Structure, Vol. 11, 947–959, August, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00160-6

Structural Basis of Proline-Specific Exopeptidase Activity as Observed in Human Dipeptidyl Peptidase-IV

Ralf Thoma, Bernd Löffler, Martine Stihle, Walter Huber, Armin Ruf, and Michael Hennig*
F. Hoffmann-La Roche AG
Pharma Research Discovery
4070 Basel
Switzerland

Summary

Inhibition of dipeptidyl peptidase IV (DPP-IV), the main glucagon-like peptide 1 (GLP1)-degrading enzyme, has been proposed for the treatment of type II diabetes. We expressed and purified the ectodomain of human DPP-IV in *Pichia pastoris* and determined the X-ray structure at 2.1 Å resolution. The enzyme consists of two domains, the catalytic domain, with an α/β hydrolase fold, and a β propeller domain with an 8-fold repeat of a four-strand β sheet motif. The β propeller domain contributes two important functions to the molecule that have not been reported for such structures, an extra β sheet motif that forms part of the dimerization interface and an additional short helix with a double Glu sequence motif. The Glu motif provides recognition and a binding site for the N terminus of the substrates, as revealed by the complex structure with diprotin A, a substrate with low turnover that is trapped in the tetrahedral intermediate of the reaction in the crystal.

Introduction

Dipeptidyl peptidase (DPP-IV; T cell activation antigen CD26, or adenosine binding protein) is a multifunctional type II cell surface glycoprotein. The protein is widely expressed in a variety of cell types, particularly on differential epithelial cells of the intestine, liver, prostate tissue, corpus luteum, and kidney proximal tubules (Hartel et al., 1988; McCaughan et al., 1990) as well as leukocyte subsets (Gorrell et al., 1991), such as T-helper lymphocytes, and subsets of macrophages (Bühling et al., 1994), and a soluble form is reported to be present in plasma and urine (Iwaki-Egawa et al., 1998). Human DPP-IV has a short cytoplasmatic tail of 6 amino acids, a 22-amino acid hydrophobic transmembrane region, and a 738-amino acid extracellular domain with ten potential glycosylation sites (Tanaka et al., 1992).

DPP-IV has roles in many biological processes, including a membrane-anchoring function for the localization of the extracellular enzyme adenosine deaminase (ADA) (Franco et al., 1998), participation in cell matrix adhesion by binding to collagen and fibronectin (Loster et al., 1995; Piazza et al., 1989), interaction as a coreceptor for the HIV envelope protein gp 120 (Ohtsuki et al., 2000), and a costimulatory function during T cell activation and proliferation (von Bonin et al., 1998) by interaction with the protein tyrosine phosphatase (CD45) (Torimoto et al., 1991).

DPP-IV (EC 3.4.14.5) has postproline dipeptidyl amino peptidase activity, preferentially cleaving X-proline or X-alanine dipeptides from the N terminus of polypeptides (Cunningham and O'Connor, 1997; Hopsu-Havu and Glenner, 1966). It belongs to the prolyl oligopeptidase family, a group of atypical serine proteases able to hydrolyze the prolyl bond (Cunningham and O'Connor, 1997) with the sequential order of the catalytic residues Ser-Asp-His (Ikehara et al., 1994), inverse to that found in classical serine proteases (His-Asp-Ser). The cleavage of N-terminal peptides with Pro in the second position is a rate-limiting step in the degradation of peptides. The natural substrates of DPP-IV include several chemokines, cytokines, neuropeptides, circulating hormones, and bioactive peptides (Lambeir et al., 2001). The wide range of substrates suggests a key regulatory role in the metabolism of peptide hormones and in amino acid transport (Hildebrandt et al., 2000). Its physiological relevance has been investigated by Augustyns et al. (1999), Hinke et al. (2000), Holst and Deacon (1998), and Mentlein (1999).

The finding that DPP-IV is responsible for more than 95% of the degradation of GLP-1 led to an elevated interest in inhibition of this enzyme for the treatment of type II diabetes. Experiments in rats and humans have provided evidence that specific DPP-IV inhibition increased half-life and total circulating GLP-1 and decreased plasma glucose. It has been demonstrated that patients with impaired glucose tolerance (IGT), type II diabetes, and a secondary failure to respond to sulfonylurea treatment benefit from increased levels of GLP1 peptides. In addition, GLP-1 is effective in type I diabetic patients because of its glucagonostatic effect. More recent investigations show a delay of gastric emptying that could have beneficial effects on satiety and might be relevant for the treatment of obesity. Protection of functional GLP-1 by inhibition of DPP-IV and concomitant activation of the GLP-1 receptor might therefore have a synergistic potential in antidiabetic drug research (Holst and Deacon, 1998). Selective and orally available small molecule inhibitors of DPP-IV have been discovered and are now in clinical trials (Pospisilik et al., 2002; Villhauer et al., 2002).

We expressed and purified the ectodomain of human DPP-IV in *Pichia pastoris* and determined the X-ray structure at 2.1 Å resolution in order to support and accelerate drug discovery activities by structure-based drug design and functional analysis. The enzyme consists of two domains, the catalytic domain, with an α/β hydrolase fold, and a β propeller domain with an 8-fold repeat of a β sheet motif that consists of four antiparallel strands. The β propeller domain contributes two important functions to the molecule that have not been reported for such structures, an extra β sheet motif that forms part of the dimerization interface and an additional short helix with a double Glu sequence motif. The Glu

*Correspondence: michael.hennig@roche.com

Table 1. Enzyme Kinetic Constants of DPP-IV

Proteins	k_{cat}[a] (s^{-1})	K_M[a] (μM)	k_{cat}/K_M($\mu M^{-1}s^{-1}$)
$sDPP\text{-}IV_{deglycos}$	43.1	17.2	2.51
$sDPP\text{-}IV_{glycos}$	37.3	15.5	2.41
$sDPP\text{-}IV_{deglycos}$/ADA	39.6	14.8	2.68

[a] Analyzed with Lineweaver-Burk plots; buffer, 50 mM Tris-HCl (pH 7.8), containing 100 mM NaCl, 0.1 mg/ml BSA, and 0.5% dimethylformamid; substrate, Ala-Pro-7-amido-4-trifluoromethylcoumarin; temperature, 25°C.

motif provides recognition and a binding site for the N terminus of the substrates, as revealed by the complex structure with diprotin A, a substrate with low turnover that is trapped in the tetrahedral intermediate of the reaction in the crystal.

Results and Discussion

Production and Purification of sDPP-IV

The soluble extracellular domain of human dipeptidyl peptidase IV (sDPP-IV; residues 31–766) was expressed in the yeast *Pichia pastoris*. The protein was secreted at the low level of 1 mg/l, as estimated from the total activity. As a first purification step, the concentrated protein was passed through a size exclusion column, which removed the main fraction of contaminating peptides from the yeast-peptone medium. Sequential chromatography on anion and cation exchanger columns and a second size exclusion chromatography were used to get protein of 95% purity, as judged by SDS-PAGE. The yield of pure protein was 0.3 mg/l of growth medium. The purified protein shows essentially an identical K_M value and identical inhibition constants for known inhibitors of DPP-IV, like NVP-DPP728 (Hughes et al., 1999), to the enzyme from human serum (Tables 1 and 2).

Overall Structure

The structure of human DPP-IV was solved by MAD with a mercury derivative (see Table 3) and subsequently refined to an R factor of 21.5% at 2.1 Å resolution. The current model consists of all residues from Ser39 to Pro766 of the amino acid sequence of the expressed ectodomain of the protein.

A homodimer of DPP-IV is situated in the asymmetric unit (Figure 2). Dimerization is also observed in solution under various conditions and is required for activity. Each subunit is made of two domains, the catalytic domain, with an α/β hydrolase fold containing the catalytic triad (Ser630, Asp708, and His740), and a domain with an eight-bladed β propeller fold, the β propeller domain (Figure 2). The assignment of the secondary structure is given in Figures 1 and 2. The only other known crystal structure of this class of enzyme is that of prolyl-oligopeptidase (POP), determined by Fülöp et al. (1998) (Protein Data Bank entry 1qfm). POP has an α/β-hydrolase and a β propeller domain, as well, but is monomeric, and the β propeller consists of seven repeats only (Figure 3C).

During the review process of this manuscript, a report describing the structure of human DPP-IV (Rasmussen et al., 2003) in complex with another peptidic ligand (valine-pyrrolidine dipeptide) was published. The authors produced human DPP-IV in a baculovirus expression system and used the enzyme without deglycosylation for crystallization. The resulting crystals exhibit a different crystal form, diffracting to 2.5 Å resolution. This overall structure of the enzyme is essentially identical to the one reported here.

Table 2. K_I and K_D Values of DPP-IV Inhibitors

	K_I (μM)	K_D (μM)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
Ile-Pro-Ile	4.63[b]	3.8[a]	—	—
NVP-DPP728	0.006[b]	0.002[a]	1.36×10^{6} [a]	2.48×10^{-3} [a]
$NVP\text{-}DPP728_{(Lit.)}$[c]	0.011	0.010	1.3×10^{5}	1.3×10^{-3}

[a] Measured with Biacore; buffer, 0.01 M Hepes (pH 7.4), containing 0.15 M NaCl, 3 mM EDTA, and 0.005% polysorbate 20 (v/v).
[b] Temperature, 25°C; in assay buffer (see Table 1); glycosylated sDPP-IV.
[c] (Hughes et al., 1999).

