is quite open and appears to be easily accessible to peptide substrates. The roof part of the cleft is formed by the helical structure of the upper domain, whereas the floor is formed mainly by the central five-stranded $\beta$-sheet. Helix $\alpha_{13}$ (connecting the two domains) forms the backside of the cleft. On the left side, the cleft is confined by the C-terminal tail going from the upper to the lower domain, whereas it opens up toward the right side of the cleft, which is formed by the loop region of the arm-like extension of the lower domain (Fig. 1).

A sequence alignment of eukaryotic DPP III enzymes reveals a high sequence identity between human, rat, and mouse DPP III (>90% identity), whereas these enzymes share only ~40% sequence identity with yeast DPP III. Previously, five conserved regions were identified in such alignments (supplemental Fig.
arrangement of these important residues (Fig. 2C). This close structural resemblance with thermolysin and neprilysin and the full conservation of these residues in DPP III enzymes indicate similarly important roles in catalysis.

**Predicted Mode of Substrate Binding**—As yet, no structural information on DPP III-substrate complexes is available. We tried to identify possible modes of substrate binding based on comparisons with structures of inhibitor complexes of neprilysin and thermolysin. Specifically, we chose the complexes of thermolysin with phosphonamidate benzyloxycarbonyl-Phe-P-Leu-Ala (Protein Data Bank code 4tmn) and of neprilysin with N-α-L-rhamnopyranosyl(hydroxyphosphonyl)-L-leucyl-L-tryptophan (phosphoramidon; Protein Data Bank code 1dmt) because these inhibitors still include peptide fractions toward both the N and C termini, allowing us to predict the orientation of the substrate in DPP III. A superposition of conserved active-site residues yields modes of binding of these inhibitors to DPP III in which the C-terminal end is oriented toward the more constricted part of the cleft, whereas the N-terminal portion points toward the much wider part of the binding site (Fig. 4A). We also calculated the electrostatic potential using a continuum Poisson-Boltzmann method as implemented in DelPhi (43) and mapped it onto the molecular surface of DPP III (Fig. 3B). This analysis yielded a more positive surface potential in the narrow region of the cleft, which supports the idea that the C-terminal end of the peptide substrate is bound in this area. The positive potential is due mainly to the charge of four fully conserved Arg residues (Arg-582, Arg-664, Arg-671, and Arg-674) in this part of the cleft.

**DISCUSSION**

The structure of DPP III from *S. cerevisiae* is the first structure of a member of the M49 family of metallopeptidases (1) and represents a new fold. It is a two-lobe structure with a wide cleft running through the middle (Fig. 1), which very likely represents the substrate-binding site and also contains the essential catalytic zinc ion (4, 23).

In mammals, DPP III is considered to be associated with a wide variety of physiological processes related to the hydrolysis of bioactive peptides, such as enkephalins, endomorphins, and angiotensins. Because of the significant sequence similarity especially in the active-site region (Fig. 3A), this structure can
The conservation score was calculated for an alignment of DPP III sequences from human, mouse, rat, and yeast (supplemental Fig. 2) using AMAS (42), and dark blue areas denote complete conservation. The catalytic zinc ion is shown as a yellow sphere. The electrostatic potential was calculated using DelPhi (43) with full charges assigned to Asp, Glu, Lys, and Arg residues and partial charges to backbone atoms. The color range goes from $-20$ kT/e (red) to $+20$ kT/e (blue).

FIGURE 3. Surface representations of DPP III colored according to sequence conservation (A) and to electrostatic potential (B). The conservation score was calculated for an alignment of DPP III sequences from human, mouse, rat, and yeast (supplemental Fig. 2) using AMAS (42), and dark blue areas denote complete conservation. The catalytic zinc ion is shown as a yellow sphere. The electrostatic potential was calculated using DelPhi (43) with full charges assigned to Asp, Glu, Lys, and Arg residues and partial charges to backbone atoms. The color range goes from $-20$ kT/e (red) to $+20$ kT/e (blue).

