Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog

Hanne B. Rasmussen[1], Sven Branner[1], Finn C. Wiberg[2,3] and Nicolai Wagtmann[2]

Published online 16 December 2002; doi:10.1038/nsb882

Dipeptidyl peptidase IV (DPP-IV/CD26) is a multifunctional type II transmembrane serine peptidase. This enzyme contributes to the regulation of various physiological processes, including blood sugar homeostasis, by cleaving peptide hormones, chemokines and neuropeptides. We have determined the 2.5 Å structure of the extracellular region of DPP-IV in complex with the inhibitor valine-pyrrolidide. The catalytic site is located in a large cavity formed between the α/β-hydrolase domain and an eight-bladed β-propeller domain. Both domains participate in inhibitor binding. The structure indicates how substrate specificity is achieved and reveals a new and unexpected opening to the active site.

The serine peptidase dipeptidyl peptidase IV (DPP-IV, CD26, EC 3.4.14.5) modulates the biological activity of several peptide hormones, chemokines and neuropeptides by specifically cleaving after a proline or alanine at amino acid position 2 from the N terminus[1]. Various genetic and pharmacological studies have revealed a prominent physiological role for this regulatory mechanism. Inhibitors of DPP-IV enzyme activity delay allograft rejection[2] and lessen experimental arthritis[3] and experimental autoimmune encephalomyelitis (EAE)[4] in animals. The DPP-IV substrates involved in these models of chronic inflammation have not yet been defined but may include chemokines. For instance, DPP-IV–mediated cleavage abrogates the activity of the monocyte chemotactic peptides MCP-1, -2 and -3, and changes the receptor specificity of the chemokine RANTES[5].

DPP-IV is important in maintaining physiological glucose homeostasis. Mice lacking the gene for DPP-IV show enhanced insulin secretion and accelerated clearance of blood glucose, partly because of increased endogenous levels of active glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP)[6]. GLP-1 and GIP are potent stimulators of insulin secretion, but their activity is rapidly abolished by DPP-IV–mediated truncation. Thus, DPP-IV contributes to the tight control of glucose levels by terminating GLP-1 and GIP signalling.

A significant, rapidly growing fraction of the human population is affected by type 2 diabetes, a disease characterized by elevated blood glucose levels and relative insufficiency of insulin. Pharmacological inhibition of DPP-IV activity increases insulin secretion and improves glucose control in diabetic animals[7–10] and in humans[11]. Thus, DPP-IV inhibition is a promising new strategy for treating type 2 diabetes, and DPP-IV inhibitors are now in clinical trials. In this context, knowledge of the three-dimensional structure of DPP-IV is important for designing new DPP-IV inhibitor drugs and understanding the structure–activity relationship of known inhibitors.

DPP-IV is expressed as a 220 kDa homodimeric type II transmembrane glycoprotein on the surface of various cell types, including epithelial and endothelial cells, and lymphocytes[5]. DPP-IV functions as a binding partner for other proteins, including adenosine deaminase (ADA)[12], the kidney Na^+/H^+ exchanger[13] and the T-cell antigen CD45 (ref. 14). Cell surface-bound DPP-IV is involved in T cell co-stimulation and tumor suppression[15,16]. Proteolytic cleavage of membrane bound DPP-IV results in a soluble form (amino acids 39–766) that circulates in the plasma. Both the membrane-bound and the soluble form show identical enzymatic activity[17].

Despite advances in understanding the biological functions of DPP-IV, the structural basis for the dipeptidyl peptidase activity of this enzyme and how it selectively chooses short peptides, such as GLP-1 and GIP, as substrates are still unknown. To address these questions, we determined the structure of a recombinant, soluble form of human DPP-IV that begins at residue Ser39, corresponding to the predominant form found in human plasma but with a more homogenous glycosylation pattern (see Methods). The recombinant soluble DPP-IV retains enzyme activity (data not shown). A complex of DPP-IV and a substrate analog, the competitive inhibitor valine-pyrrolidide (Val-Pyr) (K_i = 2 μM), was crystallized and diffraction data collected at 2.5 Å resolution. The structure was solved using the MAD method.