Catalytic Domain

The catalytic domain is built up of residues Gln508–Pro766 and contains a central eight-stranded β sheet that is flanked by 12 helices known as the α/β hydrolase fold. Twenty-one percent sequence identity to POP indicates significant structural homology (Figure 1), and superposition of the central α helix, carrying the catalytic Ser630 on its first turn, with the corresponding structure of POP gives an rms deviation of 2.5 Å[2] for 238 residues. The catalytic domain is connected to the β propeller by an N-terminal 15-residue linker, which is considerably shorter than the corresponding 76-residue region in POP. The residues lacking in DPP-IV are, however, replaced structurally and functionally by the C-terminal part of the catalytic domain of the second subunit of the dimer.

β Propeller Domain

The β propeller domain is formed by residues Lys56–Asn497. The preceding N-terminal residues Ser39–Leu55 form a loop structure with a small α helix (α1*; Figure 1) at the surface and in close proximity to the first residues of the catalytic domain. The β propeller domain consists of an 8-fold repeat of a four-stranded antiparallel β sheet motif (blade; Figure 3A and 3B). The blades are in circular arrangement, such that they form a solvent-filled tunnel with a diameter of about 13 Å.

The β propeller domain in DPP-IV does not form a joint β sheet motif (Fülöp and Jones, 1999; Paoli, 2001), but, rather, the blades show a regular arrangement (β1/1 to β7/4 or β8/4) (Figure 3A) around the central axis, forming a ring system that is not closed.

DPP-IV deviates from the regular β propeller fold by additional secondary-structural elements. An antiparallel β sheet is inserted in blade 2 between strands 1 and 2. The tip of the turn carries the residue, Arg125, that forms a salt bridge with Glu205, which is situated at the C-terminal turn of an α helix (residues Trp201–Glu205), which is inserted between the third and fourth blade. Arg125, Glu205, and the neighboring Glu204 form a significant part of the substrate binding site and are mainly

Table 3. Crystallographic Data and Refinement Statistics

Data Set	MAD (Remote)	MAD (Peak)	MAD (Inflection)	Apo	Diprotin A Complex
Wavelength	0.992	1.0065	1.009	0.9765	0.92
X-ray source	SLS	SLS	SLS	ID14, ESRF	SLS
Detector	MAR IP[a]	MAR IP[a]	MAR IP[a]	Quantum CCD	MAR CCD
Exposure time/frame (s)	10	10	10	2	4
Angular increment per frame (°)	2.0	2.0	2.0	0.25	0.25
Total rotation range (°)	110	136	140	130	130
Crystal to detector distance (mm)	410	410	410	240	260
Unit cell parameters (a, b, c) (Å)≅	65.2, 68.7, 420.1	65.2, 68.7, 420.1	65.2, 68.7, 420.1	65.5, 68.2, 419.3	65.1, 67.1, 419.6
Data Reduction					
Maximum Resolution (Å)	2.6	2.6	2.6	2.1	2.5
Number of measurements	212,619	263,910	276,921	234,528	171,090
Number of unique reflections	58,627	59,544	59,939	87,113	64,208
Completeness (%)[b]	97.5 (99.4)	99.9 (100.0)	99.9 (99.9)	82.9 (72.3)	97.5 (99.4)
R_{sym}[b,c]	9.1 (15.9)	9.0 (18.1)	8.6 (14.2)	8.4 (26.8)	9.1(15.9)
Heavy-Atom Refinement Parameters					
f′(e) / f″(e)	−7.0/9.5	−8.0/9.8	−12.1/5.0		
Phasing power[d] (anomalous)	0.95	1.0	0.7		
Refinement Statistics					
Resolution range (Å)				20–2.1	30–2.5
R_{cryst} (R_{free})[e] (%)				21.5 (26.5)	22.5 (28.2)
Number of protein atoms[f] (mean B in Å^2)				11 962 (34.6)	11 962 (27.1)
Number of water molecules				322 (33.4)	268 (25.0)
Number of ligand/heavy atoms (mean B in Å^2)				6 (77.3)	24 (28.3)
Number of NAG atoms (mean B in Å^2)				112 (59.0)	98 (51.4)
Rmsd[g] bonds (Å^2)				0.018	0.019
Rmsd[g] angles (°)				1.86	2.07

[a] mar research image plate detector (345 mm diameter, 100 μm pixel size).
[b] Values in parentheses are statistics for the highest resolution bin.
[c] $R_{sym} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i(h)$, where $I_i(h)$ and $\langle I(h)\rangle$ are the ith and mean measurements of the intensity of reflection h.
[d] Phasing power $= \Sigma_h F_H(h)/\Sigma_h|F_D(h) - |F_N(h) + F_H(h)||$.
[e] $\Sigma_h ||F_{obs}| - |F_{calc}||/\Sigma_h|F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes for the reflection h, applied to the working (R_{cryst}) and test (R_{free}) sets, respectively.
[f] Nonhydrogen atoms only.
[g] Rmsd, root-mean-square deviation from mean.

responsible for the substrate specificity. A further antiparallel β sheet motif formed by residues Asp230–Asn263 is inserted between strands 3 and 4 of blade 4 (Figure 3B). This structural element forms a significant part of the dimer interface (see below).

Whereas the N-terminal β sheet structure of the propeller has shorter strands and is somewhat tilted, the loop connecting the first and second β sheet is longer, shows high temperature factors, and may reduce the rigidity of the propeller architecture. The reduced stability of the circular domain structure at this position might be compensated by an extended hydrophobic cluster that consists of Ile63, Leu69, Ile76, Phe89, Leu90, Phe95, Phe98, Ile107, Ile114, Tyr135, Leu137, and Leu142, a salt bridge between Arg61 and Asp104, and a hydrogen bond between the main chain NH of Arg61 and Tyr105. This distortion leads to a reduced height of the propeller at the positions between blades 1 and 2 (Figure 3B).

As no residues from the α/β hydrolase domain fill this up, a cleft between the two domains of the DPP-IV molecule is formed, with a diameter of about 15 Å, enabling access to the catalytic site (Figure 4). Therefore, we propose that DPP-IV has two independent ways for the substrate and product to access and leave the active site, a cleft between the domains and the tunnel through the β propeller. The open cleft may enable large peptides and partially folded proteins to access the active site. The narrower tunnel could be an exit for the cleaved dipeptides (Figure 4). The crystal structure of POP shows that the cleft between the two domains does not exist and that the tunnel through the β propeller is narrower, about 4 Å compared to about 13 Å for DPP-IV (Figures 3A and 3C). This structural difference is supported by the observation that DPP-IV can process much larger substrates than can POP. Peptides with a length of up to about 80 residues appear to be good substrates of DPP-IV, and larger proteins may also be cleaved, depending on their tertiary structures (Augustyns et al., 1999; Mentlein, 1999). In contrast, POP is reported to hydrolyze substrates with a maximum size of about 30 residues only (Polgar, 1992). As the diameter of the β propeller tunnel in POP is significantly smaller, substrate access in POP requires movement of the propeller blades (Fülöp et al., 2000).

The β propeller motif has been found in several further proteins, but no, or only low, sequence homology could