Catalytic Mechanism—The catalytic mechanism of the bacterial zinc-dependent peptidase thermolysin has been thoroughly investigated by enzyme kinetics, mutagenesis studies, and crystallographic analyses (44) as well as by computational means (45). It is now generally accepted that a water-mediated cleavage of the scissile peptide bond is more likely than alternative mechanisms, such as the “anhydride” mechanism (41). For thermolysin, the zinc-bound water molecule is thought to be displaced by the incoming substrate carbonyl group. It stays coordinated to the metal center and is deprotonated by Glu-143 to enhance its nucleophilicity for the attack on the peptide group (41). A similar activating role has been ascribed to Glu-584 in nephrilysin (46). Subsequently, the same residue transfers a proton to the leaving nitrogen atom. In our DPP III structure, clear electron density was observed for a water molecule coordinated to the zinc ion (Fig. 2A), completing an approximate tetrahedral coordination. This water molecule is hydrogen-bonded to Glu-461, which is in an equivalent position to Glu-143 in thermolysin and Glu-584 in nephrilysin (Fig. 2C) and therefore very likely plays the same role in the catalytic mechanism. Supporting evidence comes from mutagenesis data: a mutation of Glu-461 to glutamine in yeast DPP III reduced the catalytic efficiency (because of a decrease in $k_{cat}$) by 15,000-fold, 4 Similarly, the mutation of the corresponding glutamic acid residue in rat DPP III (Glu-451) to aspartic acid or alanine completely inactivated the enzyme (24).

Stabilization of the “oxyanion” intermediate resulting from the nucleophilic attack of the water molecule on the carbonyl carbon is very likely the main factor for catalysis. Although ligation to the metal ion definitely contributes to that, in thermolysin, this reaction intermediate is additionally stabilized by hydrogen bonding interactions with neighboring residues His-231 and Tyr-157 (41, 44). In the DPP III structure, we do not observe a tyrosine residue similar to Tyr-157; however, one of the two alternative conformations of His-578 nicely superimposes with His-231 in thermolysin (and with His-711 in nephrilysin) and would be appropriately positioned for interacting with the zinc-bound substrate or intermediate (Fig. 2C). The role of the second conformation, if any, is not clear yet, but the conformational flexibility of this histidine residue could be important for catalysis. In thermolysin and nephrilysin, aspartic acid residues (Asp-226 and Asp-709, respectively) are hydrogen-bonded to those histidines indirectly influencing the stabilization of the intermediate (47). An equivalent interaction partner is missing in yeast DPP III, but the carboxamide group of Gln-576 could hydrogen bond to His-578 after a conformational change of the Gln side chain.

Although the overall structure of DPP III is completely unrelated to structures of thermolysin and nephrilysin (supplemental Fig. 3), the similarities in the active site are clearly evident. Therefore, we propose an equivalent catalytic mechanism for DPP III that involves general base activation (Glu-461) of the attacking water molecule (Fig. 4B). The differences that exist in second shell residues should not change the general mechanism but indicate variations in the stabilization of transition states and intermediates.

Substrate Recognition—A remarkable feature of the crystal structure of yeast DPP III is the extremely wide substrate-binding cleft around the catalytic zinc ion (Fig. 3). Although DPP III enzymes are known to be quite unspecific with respect to sequence and length variations of the peptide substrates (18, 48, 49), they strictly cleave off two amino acid residues from the N terminus of their substrate, i.e. they act in a dipeptidyl-peptidase manner. Thus, the question arises as to how this specificity is achieved by these enzymes.

In the absence of experimental structures of DPP III-substrate complexes, the orientation of peptide binding was predicted based on a comparison with certain inhibitor complexes of thermolysin and nephrilysin. This analysis clearly indicated that the C-terminal end of the substrate should bind in the more constricted regions of the cleft, whereas the N-terminal end points toward the more open areas (Fig. 4A). Although all three enzymes belong to completely different families of metallopeptidases (M4, M13, and M49), they share important active-site residues (Glu-461, His-578, and Arg-582 in DPP III) apart from the obvious zinc-binding residues. Despite the fundamental differences in the overall structures, both thermoly-
sin and neprilysin bind their peptide substrates in the same orientation, which increases the level of confidence for the predicted substrate orientation in DPP III.