Overall structure of DPP-IV

DPP-IV is a dimer in the crystal (Fig.1), in agreement with previous biochemical data reporting that the active enzyme is a dimer[18]. The N terminus of each subunit is located at the same site of the dimer (Fig. 1*a*), indicating that full-length, membrane-bound DPP-IV could exist as homodimers at the cell surface. The surface facing the membrane is positively charged and complements the negatively charged phospholipids at the cell-membrane surface. This indicates that DPP-IV, when bound to the membrane, is in close contact with the cell surface. Each

[1]Protein Chemistry, Research and Development, Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd, Denmark. [2]Biotechnology, Research and Development, Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd, Denmark. [3]Present address: Symphogen A/S, Elektrovej building 375, 2800 Lyngby, Denmark.

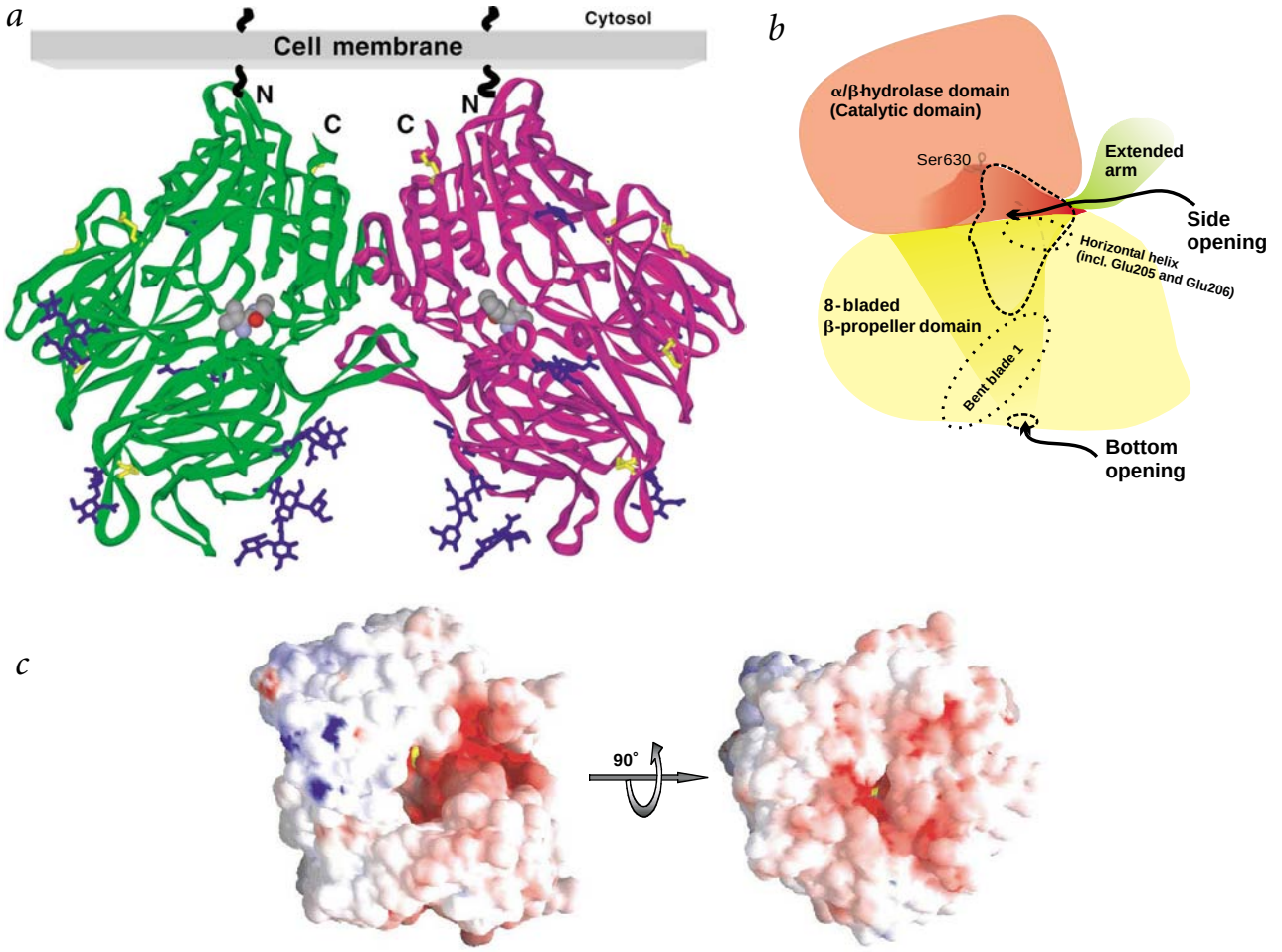

Fig. 1 Structure of DPP-IV. ***a***, DPP-IV forms a homodimer (subunit A shown in green and subunit B in magenta). Each subunit consists of two domains: an α/β-hydrolase domain and a β-propeller domain. The full-length DPP-IV is a type II transmembrane protein in which amino acids 7–28 constitute the membrane spanning region. The α/β-hydrolase domain, located closest to the membrane, consists of amino acids 39–51 and 506–766, and contains the active triad Ser630, Asp708 and His740. The eight-bladed β-propeller domain is formed by residues 55–497. The inhibitor Val-Pyr is shown in CPK and colored by element: carbon (gray), nitrogen (blue) and oxygen (red). Five S–S bridges (yellow) have been identified in each molecule: Cys328–Cys339 (blade 5), Cys385–Cys394 (blade 6), Cys444–Cys447 (blade 7), Cys454–Cys472 (blade 8) and Cys649–Cys762 (hydrolase). Carbohydrates (blue) have been located in the electron density map at seven of nine possible *N*-glycosylation positions in both subunits, holding up to three branched proximal glycoside units per position. The glycosylation sites are all situated in the β-propeller domain, except one, and clustered mainly to subdomain two (blade 2–5) of the β-propeller. ***b***, Schematic illustration