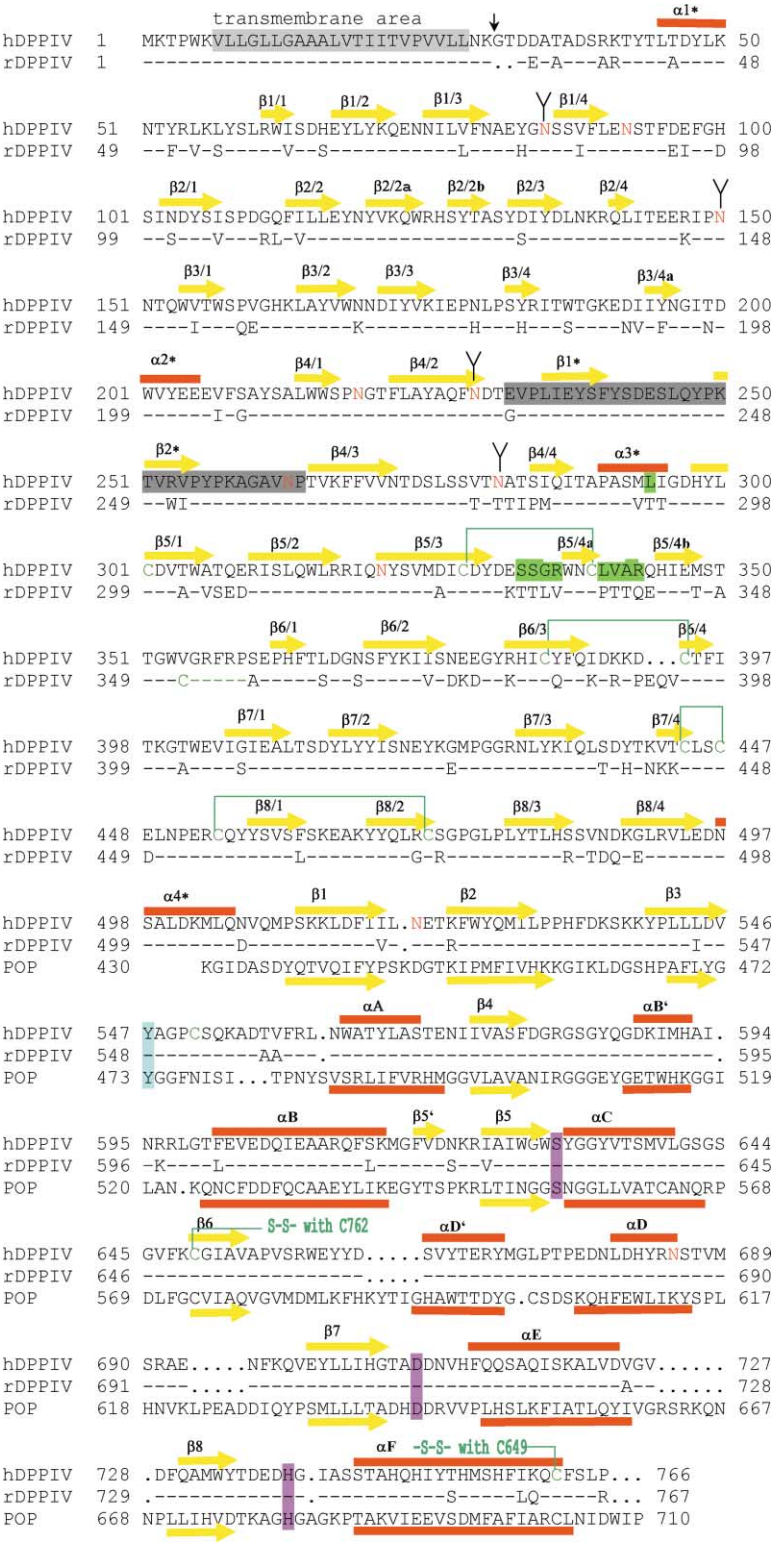

Figure 1. Sequence Alignment of DPP-IV and POP

Amino acid sequence alignment of DPP-IV from human (hDPP-IV) and rat (rDPP-IV; only different residues are shown). The alignment of POP from pork was performed with structural superposition for the α/β-hydrolase domain only because of a lack of structural homology for the β propeller domain. The top line gives additional information about the secondary structure of DPP-IV (yellow arrows and red bars), the glycosylation sites with visible electron density (Y), the potential glycosylation sites (red), the disulphide bonds (green lines between cysteines that are involved) and an arrow that indicates the start of the cloned ectodomain. Sequences are highlighted in light gray for the transmembrane part, gray for the part of the β propeller involved in dimerization, green for residues involved in adenosine deaminase binding, blue for the tyrosine that is involved in the stabilization of the oxyanion of the catalytic intermediate, and pink for the catalytic residues.

be demonstrated (Fülöp and Jones, 1999; Jawad and Paoli, 2002; Paoli, 2001). A search of the Protein Data Bank for homologous structures gave proteins with very different functions, such as clathrin (seven blades; ter Haar et al., 1998), methylamine dehydrogenase (seven blades; Chen et al., 1998) and nitrite reductase (eight blades; Fülöp et al., 1995).

Active Site

The catalytic triad (Ser630, Asp708, and His740) is located in a large cavity at the interface of the two domains. Ser630 is found at the tip of a very sharp turn between β strand 5 and helix C, called the nucleophile elbow, which is a characteristic of hydrolases of the α/β type (Ollis et al., 1992). The serine hydroxy group is

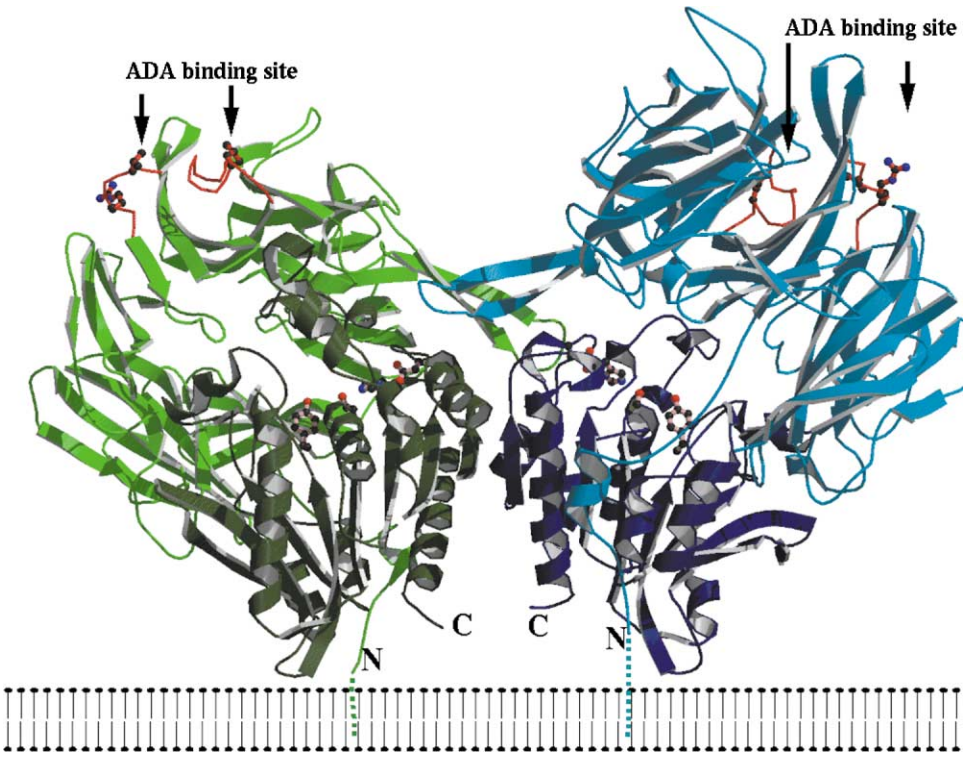

Figure 2. Overall Structure of DPP-IV

Ribbon diagram of DPP-IV viewed perpendicular to the 2-fold axis. The domains are colored dark green and light green for the α/β hydrolase and β propeller domains, respectively, of subunit A and dark/light blue for the other subunit. The overall dimensions of the molecule are about $125 \times 80 \times 60$ Å^3. The active site is highlighted by the catalytic residues in ball and stick representation as well as residues that are identified by mutagenesis data to be important for ADA binding. The proposed location at the cell surface is shown by the schematic drawing of the membrane. This figure was prepared with Molscript (Kraulis, 1991) and rendered with Raster3D (Merrit and Bacon, 1997).

well exposed to solvent and hydrogen bonded to the catalytic imidazole group of His740 on one side (2.6 Å) and accessible to the substrate on the other side. His740 is found in the middle of a loop between β strand 8 and helix F. With a distance of 2.75 Å to Nϵ of the imidazole ring, one of the oxygen atoms of Asp708 is hydrogen bonded to His740 and completes the catalytic triad (Figure 5). The other oxygen atom of the carboxylate group of Asp708 is coordinated by two main chain NH groups (Val711 and Asn710). Thus, the location and geometry of the triad are very similar to those found in other α/β hydrolases, with the "handedness" opposite to that of the classical serine peptidases (Ollis et al., 1992).

The negatively charged oxyanion of the tetrahedral intermediate is stabilized by the main chain NH group of Tyr631 and by the hydroxy group of Tyr547 (Figure 5). Furthermore, the structure shows that Gly628 and Gly632 are important for the formation of the sharp turn to bring the catalytic residue Ser630 in the correct position. This is in accordance with mutagenesis studies on rat DPP-IV (Ogata et al., 1992) showing that the sequence Gly_{628}-X-Ser_{630}-Tyr_{631}-Gly_{632} is essential for DPP-IV activity.

Substrate Binding

The substrate binding site of DPP-IV is indicated by the bound diprotin A (Ile-Pro-Ile), which is a substrate with a low turnover leading to an apparent competitive inhibition (Rahfeld et al., 1991). Inspection of the electron density map shows the ligand covalently bound to the active site Ser630 of the enzyme in both subunits. The N-terminal Ile (P2) and Pro residues (P1) are well defined and enable a detailed analysis of the interaction with the substrate binding site (according to the notation of Schechter and Berger [1968]). Less well-defined electron density is found for the C-terminal Ile (P1′), but, in subunit B, the conformation of this part of the ligand could also be observed (Figure 5). The side chain Nϵ of the catalytic His740 is in hydrogen bonding distance to the NH group of P1′ (2.90 Å) and to the Oγ of the Ser630 side chain (2.83 Å). The presence of the tripeptide in the active site indicates a very slow hydrolysis rate that might be different at the crystallization conditions compared to published data (Rahfeld et al., 1991).

DPP-IV hydrolyzes oligopeptides and proteins from the N terminus, cleaving dipeptide units when the second residue is proline, hydroxyproline, dehydroproline, pipecolic acid, or alanine. In both subunits the proline in position P1 of diprotin A is in the *trans*-configuration and fits optimally into the pocket of the active site, as expected (Fischer et al., 1983). The S1 pocket is formed by Val711, Val656, Tyr662, Tyr666, Trp659, and Tyr631, which shape a well-defined hydrophobic pocket that would be filled by proline much better than by alanine. A major contribution to binding to the pyrrolidine ring of Pro is achieved by ring stacking to Tyr662, which is equivalent to Trp595 in POP (Fülöp et al., 1998). Gly is also accepted, but with very low k_{cat}/K_M values (Brandt