In general, the specificity of peptidases is determined mainly by the residues within a range of four to five amino acids around the scissile bond. These residues can also influence the rate of hydrolysis. In contrast, based on the sequences of known peptide substrates of DPP III, no particular sequence pattern appears to exist. Adjacent to the scissile bond, hydrophobic, aromatic, or basic amino acid residues are accepted by human and rat DPP III. It appears that the binding of the N-terminal ammonium group of the substrate in a proper distance to the zinc ion, which itself has to bind to the carbonyl oxygen atom of the second amino acid for hydrolysis to occur, is the most important factor for the observed dipeptidyl-peptidase specificity of DPP III enzymes.

Our data thus indicate a possible way of substrate orientation and binding in which the electrostatic potential of the active-site cleft plays a major role. Although a binding site for the N terminus appears to exist (involving Glu-461 and Asp-381), a series of arginine residues (Arg anchors) enable the enzyme to accept peptide substrates of varying lengths. Although experimental evidence indicates a minimum length of three and a maximum length of 10 residues, the best substrates for DPP III appear to be tetra- to octapeptides. In the case of the tetrapeptide, our model indicates that the terminal carboxy group is bound by the guanidinium group of Arg-582. Apart from Arg-582, we identified a series of “Arg anchors” in the active-site cleft of yeast DPP III, which are suitably positioned for binding terminal carboxylates of peptides with different chain lengths. Although the C terminus of a bound pentapeptide could still interact with Arg-582, binding of hexa- and heptapeptides would involve interactions with Arg-674, and binding of longer substrates would involve interactions with Arg-671 or Arg-664. These positively charged residues are completely conserved with the only exception of Arg-671, which is a lysine in other DPP III enzymes. In addition to these arginines, the completely conserved polar residues Glu-565 and Gln-576 could play a role in binding of the main chain of the substrate.

Our data thus indicate a possible way of substrate orientation and binding in which the electrostatic potential of the active-site cleft plays a major role. Although a binding site for the N terminus appears to exist (involving Glu-461 and Asp-381), a series of arginine residues (Arg anchors) enable the enzyme to

![Catalytic mechanism of DPP III](image-url)
bind the C termini of peptides differing in chain length. In the metalloaminopeptidase from *E. coli* (51), the N terminus is bound by two glutamic residues, and the same binding motif was observed for DPP IV (52, 53), a serine hydrolase. The latter study concluded that the spatial arrangement of the DPP IV active-site serine with respect to the glutamates binding the N terminus is the most important factor for the alignment of the peptide before cleavage. In DPP IV, these two glutamic acid residues are situated on a small helix that leaves room for only two amino acid residues (52). In DPP III, the proposed binding site for the peptide N terminus is at an appropriate distance to the zinc ion, but in contrast to DPP IV, the binding site is wide open in this region. Similar to our proposal for DPP III, the C-terminal carboxylate is supposed to be bound by water-mediated interactions with two arginine residues in the *E. coli* aminopeptidase (51).

The observed tolerance of DPP III enzymes toward peptide sequence variation resembles that of neurolysin and thimet oligopeptidase, two other zinc metallopeptidases that also act on oligopeptide substrates (50), and can be explained by two different mechanisms. In the first mechanism, the different substrates adopt different conformations with their amino acid residues interacting with alternate subsites in the binding pocket. In the second mechanism, conformational changes of the enzyme itself increase the plasticity and alter the dimension of the binding site, permitting different residue types (50). In neurolysin, the broad specificity has been ascribed to flexible loops in the vicinity of the catalytic zinc ion (50). In DPP III, the peptide-binding site is formed by the cleft region that lies at the junction of the two domains, and no flexible loops can be identified close to the zinc ion. Instead, an interdomain movement could be assumed based
on high temperature factors at the domain junctions. A TLS analysis indeed indicated such a domain movement, which would lead to a widening and narrowing of the cleft dimensions as well as to a change in the shape of the cleft. Previous studies on the human enzyme indicated a hydrophobic S1′ subsite (12) similar to those found in thermolysin and neprilysin (44, 46). Based on our model structure, this (mostly hydrophobic) subsite could be formed by Ser-417, Leu-422, Phe-453, and Val-457, which are all conserved among DPP III enzymes.

