Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

High-resolution structure of human apo dipeptidyl peptidase IV/CD26 and its complex with 1-[({2-[(5-iodopyridin-2-yl)amino]-ethyl}amino)-acetyl]-2-cyano-(*S*)-pyrrolidine

Christian Oefner,* Allan D'Arcy, Aengus Mac Sweeney, Sabine Pierau, Rana Gardiner and Glenn E. Dale

Morphochem AG, Basel, Switzerland

Correspondence e-mail: christian.oefner@morphochem.ch

Received 15 April 2003
Accepted 7 May 2003

PDB Reference: Apo dipeptidyl peptidase IV/CD26, 1pfq.

Dipeptidyl peptidase IV is a multifunctional type II transmembrane serine protease glycoprotein. The high-resolution crystal structure of the homodimeric human apo dipeptidyl peptidase IV has been determined at 1.9 Å resolution. In addition, the structure of the binary complex with 1-[({2-[(5-iodopyridin-2-yl)amino]-ethyl}amino)-acetyl]-2-cyano-(*S*)-pyrrolidine has been solved, revealing the nature of the covalent interaction with the active-site serine.

1. Introduction

Dipeptidyl peptidase IV (DPP-IV, CD26; EC 3.4.14.5) is a multifunctional type II transmembrane serine protease glycoprotein that is also present in a soluble form in serum and semen (Hong & Doyle, 1990; Iwaki-Egawa *et al.*, 1998; Durinx *et al.*, 2000). The human DPP-IV cDNA encodes a protein of 766 amino acids, with six amino acids in the cytoplasm, 22 residues spanning the plasma membrane and 738 amino acids comprising the extracellular domain (Misumi *et al.*, 1992). DPP-IV has multiple properties, including a highly specific serine protease activity that cleaves N-terminal dipeptides from peptides with an L-proline or L-alanine as the pentultimate residue, where it inactivates or generates biologically active peptides (Torimoto *et al.*, 1992; Mentlein, 1999). Additionally, DPP-IV functions as an adenosine deaminase (ADA) binding protein and contributes to extracellular matrix binding (Houghton *et al.*, 1998; Morimoto & Schlossman, 1998; Schrader *et al.*, 1990). DPP-IV is expressed in all organs, primarily on the apical surfaces of epithelial cells and acinar cells (Mentlein, 1999). In the kidney, where the enzyme is exeptionally concentrated, DPP-IV is located in the cortex and found in the brush border and microvillus fractions. However, it is the expression of DPP-IV on the lymphocytes (immune function) and on endothelial cells (inactivation of circulating peptides) which is the clinical target of inhibitors (Drucker, 2003).

DPP-IV degrades and regulates the activity of the incretin hormone glucagon-like peptide 1 (GLP-1). Owing to the impressive antidiabetic actions of GLP-1, DPP-IV inhibitors are of particular interest for their therapeutic potential in the treatment of type II diabetes. The contribution of catalytic activity to blood-glucose control through GLP-1 inactivation has recently been confirmed (Marguet *et al.*, 2000). Several classes of DPP-IV inhibitors have been described for this enzyme, including diprotin A and B, α-amino boronic acid analogues of proline, dipeptides, amino acid pyrolidides and thiazolidides, diaryl phosphonates, 2-cyanopyrrolidines and the more recently described pyrrolidine-carbonitrile deriva-

tives (Villhauer *et al.*, 2002; De Meester *et al.*, 1997; Hughes *et al.*, 1999; Stockel-Maschek *et al.*, 2000; Flentke *et al.*, 1991). The 2-cyanopyrrolidine derivatives are predicted to form a covalent adduct with the active-site serine of DPP-IV, forming an imidate adduct (Hughes *et al.*, 1999). Recently, the structure of valine-pyrrolidide-inhibited DPP-IV has been reported, giving insight into the fold of the enzyme and indicating how substrate specificity is achieved (Rasmussen *et al.*, 2003). In this report, we describe the structure of the apo enzyme and that of the covalent complex with a pyrrolidine-carbonitrile derivative.