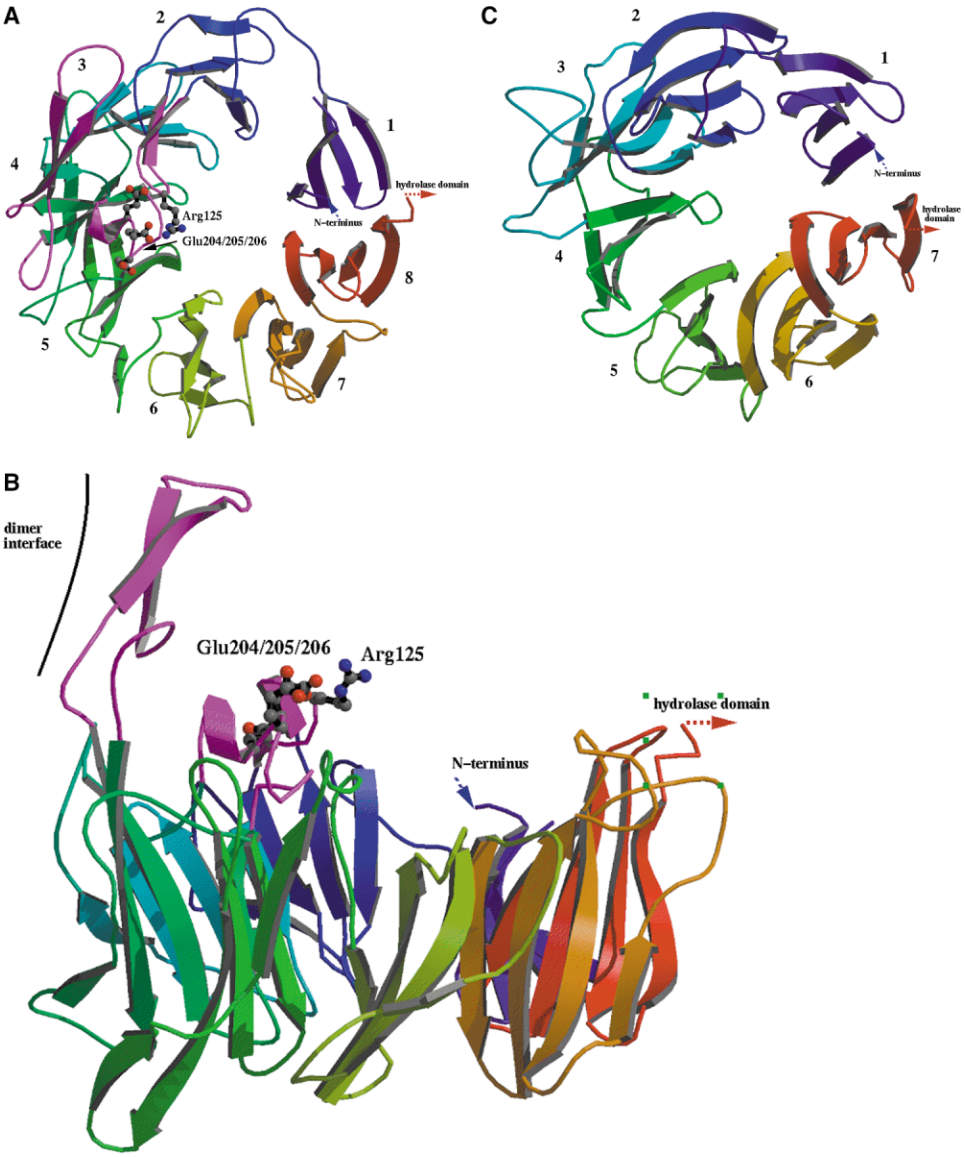

Figure 3. Ribbon Drawing of the β Propeller Domains of DPP-IV and POP

(A) DPP-IV has eight repeats of a structural motif that consists of four antiparallel β strands or blades (blades are numbered 1–8). Additional secondary structural elements are colored magenta: an antiparallel β sheet (β2/2a and β2/2b in Figure 1) that is an extension of blade 2 with Arg125 at the tip of the turn that is involved in the substrate binding, an α helix (α2* in Figure 1) with the C-terminal glutamate-rich loop that contributes to substrate recognition and specificity (Glu204/205/206), and the antiparallel β sheet that forms a main part of the dimer interface (β1* and β2* in Figure 1). The latter structural elements are extensions of blade 4.

(B) The β propeller domain of DPP-IV rotated 90°.

(C) POP has seven blades and no notable deviations from the β propeller structure. The blades are numbered 1–7.

et al., 1995). No other naturally occurring amino acids residues can occupy position P1. Either the side chains are too bulky or hydrophilic. The side chains of the residues P2 and P1′ point into the solvent, and no interaction with the protein occurs. This explains the large diversity of amino acids accepted in substrates at these positions.

Essential for substrate binding and catalysis is the N terminus of the substrates, which has to be unprotected and protonated (Brandt et al., 1996; Heins et al., 1988; Yoshimoto et al., 1978). The diprotin A complex shows that the terminal NH_3^+ group is held very precisely in position by strong interactions with the carboxylates of the double Glu motif, Glu205 and Glu206, as well as the OH of Tyr662 (Figure 5). A third glutamate, Glu204, stabilizes this substrate recognition site by a hydrogen bonding network with the backbone NH of Arg125, His126, and Ser127 as well as the hydroxy group of Ser127. The importance of the glutamate residues is confirmed by single point mutations that abolish DPP-IV activity (Abbott et al., 1999a). The double Glu motif is located at the end of a helical segment (α2*; Figure 1; see also Figure 3) that is highly conserved in the DPP IV-like gene family (Asp-Trp-X-Tyr-Glu-Glu-Glu-X). The

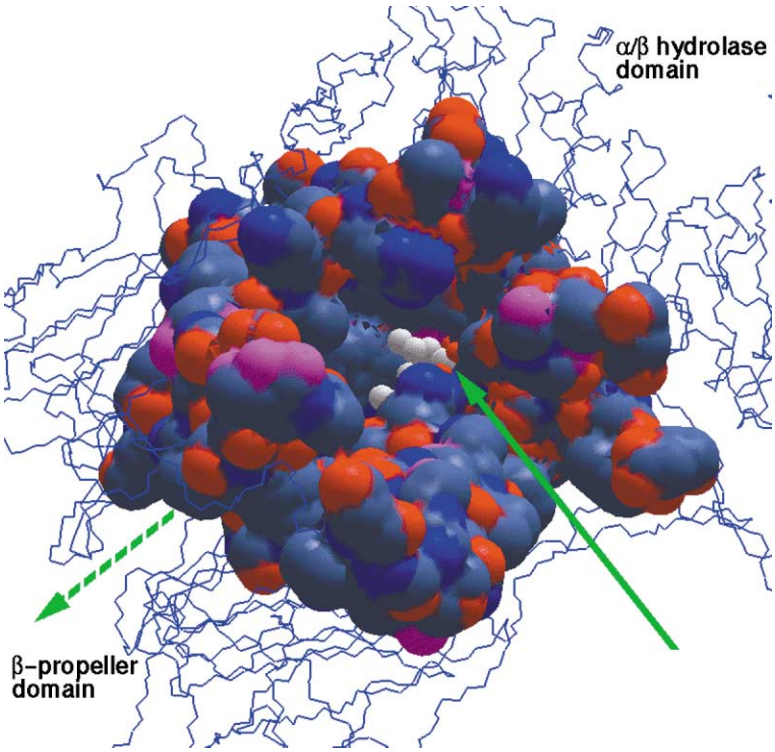

Figure 4. Access to the Active Site

Schematic view on the subunit of DPP-IV with the active site surface colored according to the atom types. The substrate diprotin A is shown with white carbons indicating the substrate binding site. Arrows illustrate that the substrate may enter the active site at the well-accessible and open active site cleft and that the dipeptidic product of the catalytic reaction may leave the active site cavity via the narrower tunnel that is formed by the β propeller.

helix represents a deviation from the regular β sheet architecture of the β propeller domain (Figures 1 and 3A). A superposition of the active sites of the complexes of the exopeptidase DPP-IV with diprotin A and the endopeptidase POP with an octapeptide (Fülöp et al., 2001) shows clear differences. The P3 and P4 residues of the octapeptide substrate of POP coincide with the double Glu motif in DPP-IV, indicating that this additional structural element in the exopeptidase is very important for substrate selectivity. Thus, the double Glu motif is a recognition site for the N terminus of substrates and restricts the cleavage to dipeptides, and the S1 pocket provides an optimal binding to proline and alanine residues, leading to a highly specific peptidase.

Mode of Inhibition by Diprotin A

Inspection of the electron density of the bound inhibitor shows a covalent linkage to Ser630 and a tetrahedral configuration for the C atom of the former carbonyl group of the scissile peptide. Obviously, a tetrahedral intermediate is observed in the complex structure with the substrate diprotin A (Figure 5), with the oxyanion stabilized by hydrogen bonds to the hydroxy group of the side chain of Tyr547 (2.70 Å) and the main chain amine of Tyr631 (3.15 Å). As much of the catalytic power of serine proteases derives from its preferential binding of this transition state, the tetrahedral intermediate is a well-defined, but normally high-energy, state with a short lifetime. The accumulation of this intermediate could be a result of a kinetic barrier, and the inspection of the active site structure reveals several structural features that could lead to a trapping of diprotin A. First, the two hydrophobic isoleucine side chains point in the same direction in proximity, and, therefore, this hydrophobic interaction may stabilize the tripeptide in an unsuitable conformation for the progress of the reaction. Second, a large network of salt bridges and hydrogen bonds stabilize the complex. It involves the carboxyl groups of Glu205/206 that interact with the N terminus of the tripeptide, but Glu205 makes another salt bridge to Arg125, and this, in turn, interacts with the C-terminal carboxyl group of the tripeptide (Figure 5). It is obvious that this interaction is only present in tripeptidic substrates and may stabilize the observed intermediate by protection of the leaving group.

We note that the pyrrolidine moieties of the structure presented here and of the complex with the dipeptide valine-pyrrolidine (Rasmussen et al., 2003) shift about 1 Å, whereas the conformations of the active side residues are very similar. The observed differences can be attributed to the covalent and noncovalent binding of the ligands to the enzyme, respectively.