of DPP-IV showing the individual domains and important structural elements. ***c***, The surface of molecule A of DPP-IV colored by electrostatic potential. The negatively charged surface is red and positively charged surface is blue, viewed from the side and the bottom of the propeller. The inhibitor, Val-Pyr is shown in CPK (yellow). The figure was generated by GRASP[35].

subunit consists of two domains, an α/β-hydrolase domain and an eight-bladed β-propeller domain (Fig. 1*b*). Between these two domains, a large cavity of ~30–45 Å width is found. The Val-Pyr molecule is bound in a smaller pocket within the cavity, next to the serine-protease active triad, Ser630, Asp708 and His740. The large cavity is accessible *via* two openings. Substrates and products may pass either through a funnel in the center of the propeller domain or through a much bigger opening in the side, between the hydrolase and propeller domains (discussed below).

The α/β-hydrolase domain consists of a central β-sheet sandwiched by α-helices, as in other members of the α/β-hydrolase family. In DPP-IV, the hydrolase domain is assembled by the C-terminal sequence (residues 506–766) and a short stretch of the N-terminal sequence (residues 39–51). The eight β-strands form a central β-sheet, with six arranged in a parallel and two in an antiparallel manner. The sheet is strongly twisted, with the α-helices pack against the sheet, with two on one side and four on the other. Furthermore, two small α-helices extend the β-propeller domain and become part of the hydrolase domain. The closest structural homolog of the α/β-hydrolase domain of DPP-IV for which the three-dimensional structure is known is prolyl oligopeptidase (POP)[19], an endopeptidase cleaving after proline (r.m.s. deviation is 1.84 Å for 224 Cα atoms in the α/β-hydrolase domain). DPP-IV and POP belong to distinct subfamilies of the S9 clan of serine peptidases[20]. The active Ser630 in DPP-IV is situated in a so-called 'nucleophile elbow' in the sequence Gly-Trp-Ser-Tyr-Gly, compared with Gly-Gly-Ser-Asn-Gly in POP, which are both in agreement with the conserved sequence Gly-X-Ser-X-Gly for the α/β hydrolase family.