2. Materials and methods

2.1. Expression and purification of recombinant DPP-IV

For the production of recombinant protein, the soluble extracellular domain of human DPPIV (sDPPIV; residues 29–776) was amplified by PCR using cDNA isolated from activated T-cells, cloned into the Zero Blunt TOPO cloning vector (Invitrogen) and sequenced (Durinx *et al.*, 2000). The primers used for the PCR were Fwd, 5′-TGAATTC*AACAAAGGCACAAGTGATGCTACAGCTGACAGTCGC*-3′, and Rev, 5′-AGGTACC*TTGAGGTGCTAAGGTAAAGAGAAACATTGTTTTATGAAG*-3′ (italicized sequences denote nucleotides complementary to the DPPIV cDNA). The forward primer used for the PCR introduced an *Eco*RI restriction site and consequently added the amino acids glutamate and phenylalanine to the N-terminus of the protein. The reverse primer used in the PCR introduced a *Kpn*I site following the stop codon. The subsequent cloning and expression proceeded identically to that described for human NEP (Oefner *et al.*, 2000). The supernatant from 8 l of culture was passed through a 0.2 µm filter and finally concentrated to 50 ml by crossflow ultrafiltration (Skanette) using a 30 kDa microfiltration module (Skan AG). The concentrate was applied in 10 ml aliquots to a Superdex 75XK 26/100 column equilibrated with 25 m*M* Tris pH 8.0, 100 m*M* NaCl, 1 m*M* EDTA, 5% glycerol and 0.02% NaN_3 (buffer *A*). Fractions containing sDPPIV were collected and $(NH_4)_2SO_4$ was added to a concentration of 1.5 *M*. The protein was applied to a Fractogel EMD propyl 650 (S) XK 26/30 column equilibrated with 25 m*M* Tris pH 8.0, 1.5 *M* $(NH_4)_2SO_4$, 0.02% NaN_3 (buffer *B*) and eluted with 25 m*M* Tris pH 8.0, 1 m*M* EDTA, 5% glycerol and 0.02% NaN_3 (buffer *C*). The fractions were pooled, dialyzed against buffer *C* and loaded onto a Mono Q HR5/5 column equilibrated with buffer *C*. The protein was eluted with a shallow gradient of 25 m*M* Tris pH 8.0, 1 *M* NaCl, 1 m*M* EDTA, 5% glycerol and 0.02% NaN_3 (buffer *D*). The active fractions were pooled, concentrated and applied onto a BioSec 16/600 column equilibrated with buffer *A*. The protein eluted as a single peak at the volume consistent with a dimeric protein. Analytical size-exclusion chromatography was performed on a Superdex-200 column (PC 3.2/30) using a Agilent 1100 HPLC instrument with 50 µg of purified protein and buffer *A* as the mobile phase.