Dimerization

Both the crystal structure and analytical ultracentrifugation indicate dimeric oligomerization for deglycosylated sDPP-IV, with a molecular weight of 169 kDa and noncrystallographic 2-fold symmetry (Figure 2). Six percent, or 1837 Å^2, of the total solvent-accessible surface area of each subunit is buried in the dimer interface (XSAE program, C. Broger, personal communication). This interface is mainly built up by two extra β strands (β1* and β2*) in the loop between strands 2 and 3 of the fourth blade of the β propeller domain (Figures 3A and 3B). Further interaction is provided by the α/β hydrolase domain with helix αE, β strand β8, and helix αF, with mainly hydrophobic interactions. The active site is very close to this dimer interface (Figure 2), with His740 from

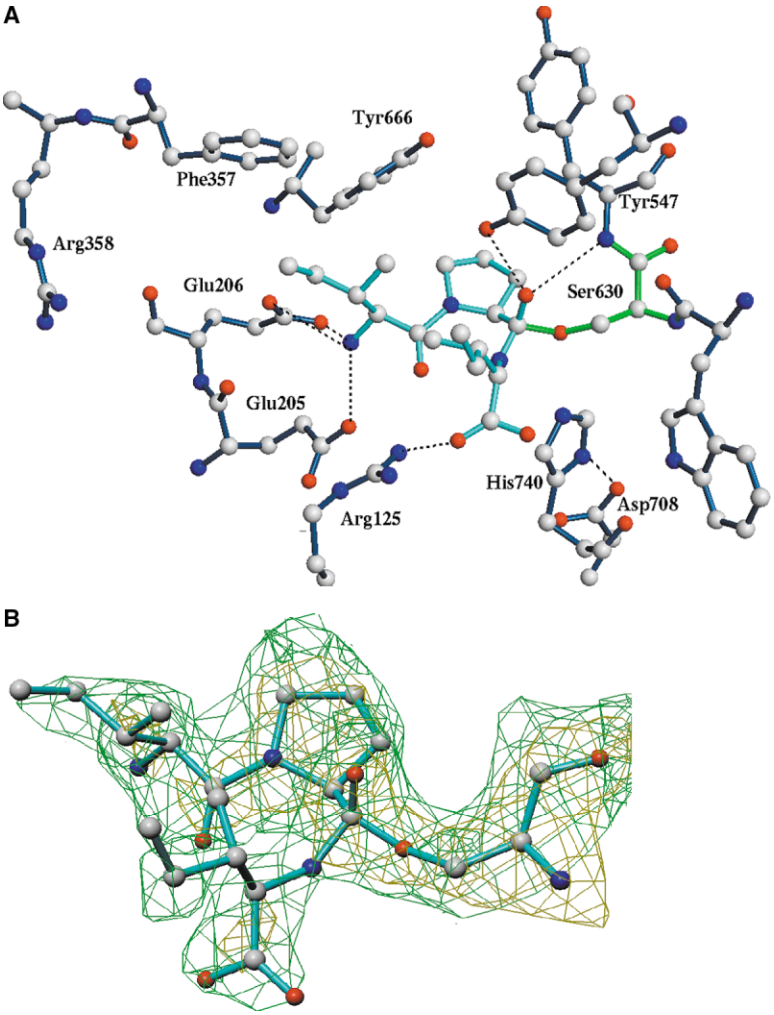

Figure 5. Active Site of DPP-IV with Diprotin A (Ile-Pro-Ile)

(A) The substrate diprotin A is trapped as a tetrahedral intermediate covalently bound to the active site Ser630. Bonds are dark blue for the protein, light blue for the ligand, and green for the active site Ser630. Dashed lines indicate hydrogen bonds. This figure was drawn with MOLOC (Gerber, 1992).

(B) The insert shows the ligand and Ser630 with the omit electron density superimposed (atoms shown were omitted from the calculations) contoured at 2.5 σ (green) and 4 σ (yellow).

the catalytic triad located in the loop connecting αF and β7 (Figure 1). Consequently, disruption of the dimer interface would also strongly affect the catalytic activity, and dimerization is required for activity.

Stability of DPP-IV

As a cell surface protein, DPP-IV is extremely stable. Consequently, the recombinant sDPP-IV shows a half-life of 5 min at 71°C in irreversible heat inactivation experiments, independent of the protein concentration and the degree of glycosylation, indicating high thermal stability. In unfolding experiments (Lambeir et al., 1997) with protein purified from human seminal plasma, DPP-IV retained its native conformation up to 8 M urea.

The crystal structure points to several factors that may contribute to this stability. First, the structural organization as a dimer with an extended hydrophobic interface stabilizes the molecule, as shown for several other proteins (Thoma et al., 2000). Second, we observed five disulphide bonds and two free sulfhydryl groups by SH titration experiments under denaturing conditions that have now been confirmed by the X-ray structure. All disulphide bridges in the β propeller connect different strands in blades or stabilize loops (Cys444/Cys447, Cys385/Cys394, Cys454/Cys472, and Cys328/Cys339). One disulfide bond is observed in the α/β-hydrolase domain (Cys649/Cys762) and covalently links the C-terminal helix αF to the core of the α/β hydrolase domain.

Glycosylation

sDPP-IV overexpressed in *P. pastoris* shows a decreasing molecular weight over the elution peak in the analytical gel filtration, as analyzed online with a multiangle laser light-scattering detector. In contrast, sDPP-IV deglycosylated with EndoF glycosidase shows a uniform molecular weight over the whole peak range because of the specific cleavage of asparagines-linked oligomannose after the first N-acetylglucosamines residue (GlcNAc). This leads to a decrease in molecular weight of 20 kDa, as estimated by SDS-PAGE. Crystals suitable for X-ray diffraction have been observed with protein that was treated with Endo F. Four GlcNAc's at the positions Asn85, Asn150, Asn229, and Asn281 could be identified in subunit A. In subunit B, the GlcNAc connected to Asn85, Asn150, and Asn229 is again visible, and an additional site could be identified at Asn92, but no electron density for an GlcNAc was found for Asn281. The GlcNAc of Asn85 is involved in a crystal contact in both subunits.

DPP-IV expressed in human has a more complex type of glycosylation than in *P. pastoris* (Cremata et al., 1998) and contains terminal sialic acid (Lambeir et al., 1997). However, no change in catalytic parameters (see Table 2) or folding is expected.

Interaction with ADA

Adenosine deaminase (ADA; EC 3.5.4.4) is a 41 kDa protein expressed in all mammalian tissues that catalyzes the deamidation of adenosine and 2′-deoxyadenosine to inosine and 2′-deoxyinosine, respectively. It is important for the regulation of the extracellular concentration of adenosine and for the regulation of the immune response. ADA is involved in T cell activation in general, and the pathogensis of autoimmune disorders (such as rheumatoid arthritis) as well as the mechanism of immunodeficiency disease (such as SCID or AID; Franco et al., 1998). Binding of the soluble extracellular ADA is a unique property of DPP-IV molecules of higher mammals and is not observed in mouse or rat DPP-IV (Dinjens et al., 1989; Iwaki-Egawa et al., 1997). Using analytical ultracentrifugation, we observed a 1:1 complex of ADA molecules with a sDPP-IV subunit, giving a molecular weight of 252 kDa. Surface plasmon resonance (Biacore) measurements show a binding constant of 3.15 ± 2 nM to ADA from bovine with a very low dissociation rate ($k_{off} = 8.75 \times 10^{-5}\,s^{-1}$; $k_{on} = 2.98 \times 10^{4}\,M^{-1}s^{-1}$), indicating a strong interaction.

Mutagenesis studies (Abbott et al., 1999b; Dong et al., 1997) identified two important regions in DPP-IV, Leu_{340}-Val_{341}-Ala_{342}-Arg_{343} (at the beginning of β5/4) and Leu294 (α4, at the end of blade 4), and a less important region, Glu_{332}-Ser_{333}-Ser_{334}-Gly_{335}-Arg_{336} (a loop region at the end of β5/3), that are all located at the surface of the β propeller domain (Figure 1). Mutation to amino acids found in rat DPP-IV reduces binding affinity to ADA. These residues form a binding site that is located far away from the active site (Figure 2), confirming the independence of DPP-IV activity on ADA binding (Table 1; De Meester et al., 1994). We conclude that the function of DPP-IV is the localization and orientation of ADA for proper catalysis. The structure gives an indication for the orientation and localization at the cell surface, because the N terminus must be close to the membrane and the ADA binding would be on the opposite site of the molecule—pointing away from the cell surface (Figure 2). Further, there would be sufficient space enabling interaction of ADA with the A1-adenosine receptor (Ciruela et al., 1996; Franco et al., 1998), which probably plays an important role in the ontogenesis of immune tissues. This view would also support the hypothesis by Ciruela and Franco (Ciruela et al., 1996; Franco et al., 1998), who proposed a link for cell-cell interaction via the binding of DPP-IV, ADA, and A1-adenosine.