The N-terminal β-propeller domain of DPP-IV (amino acids 55–497) consists of eight blades, each made up of four antiparallel β-strands (Fig. 2*a*,*b*). Although the propeller is topolog-

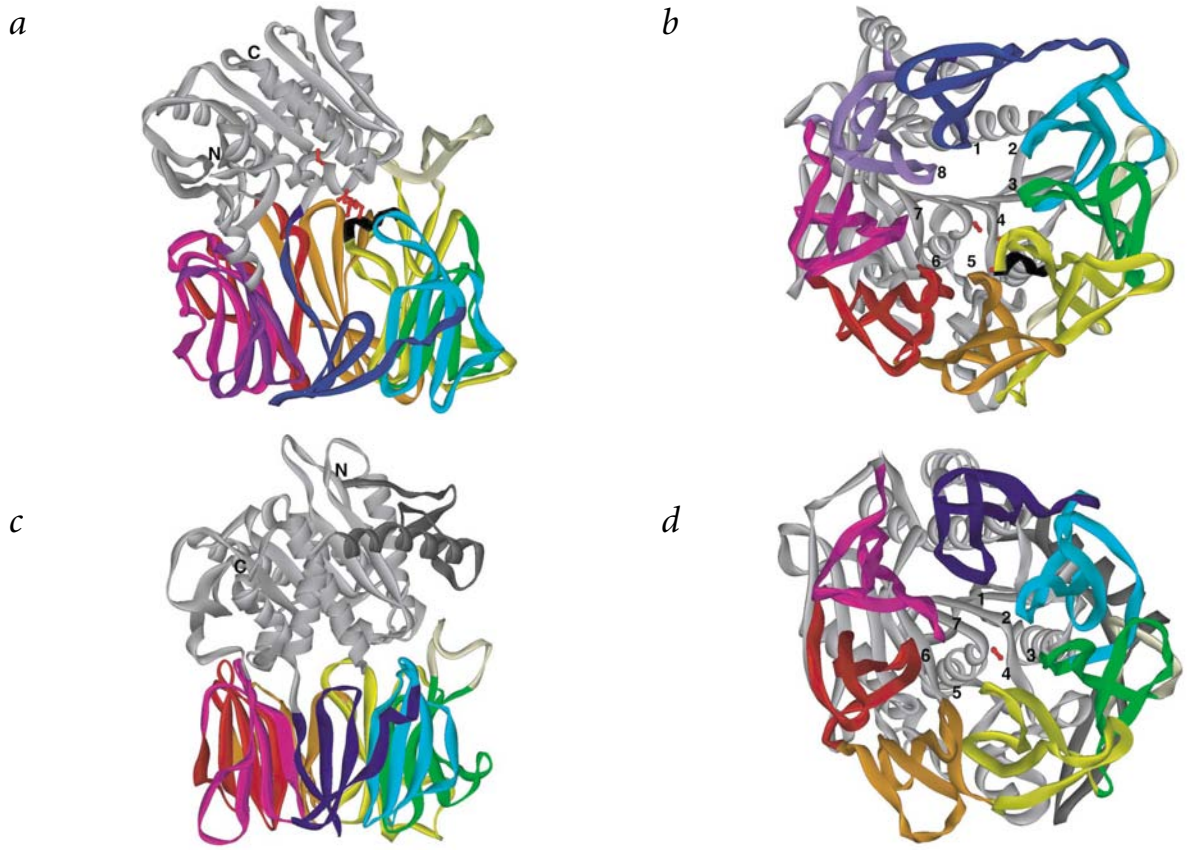

Fig. 2 Structural comparison of DPP-IV and POP. The eight (DPP-IV) and seven (POP) bladed β-propeller domains are illustrated by solid ribbon and colored by blade number. The α/β-hydrolase domain is gray. ***a***, DPP-IV viewed from the side. Glu205, Glu206 and Ser630 are illustrated by ball and stick (red). The extended arm is yellow, and the horizontal helix holding Glu205 and Glu206 is black. ***b***, DPP-IV viewed from the bottom. ***c***, POP viewed from the side. The N-terminal extension relative to DPP-IV is dark gray. ***d***, POP viewed from the bottom, with Ser554 illustrated in ball and stick (red).

divided into two subdomains (Fig. 2*a*), with blades 2–5 forming one subdomain; and blades 1 and 6–8, forming another.

All the blades of subdomain one (blades 2–5) bend toward the cavity, whereas those of subdomain two (blades 1 and 6–8) bend away from the cavity, especially blade 1, which is strongly bent outward (Fig. 2*a*). This results in an ellipsoid shape of the inside of the propeller (the major axis going from blade 2 to blade 6) rather than the circular shape, which is more common for this fold. POP has a seven-bladed propeller[19] that is much more regular than the one observed in DPP-IV. It is not possible to make meaningful sequence or structure alignments, considering the full β-propeller domains of DPP-IV and POP.