2.2. Crystallization and data collection

Prior to crystallization, the purified protein was examined using dynamic light scattering (Zulauf & D'Arcy, 1992). The protein showed a monodisperse size distribution and an estimated molecular weight of approximately 200 kDa, corresponding to a dimer in solution. The protein in 25 m*M* Tris pH 7.5, 100 m*M* NaCl, 2 m*M* TCEP and 5% glycerol was concentrated to 20 mg ml^{-1} and treated with glycosidase endo-F to remove N-linked sugars as described previously (Grueninger-Leitch *et al.*, 1996). The deglycosylated protein was screened for initial crystallization conditions using the modified microbatch method with an INDEX screen (Hampton Research cat. No. HR2-134; D'Arcy *et al.*, 2003; Chayen, 1998). Crystals appeared in a number of conditions containing polyethylene glycol as the precipitation agent. The most promising (25% PEG 3350, 200 m*M* $MgCl_2$, 100 m*M* HEPES 7.5) was optimized in a hanging-drop vapour-diffusion experiment. The crystals proved difficult to reproduce and seeding was necessary to induce nucleation of the preferred crystal form. A binary complex of the endo-F glycosidase-treated human DPP-IV has been formed by soaking in 5 m*M* 1-[({2-[(5-iodopyridin-2-yl)amino]-ethyl}amino)-acetyl]-2-cyano-(*S*)-pyrrolidine. Crystals belong to the orthorhombic space group $P2_12_12_1$ and contain two molecules in the asymmetric unit. For data collection, all crystals were flash-frozen in a cryoprotectant solution corresponding to the reservoir condition containing 25% PEG 3350 and 20% 1,6-hexanediol and data were collected at 100 K. A heavy-atom derivative was obtained after a 40 min soak in the presence of 0.2 mg ml^{-1} $HgCl_2$ in freezing solution. All data sets, except Native-2, were measured with Cu $K\alpha$ radiation from a Nonius FR591 rotating-anode generator equipped with an Osmic mirror system and were recorded on a MAR Research image-plate area detector. The Native-2 data set used for refinement was collected at the SLS beamline X06SA at PSI, Villigen, Switzerland on a MAR CCD detector at a wavelength of 1.00563 Å. All diffraction data were processed and scaled with *DENZO* and *SCALEPACK* (Otwinowski, 1993) and were further analyzed with the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

2.3. Structure determination and refinement

A $HgCl_2$ derivative was solved from the isomorphous amplitude differences relative to Native-1 using the program *SHELXS*97 (Sheldrick *et al.*, 1993). Four sites were identified and heavy-atom refinement and phasing was subsequently performed with the program *SHARP* (de La Fortelle & Bricogne, 1997). The initial *SHARP* electron density was modified and further improved with the program *SOLOMON* by applying iterative rounds of solvent flattening and a final cycle of solvent flipping using a solvent content of 46% (Abrahams & Leslie, 1996). This led to a partially interpretable electron density at 2.7 Å resolution. Initial model building performed with the graphics program *MOLOC* (Gerber, 1992) allowed the determination of the non-crystallographic twofold axis. Subsequent non-crystallographic

symmetry averaging resulted in a modified density of excellent quality, which allowed the determination of both a nearly complete dimeric structure and a reliable sequence assignment. Iterative rounds of model building using *MOLOC* and stereochemically restrained positional and temperature-factor refinement with *REFMAC* (Murshudov *et al.*, 1997) were performed, using parameters for ideal stereochemistry as described by Engh & Huber (1991). Progressive introduction of solvent molecules with good geometry and asparagine-linked sugar residues led to a nearly complete uncomplexed structure with good stereochemistry as examined by *PROCHECK* (Laskowski *et al.*, 1993). All residues fall within the allowed regions of the Ramachandran plot, with the exception of Glu73 in both monomers, which however has a clearly defined backbone conformation. The structure of the apo enzyme converged to an R factor of 0.254 using all data in the resolution range 20.0–1.9 Å (122 631 reflections), with an $R_{\rm free}$ of 0.298. The refined apo structure of deglycosylated human DPP-IV was used to solve the binary complex with the iodinated cyanopyrrolidide derivative at 2.55 Å resolution. An initial Fourier difference map revealed residual electron density located in the active site corresponding to the covalently bound small molecule. Progressive introduction of solvent molecules with good geometry using the methods described above led to a refined complex with an R factor of 0.24. The results are summarized in Table 1.

Table 1
Crystallographic data and refinement statistics.