Biological Implications

DPP-IV is a serine exopeptidase localized at the plasma membrane by an N-terminal transmembrane anchor. Besides its catalytic activity, it has roles in many biological processes, including binding of adenosine deaminase (Franco et al., 1998), participation in cell matrix adhesion (Loster et al., 1995; Piazza et al., 1989), interaction with HIV proteins (Ohtsuki et al., 2000), and costimulatory function during T cell activation (von Bonin et al., 1998). With its unique substrate specificity, cleaving off X-Pro or X-Ala dipeptides from the N terminus, DPP-IV plays a critical role in initiating N-terminal degradation of bioactive peptides in endocrine, neuroendocrine, and immune functions (Augustyns et al., 1999; Hinke et al., 2000; Holst and Deacon, 1998; Mentlein, 1999). Inhibition of DPP-IV leads to an increased half-life of GLP-1 (von Bonin et al., 1998), corresponding to a GLP-1 augmentation and stimulation of insulin secretion. DPP-IV has therefore been selected as a target for the treatment of type II diabetes. The structure of DPP-IV and its complex with the tripeptide Ile-Pro-Ile will provide essential information to drug discovery, as it explains how substrates are specifically recognized and gives insights into key interactions that can be used for the inhibition of the catalytic activity. Two specific negative charges from the double Glu sequence motif guide the N terminus of the peptide to the active site and fix the substrate in the correct position for cleavage. The distance between the double Glu and the catalytic serine limits the cleavage to dipeptides, and the size and shape of the P1 pocket restricts the S1 residue to proline or with less affinity to alanine—valuable information for the design of new inhibitors (Augustyns et al., 1999; Hinke et al., 2000; Holst and Deacon, 1998; Mentlein, 1999). The apparent competitive inhibition caused by a low turnover rate of the substrate Ile-Pro-Ile can be explained by the trapping of the tetrahedral intermediate by several interactions that cannot be expected for normal substrates. In particular, the hydrophobic interactions made by the Ile side chains of the ligand and participation of the negatively charged C terminus of the peptide in an extensive salt bridge cluster account for this kinetic barrier.

Besides the impact for drug design, the structure gives an understanding of the importance of dimerization for DPP-IV activity, the regulation of substrate access via two active site entrances, and the interaction with other proteins, such as adenosine deaminase.

Experimental Procedures

DNA Manipulation and Sequence Analysis

Preparation of DNA probes, digestion with restriction endonucleases, DNA ligation, and transformation of *E.coli* strains were performed as described (Sambrook et al., 1989). For DNA sequencing, the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit and ABI PRISM 310 Genetic Analyzer were used. PCR was performed in the T3 Thermocycler (Whatman Biometra), with *Pfu* polymerase (Stratagene).

Expression of Recombinant Human sDPP-IV in *P. pastoris*

The ectodomain of DPP-IV, residues 31–766 (sDPP-IV), was amplified by PCR with the plasmid pcHDP1-23 as a template (Misumi et al., 1992). The oligonucleotides used for PCR were 5′-TGCTG**GAAT TC**GGCACAGATGATGCTAC-3′ with an EcoRI site (bold letters) and 5′-GCAT**GGTACC**TT GAGGTGCTAAG -3′ with a KpnI site (bold letters). With the two new restriction sites, the amplified DNA fragment was cloned into pPICZα-A vector (Invitrogen) to create a fusion with the α-mating factor signal sequence for the secretion of the protein. Truncation of the α-mating factor by the signal peptidases will leave two additional amino acids, glutamine and phenylalanine, at the N terminus of sDPP-IV. The sequence was confirmed by sequencing. pPICZα-sDPP-IV was linearized with SacI and transformed by elec-

troporation in *P. pastoris* strain GS115, and the phenotype of the colonies obtained was checked as recommended by the distributor, Invitrogen.

Eight transformants with the phenotype MutS were screened for the expression of DPP-IV. Colonies were grown at 30°C in YPD medium (1% yeast extract, 2% peptone, and 2% glucose) with zeocin (100 μg/ml) to an OD_{600} of 8–10. Cells were collected by centrifugation and resuspended in YP medium plus 2% methanol. The same amount of methanol was added every 24 hr. After 48 hr the medium of each clone was tested for activity (see below). sDPP-IV was then produced in a large-scale culture with the transformed cell line with the highest activity per volume, as described (Dale et al., 2000).

Purification of sDPP-IV

Ten liters of the collected sDPP-IV supernatant of the selected transformed *P. pastoris* cell line was filtered and concentrated to 180 ml by crossflow ultrafiltration (skannette) with a 30 kDaA filtration module (AGT Technology). The concentrate was passed over a Sephacryl 200 XK 50/100 size exclusion column (5 × 95 cm; Pharmacia) and equilibrated with 50 mM Tris-HCl (pH 7.8) and 100 mM NaCl (S buffer). Collected fractions were screened on SDS-PAGE and for activity. Fractions containing sDPP-IV were dialyzed against 50 mM Tris-HCl (pH 7.9). The protein solution was loaded on a Fractogel-TMAE column (2.6 × 13 cm; Merck), equilibrated with 50 mM Tris-HCl (pH 7.9), washed with two column volumes of the same buffer, and eluted with 500 ml of a linear gradient from 0–200 mM NaCl. Fractions containing sDPP-IV were dialyzed against 20 mM sodium acetate (pH 4.8). The protein solution was loaded on a Fractogel-COO^- column (1 × 12 cm; Merck), equilibrated with the same buffer, and washed with two column volumes of this buffer. Bound proteins were eluted with 200 ml of a linear gradient from 50 to 500 M NaCl. The elution profile showed a major peak at 250 mM NaCl. Preparation of enzymatically deglycosylated sDPP-IV ($sDPPIV_{deglycos}$) was carried out prior to loading on the last gel filtration column. EndoF1-GST (0.1%) was added to the pooled fractions of DPP-IV and incubated for 20 hr at 21°C. The concentrated protein solution was loaded on a Biosec size exclusion column (1.6 × 60 cm; Merck) that was equilibrated with S buffer. Fractions were analyzed by SDS-PAGE, showing a purity of >95%. N-terminal sequencing showed that the protein was efficiently processed by the STE13 signal peptidase, which cleaves off the α-mating factor. Preparation of the $sDPPIV_{deglycos}$:ADA complex was performed by addition of a two times excess of ADA (Sigma Type IV, from calf intestinal Mucosa) and purification with a Biosec size exclusion column.

Analytical Methods

Purification of sDPP-IV was followed by electrophoresis on 10%–20% Tricine SDS polyacrylamide gradient gels (Lämmli, 1970). Protein concentrations were determined according to Bradford (Bradford, 1976) or for pure protein by absorption spectroscopy (Pace et al., 1995) with the calculated molecular extinction coefficient at 280 nm of 193′920 $M^{-1}cm^{-1}$ (A_{280} = 2.27 for a 1 cm path and a protein concentration of 1mg/ml). Analytical gel filtration chromatography was performed on a Superdex 200 12 HR 10/30 column (Pharmacia) equilibrated with S buffer. The eluate was monitored with a miniDAWN multiangle laser light-scattering detector (Wyatt) and a refractive index detector (Shodex), which allows the determination of the molecular weight and dispersity over the elution peak (Wyatt, 1993). Sedimentation equilibrium runs in a Beckman analytical ultracentrifuge (model Optima XL A) were performed at 20°C and 9000 rpm for $sDPP\text{-}IV_{deglycos}$ and at 7000 rpm for the $sDPP\text{-}IV_{deglycos}$:ADA-complex. The initial protein concentrations were 0.22–0.25 mg/ml in S buffer. The absorption was followed at 280 nm. Assumed partial specific volumes for sDPP-IV of 0.729 cm^3/g and ADA of 0.735 cm^3/g were used to determine the molecular masses.

Free sulfhydryl groups were determined according the procedure described by Ellman (Ellman, 1959) under denaturing conditions (0.3% SDS in 50 mM Tris [pH 8.0]).

Thermostability Measurements

The irreversible loss of activity after incubation at various temperatures was used as an operational criterion of the thermostability of sDPP-IV. Kinetics of irreversible heat inactivation were performed as described by Sterner et al. (1996) with a final protein concentration of 20 μg/ml in 50 mM potassium phosphate buffer at pH 7.5, containing 100 mM NaCl. The residual activity was determined by recording the initial velocity at 25°C of the enzyme-catalyzed reaction (see below), and the averaged values obtained were plotted against the incubation temperature.

Biacore

DPP-IV was immobilized on a CM5 surface plasmon resonance sensor (Biacore) by standard amide coupling chemistry. The organic adlayer on this sensor type consists of carboxymethylated dextran (MW of ~100 kDa). After activation of the carboxylic acid groups with carbodiimide/N-hydroxysuccinimide solutions, the surface was contacted with a DPP-IV solution (80 μl) containing about 100 μg/ml protein in acetate buffer (10 mM [pH 4.5]). The amount immobilized corresponded to a sensor response of roughly 10,000 RU. The surfaces of two flow cells were modified with protein. To suppress baseline drift—possibly due to slow dimer dissociation—the protein of one cell was crosslinked by short contact with carbodiimide/N-hydroxysuccinimide solution. This treatment did not influence the protein activity, since binding constants determined with crosslinked protein were similar to those determined with noncrosslinked protein. HEPES buffer (0.01 M HEPES [pH 7.4], 0.15 M NaCl, 3 mM EDTA, and 0.005% polysorbate 20 [v/v]) was used as the running buffer. Diprotin A was disolved directly in this buffer. NVP-DPP728 was first dissolved in pure DMSO and then diluted into running buffer. The final inhibitor solution contained less than 0.1% DMSO. Binding experiments were carried out by contacting the immobilized protein surfaces with inhibitor solutions of varying concentrations at a flow rate of 10 μl/min or 30 μl/min. After each contact with inhibitor, the protein surfaces were regenerated by extensively washing with running buffer.

Activity Assay

The activity assay is based on the increase of fluorescence of products compared to the substrate Ala-Pro-7-amido-4-trifluoromethylcoumarin (Calbiochem; Smith et al., 1992). A 20 mM stock solution in 10% DMF is stored at −20°C until use. Purification was followed with a final substrate concentration of 50 μM, and, for the determination of kinetic parameters, it was varied between 1.5 μM and 500 μM in the assay. DPP-IV activity assays were performed in 96-well plates in a total assay volume of 100 μl. The assay buffer consists of S buffer containing 0.1 mg/ml BSA. Fluorescence is detected in an LS 50B luminescence spectrometer LS 50B (Perkin Elmer) at an excitation wavelength of 400 nm and an emission wavelength of 505 nm. Initial rate constants are calculated by best-fit linear regression.