The inside of the propeller of DPP-IV forms a funnel-shaped tunnel between the active site and the bottom of the monomer. The bottom opening (defined according to the side-on view of the molecule, Fig. 2*a*), distal to the hydrolase domain, is ~7 Å × 14 Å; at the top, the opening widens to a diameter of ~24 Å. The length of the funnel is ~28 Å; however, the funnel is not the only opening to the active site. A pronounced side opening is created by (i) the pronounced bending of blade 1 in the propeller (Fig. 2*c*), (ii) a cleft formed by the smaller bending of blades 2–4 and some shorter β-strands in these blades and the β-turn holding the active Asp708, and (iii) an extended arm from the β-propeller (Fig. 2*a*). A calculation of the surface electrostatic potential shows that the cleft, together with the binding site for the inhibitor, is negatively charged (Fig. 1*c*). The cavity although not as strongly as in the side-opening cleft. On the basis of size and electrostatic characteristics, both the funnel and the side opening may serve as an entrance/exit to the active site.

The ADA-binding property of DPP-IV has been ablated by site directed mutagenesis of L294R or V341K[21]. L294 is situated in the small α-helix between blade 4 and 5, and V341 is situated in a long loop between β-strands 3 and 4 of blade 5. Both residues are exposed to solvent, located only 15 Å (Cα distance) apart from one another on the backside of the propeller looking at the molecule from the side (Fig. 2*a*).

A sequence motif has been associated with the β-propeller proteins, the 'WD motif'[22]. The β-propeller of DPP-IV has some of the WD repeat characteristics in some of the eight blades; however, a consensus throughout the whole propeller is not found. The WD repeat is named from the motif found at the end of the third strand, namely amino acids Trp-Asp. This motif is not found in the propeller blades of DPP-IV, although homologous residues are observed in several of the blades.

The active dimer

In POP the hydrolase domain has a 56-amino acid N-terminal extension relative to DPP-IV (Fig. 2*c*). Part of this extension (residues 1–10 and 39–54) forms a structural patch that, based on superimposition of the POP monomer on the DPP-IV dimer, aligns with a corresponding patch in DPP-IV that is formed by the C-terminal residues 745–761 (Fig. 3) and residues 729–737

Fig. 3 Secondary structure of DPP-IV shown over the amino acid sequence. Arrows indicate β-strands, and bars indicate helices. Color code is according to propeller blade number as in Fig. 2. The hydrolase domain is gray. Glycosylated residues located in the electron density map are marked with a cyan squared background. Cysteines involved in S–S bonds are marked with a yellow circular background and the corresponding residue number is listed above. The active triad residues are marked by a red asterisk, and other important residues for inhibitor binding are marked by an orange pentagonal background.

monomeric POP in itself fills part of the dimeric structure observed in DPP-IV.

The extended arm from the propeller participates strongly in the dimerization of DPP-IV (Fig. 1). β-strand 2 of blade 4 of the propeller extends into a small domain (amino acids 234–260) that includes an antiparallel two-stranded β-sheet. The function of this arm is clearly to stabilize the dimeric structure. POP also has an extended arm (residues 192–206) at the same spatial position relative to the hydrolase domain as in DPP-IV, but protruding from blade 3 rather than blade 4. Although the loop in POP is much shorter than that in DPP-IV, it is noteworthy that this loop is present in both structures despite their apparent monomeric *versus* dimeric structures and the low level of sequence and structural alignment between the propellers in the two peptidases.

Besides participating in the dimerization of DPP-IV, the extended arm also gives the impression of a lid function, as it is situated close to the active site and has a size that matches the side opening. If the dimer dissociates, this arm could move towards the cavity, thereby closing the side opening of the active site. Thus, the extended arm may provide a structural explanation for the observations that DPP-IV is enzymatically active as a dimer[18]. In addition to the extended arm, residues 658–661, 713–736, and 746–757 contribute to dimerization. The dimer interface is hydrophilic, indicating that DPP-IV might be able to exist in solution as a monomer. The dimer interface is 2,220 Å^2, corresponding to only 7% of the total accessible surface area of one subunit. These structural observations, in combination with reported biochemical data, raise the possibility that DPP-IV may exist either as an open, enzymatically active dimer or as a closed, inactive monomer, and that dimerization-mediated loop movement may represent a new type of mechanism for regulating the enzyme activity.