Conclusion—The structure of DPP III from S. cerevisiae represents the prototype structure for the M49 family of metallopeptidases and exhibits a novel fold. It indicates a catalytic mechanism similar to that of other zinc-dependent peptidases harboring the HE motif and provides a rationale for substrate recognition. The yeast enzyme shows a high degree of similarity to other zinc-dependent metallopeptidases and exhibits a novel fold. It indicates a catalytic mechanism similar to that of other zinc-dependent peptidases.

REFERENCES

1. Rawlings, N. D., Morton, F. R., Kok, C. Y., Yong, K., and Barrett, A. J. (2008) Nucleic Acids Res. 36, D320–D325
2. Waeger, B., Nussberger, J., and Brunner, H. S. (1990) in Hypertension: Pathophysiology, Diagnosis and Management (Laragh, J. H., and Brenner, B. M., eds) pp. 2209–2232, Raven Press, New York
3. Chen, J.-M., and Barrett, A. J. (2004) in Handbook of Proteolytic Enzymes (Barrett, A. J., Rawlings, N. D., and Woessner, J. F., eds) pp. 809–812, Elsevier Science Publishers B.V., Amsterdam
4. Fukasawa, K., Fukasawa, K. M., Kanai, M., Fujii, S., Hirose, J., and Harada, M. (1998) Biochem. J. 329, 275–282
5. Mazzocco, C., Fukasawa, K. M., Auguste, P., and Puiroux, J. (2003) Eur. J. Biochem. 270, 3074–3082
6. Abramic, M., Špoljarč, J., and Šimaga, Š. (2004) Period. Biol. 106, 161–168
7. Abramic, M., Šimaga, Š., Osmak, M., Čitlarić, S., Kolaš, B., Ljubick, J., and Dolovac, L. (2004) Int. J. Biochem. Cell Biol. 36, 434–446
8. Ohkubo, I., Li, Y. H., Maeda, T., Yamamoto, Y., Yamane, T., Du, P. G., and Nishi, K. (1999) Biochem. Biophys. Res. Commun. 260, 1421–1430
9. Lee, C.-M., and Snyder, S. H. (1982) J. Biol. Chem. 257, 12043–12050
10. Vanha-Perttula, T., and Snyder, S. H. (1992) Clin. Chim. Acta 177, 199–195
11. Swanson, A., Davis, R. M., and McDonald, J. K. (1984) Curr. Eye Res. 3, 287–291
12. Abramic, M., Lukanović, M., and Vitale, L. (1988) Biochim. Biophys. Acta 939, 29–38
13. Ellis, S., and Nuenke, J. M. (1967) J. Biol. Chem. 242, 4623–4629
14. Huang, J., Kim, J., Ramamurthy, P., and Jones, T. H. D. (1992) Exp. Mycol. 16, 102–109
15. Mazzocco, C., Gillibert-Duplantier, J., Neaud, V., Fukasawa, K. M., Claverol, S., Bonneu, M., and Puiroux, J. (2006) FEBS J. 273, 1056–1064
16. Sató, H., Kimura, K., Yamamoto, Y., and Hazato, T. (2003) Masui 52, 257–263
17. Chiba, T., Li, Y. H., Yamane, T., Ojikubo, O., Fukuoka, M., Arai, R., Takahashi, S., Ohtsuka, T., Ohkubo, I., and Matsui, N. (2003) Peptides 24, 773–778
18. Baršun, M., Jajičanin, N., Vukelić, B., Špoljarč, J., and Abramic, M. (2007) Biol. Chem. 388, 343–348
19. Agić, D., Hranjec, M., Jajičanin, N., Starčević, K., Karminski-Zamola, G., and Abramic, M. (2007) Biogeochem. 35, 153–169
20. Simaga, Š., Babić, D., Osmak, M., Šprem, M., and Abramic, M. (2003) Histochem. J. 33, 511–521
21. Liu, Y., Kern, J. T., Walker, J. R., Johnson, J. A., Schultz, P. G., and Luæes, H. (2007) Proc. Natl. Acad. Sci. U. S. A. 104, 5205–5210
22. Zhang, H., Yamamoto, Y., Shumiya, S., Kitamitsu, M., Nishi, K., Ohkubo, I., and Kani, K. (2001) Biochem. J. 339, 199–208