Crystallization

For crystallization trials, $sDPP\text{-}IV_{deglycos}$ was concentrated to approximately 10 mg/ml. A reduced factorial screen was carried out by the vapor diffusion method. Crystals were obtained with 20%–25% PEG 3350, 200 mM $MgCl_2$, Tris (pH 8.5), and 15% glycerol. The crystals were flash-frozen in liquid nitrogen and exhibit the orthorhombic space group $P2_12_12_1$, with cell dimensions of approximately 65 Å, 68 Å, and 420 Å and one dimer per asymmetric unit. They diffract to a maximum of 2.1 Å resolution with synchrotron radiation and show rather high mosaicity (0.5°–1.2°). Addition of 1 mM diprotin A prior to crystallization led to crystals of the complex. The mercury derivative was produced by cocrystallization with 0.1 mM $HgCl_2$.

Structure Determination

Data collection was performed with synchrotron radiation (Swiss Light Source [SLS], Villigen, Switzerland, and ID14, ESRF, Grenoble, France) as well as in-house facilities (search for heavy-atom derivatives and evaluation of crystal quality) and processed with DENZO (Otwinowski, 1993). Details of the data collection statistics are given in Table 3. All programs used are part of the CCP4 (CCP4, 1994) suite, except where indicated. The structure was determined by MAD with a mercury derivative. One major mercury binding site per subunit (Cys 551, one of the two free SH-groups, Cys301 and Cys551, that are located near the active site) was identified by inspection of the difference Patterson maps calculated from the peak wavelength data and was subsequently refined with SHARP

(de la Fortelle and Bricogne, 1997). Location of the 2-fold noncrystallographic axis was performed with this mercury site and the program find2folds (Dunten and Hennig, 2002). Further analysis revealed another site per subunit (Cys301), with less occupancy and disorder (two main positions with about 2.4 Å distance). Subsequently, the phases were improved by application of 2-fold averaging combined with solvent flattening and histogram matching as implemented in DM. The initial electron density at 2.6 Å resolution was readily interpretable, and about 90% of the polypeptide chain could be built. The molecular model was refined against 2.1 Å data. Subsequent rounds of manual rebuilding and refinement with REFMAC (Murshudov et al., 1999) led to a complete molecular structure of the polypeptide chain from residues Ser39 to Pro766. Details of the refined structures are reported in Table 3.

Acknowledgments

We would like to thank the staff of the synchrotron facilities at the ID14 in Grenoble, France; C. Schulze-Briese and T. Tomizaki at the SLS in Villigen, Switzerland, for their assistence; Yoshio Misumi for the gift of DPP-IV cDNA; A. D'Arcy for help in the initial crystallization experiments; F. Müller and E. Kusznir for performing the analytical ultracentrifugation experiments; A. Alker for technical assistance of the in-house X-ray facilities; C. Vonrhein for discussions regarding SHARP; P. Dunten for help in identification of the NCS symmetry; B. Kuhn for support in computational chemistry; and M. Böhringer and D.W. Banner for discussion and critical reading of the manuscript.

Received: November 6, 2002
Revised: March 6, 2003
Accepted: May 20, 2003
Published: August 5, 2003

References

Abbott, C.A., McCaughan, G.W., and Gorrell, M.D. (1999a). Two highly conserved glutamic acid residues in the predicted beta propeller domain of dipeptidyl peptidase IV are required for its enzyme activity. FEBS Lett. *458*, 278–284.

Abbott, C.A., McCaughan, G.W., Levy, M.T., Church, W.B., and Gorrell, M.D. (1999b). Binding to human dipeptidyl peptidase IV by adenosine deaminase and antibodies that inhibit ligand binding involves overlapping, discontinuous sites on a predicted beta propeller domain. Eur. J. Biochem. *266*, 798–810.

Augustyns, K., Bal, G., Thonus, G., Belyaev, A., Zhang, X.M., Bollaert, W., Lambeir, A.M., Durinx, C., Goossens, F., and Haemers, A. (1999). The unique properties of dipeptidyl-peptidase IV (DPP IV / CD26) and the therapeutic potential of DPP IV inhibitors. Curr. Med. Chem. *6*, 311–327.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Brandt, W., Lehmann, T., Thondorf, I., Born, I., Schutkowski, M., Rahfeld, J.U., Neubert, K., and Barth, A. (1995). A model of the active site of dipeptidyl peptidase IV predicted by comparative molecular field analysis and molecular modelling simulations. Int. J. Pept. Protein Res. *46*, 494–507.

Brandt, W., Ludwig, O., Thondorf, I., and Barth, A. (1996). A new mechanism in serine proteases catalysis exhibited by dipeptidyl peptidase IV (DP IV)—results of PM3 semiempirical thermodynamic studies supported by experimental results. Eur. J. Biochem. *236*, 109–114.

Bühling, F., Kunz, D., Reinhold, D., Ulmer, A.J., Ernst, M., Flad, H.D., and Ansorge, S. (1994). Expression and functional role of dipeptidyl peptidase IV (CD26) on human natural killer cells. Nat. Immun. *13*, 270–279.

CCP4 (Collaborative Computational Project, Number 4). (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Chen, L., Doi, M., Durley, R.C., Chistoserdov, A.Y., Lidstrom, M.E., Davidson, V.L., and Mathews, F.S. (1998). Refined crystal structure of methylamine dehydrogenase from Paracoccus denitrificans at 1.75 Å resolution. J. Mol. Biol. *276*, 131–149.

Ciruela, F., Saura, C., Canela, E.I., Mallol, J., Lluis, C., and Franco, R. (1996). Adenosine deaminase affects ligand-induced signaling by interacting with cell surface adenosine receptors. FEBS Lett. *380*, 219–223.

Cremata, J., Montensino, R., Quintero, O., and Garcia, R. (1998). Glycosylation profiling of heterologous proteins. In Pichia Protocols, D.R. Higgins and J.M. Cregg, eds. (Totowa, New Jersey: Humana Press), pp. 95–106.

Cunningham, D.F., and O'Connor, B. (1997). Proline specific peptidases. Biochim. Biophys. Acta *1343*, 160–186.

Dale, G.E., D'Arcy, B., Yuvaniyama, C., Wipf, B., Oefner, C., and D'Arcy, A. (2000). Purification and crystallization of the extracellular domain of human neutral endopeptidase (neprilysin) expressed in Pichia pastoris. Acta Crystallogr. D Biol. Crystallogr. *56*, 894–897.

de la Fortelle, E., and Bricogne, G. (1997). Maximum likelihood heavy-atom parameter refinement for multiple isomorphus replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

De Meester, I., Vanham, G., Kestens, L., Vanhoof, G., Bosmans, E., Gigase, P., and Scharpe, S. (1994). Binding of adenosine deaminase to the lymphocyte surface via CD26. Eur. J. Immunol. *24*, 566–570.

Dinjens, W.N., ten Kate, J., Wijnen, J.T., van der Linden, E.P., Beek, C.J., Lenders, M.H., Khan, P.M., and Bosman, F.T. (1989). Distribution of adenosine deaminase-complexing protein in murine tissues. J. Biol. Chem. *264*, 19215–19220.

Dong, R.P., Tachibana, K., Hegen, M., Munakata, Y., Cho, D., Schlossman, S.F., and Morimoto, C. (1997). Determination of adenosine deaminase binding domain on CD26 and its immunoregulatory effect on T cell activation. J. Immunol. *159*, 6070–6076.

Dunten, P., and Hennig, M. (2002). Locating non-crystallographic symmetry elements: the program Find2Folds. Acta Crystallogr. A *58*, C76.

Ellman, G.L. (1959). Tissue sulfhydryl groups. Arch. Biochem. Biophys. *82*, 70–77.

Fischer, G., Heins, J., and Barth, A. (1983). The conformation around the peptide bond between the P1- and P2-positions is important for catalytic activity of some proline-specific proteases. Biochim. Biophys. Acta *742*, 452–462.

Franco, R., Valenzuela, A., Lluis, C., and Blanco, J. (1998). Enzymatic and extraenzymatic role of ecto-adenosine deaminase in lymphocytes. Immunol. Rev. *161*, 27–42.

Fülöp, V., and Jones, D.T. (1999). Beta propellers: structural rigidity and functional diversity. Curr. Opin. Struct. Biol. *9*, 715–721.

Fülöp, V., Moir, J.W., Ferguson, S.J., and Hajdu, J. (1995). The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd1. Cell *81*, 369–377.

Fülöp, V., Bocskei, Z., and Polgar, L. (1998). Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. Cell *94*, 161–170.

Fülöp, V., Szeltner, Z., and Polgar, L. (2000). Catalysis of serine oligopeptidases is controlled by a gating filter mechanism. EMBO Rep. *1*, 277–281.

Fülöp, V., Szeltner, Z., Renner, V., and Polgar, L. (2001). Structures of prolyl oligopeptidase substrate/inhibitor complexes. Use of inhibitor binding for titration of the catalytic histidine residue. J. Biol. Chem. *276*, 1262–1266.

Gerber, P.R. (1992). Peptide mechanics: a force field for peptides and proteins working with entire residues as small unites. Biopolymers *32*, 1003–1017.

Gorrell, M.D., Wickson, J., and McCaughan, G.W. (1991). Expression of the rat CD26 antigen (dipeptidyl peptidase IV) on subpopulations of rat lymphocytes. Cell. Immunol. *134*, 205–215.