The topology of the propeller allows for some flexibility. All individual blades (Fig. 1*a*), rather than between the blades. Furthermore, the N and C termini of the propeller do not both participate in the same β-sheet — that is, there is no 'Velcro' closure of the propeller[23], a common stabilizing feature in many β-propellers. Finally, the crystallographic *B*-values of blade 1 (48 and 64 Å^2 for molecule A and B, respectively) are increased relative to the *B*-values of the rest of the blades (30 and 32 Å^2 for molecule A and B, respectively). This suggests flexibility for this bent blade, which may even straighten and thereby close the propeller. By itself or together with the extended arm, blade 1 may regulate substrate entry or exit from the active site by an opening/closing mechanism.

Several features of the DPP-IV structure suggest that the entrance to the active site is *via* the large side opening. This is the shortest and most easily accessible way to the active site. The negative potential in the cleft would attract the positively charged N terminus of the peptide substrate into the cavity. By entering this way rather than through the propeller, the peptide is orientated correctly for cleavage, as determined from analogy with the orientation of the pseudopeptide Val-Pyr. In contrast, a peptide entering through the propeller would have to make a turn inside the cavity. Once the dipeptide has been cleaved from the substrate, it may exit through the propeller funnel. Although we favor the hypothesis that substrate may access the active site *via* the side opening in DPP-IV, we can not rule out that entry may occur *via* the propeller funnel, as suggested for POP, whose structure revealed a single opening through the propeller funnel. As in POP, the bottom propeller opening of DPP-IV is negatively charged and, therefore, might attract the positively charged N terminus of peptides. Introduction of a S–S bridge by mutagenesis between blade 7 and blade 1 in POP abolished enzyme activity[24], supporting the theory of an entry through the propeller. POP was suggested to possibly be able to open the propeller even further than the 4 Å seen in the structure because of

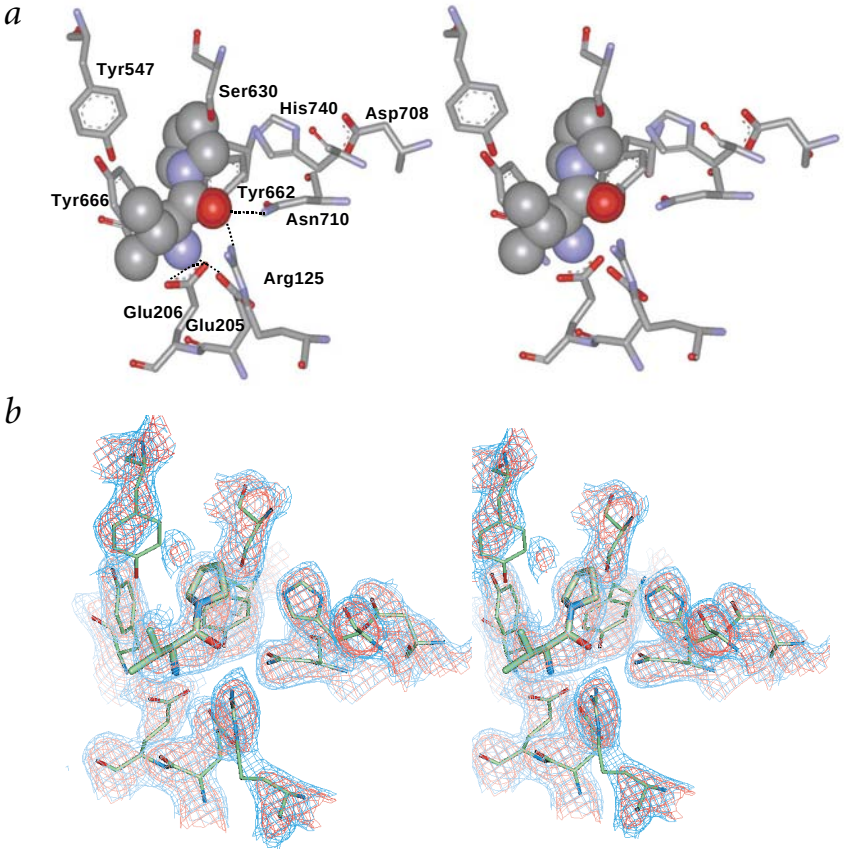

Fig. 4 Active site of DPP-IV. ***a***, Stereo drawing of the active site of DPP-IV. Val-Pyr is shown in CPK, with selected residues are shown in stick and colored by element type as in Fig. 1*a*. Hydrogen bonds/electrostatic interactions are shown with dotted lines. Glu205 and Glu206 coordinate the N terminus. Arg125 and Asn710 coordinate the carbonyl group of the peptide bond before the cleavage point. The pyrollidide ring is packed between Tyr547, Tyr762 and Tyr666. The valine moiety of the inhibitor has no contacts but points towards the pocket. ***b***, Val-Pyr is shown in thick sticks and selected residues (as in *a*) in thin sticks. The initial MAD electron density map after SHARP refinement at 2.5 Å resolution is overlaid contoured at 1 σ (blue) and 2 σ (red). ***c***, The chemical structure of the inhibitor Val-Pyr.

nation is that the flexible closure of the propeller may allow the enzyme to open its propeller on the side by bending blade 1, as seen in the structure of DPP-IV. Further functional studies are needed to address these possibilities.

Although the side opening in DPP-IV is large, the active site is located in a small pocket within the large cavity. Thus, only elongated peptides, or unfolded or partly unfolded protein fragments, can reach the site. This explains why most natural DPP-IV substrates are peptides <80 amino acids. However, larger proteins might be substrates as long as they have an unfolded N-terminal region.

The binding site of DPP-IV

The structure of the active site reveals how substrate specificity is achieved in DPP-IV, which preferentially cleaves after Xaa-Pro or Xaa-Ala (Xaa being any amino acid). Residues from both the hydrolase domain and the propeller domain take part in the binding of the substrate analog inhibitor Val-Pyr (Fig. 4). Two glutamic acids, Glu205 and Glu206, form salt bridges to the free amino group of Val-Pyr, which corresponds to the N terminus of a peptide substrate. Glu205 and Glu206 are situated in a small horizontal helix (residues 201–207) in blade 4 of the propeller domain (Fig. 2*a*). This small horizontal helix narrows the active site itself, leaving room for only two amino acids before the peptide reaches the active-site serine, making DPP-IV a dipeptidyl peptidase. The spatial arrangement of the active-site Ser630 with respect to Glu205 and Glu206 in DPP-IV probably makes the two glutamic acids the most important feature for alignment of the peptide before cleavage. This notion is consistent with studies showing that mutation of either one of these glutamic acids destroys the enzymatic activity of DPP-IV[25]. The pyrollidide active serine. Only amino acids with smaller side chains (proline, alanine or glycine[1]) will be able to fit into this narrow pocket, thereby restricting possible residues at the P1 position in substrates. Tyr662 and Tyr666 stack at each side of the pyrollidide ring of Val-Pyr — Tyr662 in a parallel fashion and Tyr666 in an orthogonal fashion — and are therefore part of the features determining the preference for a proline preceding the scissile bond. The valine side chain of Val-Pyr points into the large cavity and does not make specific contacts with DPP-IV, explaining why DPP-IV has no specific requirements for the N-terminal amino acid in the P2 position. The carbonyl oxygen of Val-Pyr forms hydrogen bonds to Asn710 and Arg125. The oxyanion hole is probably formed by Tyr547, as predicted[21,26,27], together with the backbone NH of Tyr631.

The residue of POP equivalent to Asn710 of DPP-IV is an arginine. The guanidine part of this arginine is situated close to the guanidine of Arg125 in DPP-IV, although it is distant in the Cα alignment. The side chain of Tyr662 in DPP-IV aligns structurally with the side chain of Trp595 in POP, whereas the backbone of Trp595 of POP aligns structurally with Tyr666 of DPP-IV. The essential Cys255 in POP, sitting in the large P2 pocket of DPP-IV in a superimposition based on the active triad of the two enzymes, does not have a homologous counterresidue in DPP-IV. Thus, DPP-IV and POP show some homology in the upper part of the catalytic cavity orienting the molecule (Fig. 4*a*). In contrast, significant differences are present in the bottom part of these molecules, explaining their different substrate specificities.

Concluding remarks

Several homology models of DPP-IV have been made based on

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.