Hartel, S., Gossrau, R., Hanski, C., and Reutter, W. (1988). Dipeptidyl

peptidase (DPP) IV in rat organs. Comparison of immunohistochemistry and activity histochemistry. Histochemistry *89*, 151–161.

Heins, J., Welker, P., Schonlein, C., Born, I., Hartrodt, B., Neubert, K., Tsuru, D., and Barth, A. (1988). Mechanism of proline-specific proteinases: (I) substrate specificity of dipeptidyl peptidase IV from pig kidney and proline-specific endopeptidase from Flavobacterium meningosepticum. Biochim. Biophys. Acta *954*, 161–169.

Hildebrandt, M., Reutter, W., Arck, P., Rose, M., and Klapp, B.F. (2000). A guardian angel: the involvement of dipeptidyl peptidase IV in psychoneuroendocrine function, nutrition and immune defence. Clin. Sci. *99*, 93–104.

Hinke, S.A., Pospisilik, J.A., Demuth, H.U., Mannhart, S., Kuhn-Wache, K., Hoffmann, T., Nishimura, E., Pederson, R.A., and McIntosh, C.H. (2000). Dipeptidyl peptidase IV (DPIV/CD26) degradation of glucagon. Characterization of glucagon degradation products and DPIV-resistant analogs. J. Biol. Chem. *275*, 3827–3834.

Holst, J.J., and Deacon, C.F. (1998). Inhibition of the activity of dipeptidyl-peptidase IV as a treatment for type 2 diabetes. Diabetes *47*, 1663–1670.

Hopsu-Havu, V.K., and Glenner, G.G. (1966). A new dipeptide naphthylamidase hydrolyzing glycyl-prolyl-beta-naphthylamide. Histochemie *7*, 197–201.

Hughes, T.E., Mone, M.D., Russell, M.E., Weldon, S.C., and Villhauer, E.B. (1999). NVP-DPP728 (1-[[[2-[(5-cyanopyridin-2-yl)amino]ethyl]amino]acetyl]-2-cyano-(S)-pyrrolidine), a slow-binding inhibitor of dipeptidyl peptidase IV. Biochemistry *38*, 11597–11603.

Ikehara, Y., Ogata, S., and Misumi, Y. (1994). Dipeptidyl-peptidase IV from rat liver. Methods Enzymol. *244*, 215–227.

Iwaki-Egawa, S., Watanabe, Y., and Fujimoto, Y. (1997). CD26/dipeptidyl peptidase IV does not work as an adenosine deaminase-binding protein in rat cells. Cell. Immunol. *178*, 180–186.

Iwaki-Egawa, S., Watanabe, Y., Kikuya, Y., and Fujimoto, Y. (1998). Dipeptidyl peptidase IV from human serum: purification, characterization, and N-terminal amino acid sequence. J. Biochem. *124*, 428–433.

Jawad, Z., and Paoli, M. (2002). Novel sequences propel familiar folds. Structure *10*, 447–454.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Lambeir, A.M., Diaz Pereira, J.F., Chacon, P., Vermeulen, G., Heremans, K., Devreese, B., Van Beeumen, J., De Meester, I., and Scharpe, S. (1997). A prediction of DPP IV/CD26 domain structure from a physico-chemical investigation of dipeptidyl peptidase IV (CD26) from human seminal plasma. Biochim. Biophys. Acta *1340*, 215–226.

Lambeir, A.M., Durinx, C., Proost, P., Van Damme, J., Scharpe, S., and De Meester, I. (2001). Kinetic study of the processing by dipeptidyl-peptidase IV/CD26 of neuropeptides involved in pancreatic insulin secretion. FEBS Lett. *507*, 327–330.

Lämmli, U.K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Loster, K., Zeilinger, K., Schuppan, D., and Reutter, W. (1995). The cysteine-rich region of dipeptidyl peptidase IV (CD 26) is the collagen-binding site. Biochem. Biophys. Res. Commun. *217*, 341–348.

McCaughan, G.W., Wickson, J.E., Creswick, P.F., and Gorrell, M.D. (1990). Identification of the bile canalicular cell surface molecule GP110 as the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue distribution, purification and N-terminal amino acid sequence. Hepatology *11*, 534–544.

Mentlein, R. (1999). Dipeptidyl-peptidase IV (CD26)–role in the inactivation of regulatory peptides. Regul. Pept. *85*, 9–24.

Merrit, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Misumi, Y., Hayashi, Y., Arakawa, F., and Ikehara, Y. (1992). Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine proteinase on the cell surface. Biochim. Biophys. Acta *1131*, 333–336.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr. D Biol. Crystallogr. *55*, 247–255.

Ogata, S., Misumi, Y., Tsuji, E., Takami, N., Oda, K., and Ikehara, Y. (1992). Identification of the active site residues in dipeptidyl peptidase IV by affinity labeling and site-directed mutagenesis. Biochemistry *31*, 2582–2587.

Ohtsuki, T., Tsuda, H., and Morimoto, C. (2000). Good or evil: CD26 and HIV infection. J. Dermatol. Sci. *22*, 152–160.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., et al. (1992). The alpha/beta hydrolase fold. Protein Eng. *5*, 197–211.

Otwinowski, Z. (1993). Oscillation data reduction program. In Proceedings of the CCP4 Study Weekend: Data Collection and Processing, L. Wawyey, N. Isaacs, and S. Bailey, eds. (Daresbury, UK: SERC Daresbury Laboratory), pp. 56–62.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci. *4*, 2411–2423.

Paoli, M. (2001). Protein folds propelled by diversity. Prog. Biophys. Mol. Biol. *76*, 103–130.

Piazza, G.A., Callanan, H.M., Mowery, J., and Hixson, D.C. (1989). Evidence for a role of dipeptidyl peptidase IV in fibronectin-mediated interactions of hepatocytes with extracellular matrix. Biochem. J. *262*, 327–334.

Polgar, L. (1992). Unusual secondary specificity of prolyl oligopeptidase and the different reactivities of its two forms toward charged substrates. Biochemistry *31*, 7729–7735.

Pospisilik, J.A., Stafford, S.G., Demuth, H.U., McIntosh, C.H., and Pederson, R.A. (2002). Long-term treatment With dipeptidyl peptidase IV inhibitor improves hepatic and peripheral insulin sensitivity in the VDF Zucker rat: A euglycemic-hyperinsulinemic clamp study. Diabetes *51*, 2677–2683.

Rahfeld, J., Schierhorn, M., Hartrodt, B., Neubert, K., and Heins, J. (1991). Are diprotin A (Ile-Pro-Ile) and diprotin B (Val-Pro-Leu) inhibitors or substrates of dipeptidyl peptidase IV? Biochim. Biophys. Acta *1076*, 314–316.

Rasmussen, H.B., Branner, S., Wiberg, F.C., and Wagtmann, N. (2003). Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. Nat. Struct. Biol. *10*, 19–25.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Schechter, I., and Berger, A. (1968). On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. Biochem. Biophys. Res. Commun. *32*, 898–902.

Smith, R.E., Reynolds, C.J., and Elder, E.A. (1992). The evolution of proteinase substrates with special reference to dipeptidylpeptidase IV. Histochem. J. *24*, 637–647.

Sterner, R., Kleemann, G.R., Szadkowski, H., Lustig, A., Hennig, M., and Kirschner, K. (1996). Phosphoribosyl anthranilate isomerase from Thermotoga maritima is an extremely stable and active homodimer. Protein Sci. *5*, 2000–2008.

Tanaka, T., Camerini, D., Seed, B., Torimoto, Y., Dang, N.H., Kameoka, J., Dahlberg, H.N., Schlossman, S.F., and Morimoto, C. (1992). Cloning and functional expression of the T cell activation antigen CD26. J. Immunol. *149*, 481–486.

ter Haar, E., Musacchio, A., Harrison, S.C., and Kirchhausen, T. (1998). Atomic structure of clathrin: a beta propeller terminal domain joins an alpha zigzag linker. Cell *95*, 563–573.

Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. (2000). Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from Thermotoga maritime. Structure *8*, 265–276.

Torimoto, Y., Dang, N.H., Vivier, E., Tanaka, T., Schlossman, S.F., and Morimoto, C. (1991). Coassociation of CD26 (dipeptidyl peptidase IV) with CD45 on the surface of human T lymphocytes. J. Immunol. *147*, 2514–2517.

Villhauer, E.B., Brinkman, J.A., Naderi, G.B., Dunning, B.E., Mangold, B.L., Mone, M.D., Russell, M.E., Weldon, S.C., and Hughes, T.E. (2002). 1-[2-[(5-Cyanopyridin-2-yl)amino]ethylamino]acetyl-2-(S)-pyrrolidinecarbon nitrile: a potent, selective, and orally bioavailable dipeptidyl peptidase IV inhibitor with antihyperglycemic properties. J. Med. Chem. *45*, 2362–2365.

von Bonin, A., Huhn, J., and Fleischer, B. (1998). Dipeptidyl-peptidase IV/CD26 on T cells: analysis of an alternative T-cell activation pathway. Immunol. Rev. *161*, 43–53.

Wyatt, P.J. (1993). Light scattering and the absolute characterisation of macromolecules. Anal. Chim. Acta *272*, 1–40.

Yoshimoto, T., Fischl, M., Orlowski, R.C., and Walter, R. (1978). Post-proline cleaving enzyme and post-proline dipeptidyl aminopeptidase. Comparison of two peptidases with high specificity for proline residues. J. Biol. Chem. *253*, 3708–3716.

Accession Numbers

The coordinates have been deposited in the Protein Data Bank under accession codes 1NU6 and 1NU8.