Data sets	Native-1	Native-2	$HgCl_2$	Inhibitor
Soaking conditions			0.2 mg ml^{-1}, 40 min	5 m*M*, 45 min
Data collection				
Exposure time per frame (s)	300	3	300	300
Angular increment per frame (°)	0.5	1.0	0.5	0.5
Total rotation range (°)	250	160	105	90
Crystal-to-detector distance (mm)	220	120	250	230
Unit-cell parameters *a*, *b*, *c* (Å)	70.9, 117.4, 184.6	71.0, 118.1, 184.6	71.0, 117.6, 184.7	70.9, 117.2, 183.8
Data reduction				
Maximum resolution (Å)	2.40	1.9	2.7	2.55
No. of measurements	552838	803624	142136	144890
No. of unique reflections	61137	122631	41592	45059
Completeness† (%)	99.9 (99.5)	99.7 (100)	95.9 (81.0)	88.6 (60.1)
$R_{\rm sym}$†‡ (%)	10.9 (54.7)	12.6 (64.8)	15.0 (59.3)	15.1 (62.6)
$I/\sigma(I)$†	17.0 (3.2)	12.3 (1.0)	7.3 (1.4)	5.8 (1.0)
Heavy-atom refinement parameters				
Figure of merit (acentric/centric)	0.35/0.42			
$R_{\rm cullis}$§ (acentric/centric)			0.73/0.73	
Phasing power¶ (acentric/centric)			1.19/0.99	
Refinement statistics				
Resolution range (Å)		20–1.9		20–2.55
$R_{\rm cryst}$ ($R_{\rm free}$)†† (%)		25.4 (29.8)		24.0 (32.0)
No. of protein atoms‡‡ (mean B in Å^2)		11891 (42.5)		11891 (25.7)
No. of water molecules		473		299
No. of ligand atoms (mean B in Å^2)				42 (31.5)
No. of NAG atoms (mean B in Å^2)		56 (51.2)		56 (37.5)
R.m.s.d.§§ bonds (Å^2)		0.004		0.013
R.m.s.d.§§ angles (°)		0.65		1.407

† Values in parentheses are statistics for the highest resolution bin. ‡ $R_{\rm sym} = \sum_h \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I(h)$, where $I_i(h)$ and $\langle I(h)\rangle$ are the ith and mean measurement of the intensity of reflection h, respectively. § $R_{\rm cullis} = \sum_h F_D(h) - |F_N(h) + |F_H(h)|| / \sum_h |F_D(h) - F_N(h)|$, where $F_H(h)$, $F_D(h)$ and $F_N(h)$ are the heavy-atom, derivative and native structure-factor amplitudes of reflection h, respectively. ¶ Phasing power $= \sum_h F_H(h) / \sum_h |F_D(h) - |F_N(h) + F_H(h)||$. †† $\sum ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_h |F_{\rm obs}|$, where $|F_{\rm obs}|$ and $|F_{\rm calc}|$ are the observed and calculated structure-factor amplitudes for the reflection h, applied to the working ($R_{\rm cryst}$) and test ($R_{\rm free}$) sets, respectively. ‡‡ Non-H atoms only. §§ R.m.s.d.: root-mean-square deviation from mean.

3. Results and discussion

3.1. Overall structure of DPP-IV

The crystal structure of fully glycosylated human DPP-IV has been solved and reported in complex with the substrate analogue valine-pyrrolidide at 2.5 Å resolution (Rasmussen *et al.*, 2003). Here, we present the high-resolution glycosidase-treated homodimeric apo structure (Fig. 1), which has a bent dumbell shape of approximately 120 Å in length, and its complex with 1-[({2-[(5-iodopyridin-2-yl)amino]-ethyl}amino)-acetyl]-2-cyano-(*S*)-pyrrolidine (Fig. 2). Each end of the biological dimer consists of an eight-bladed β-propeller domain embedded within the sequence of the central α/β-hydrolase domain. The central α/β-hydrolase domain ranges from residues 39 to 51 and from 506 to 766 and participates in the formation of the rigid dimer interface. It contains the catalytic triad Ser630, Asp708 and His740, which has been confirmed by point mutation of mouse CD26 as Ser624, Asp702 and His734 (David *et al.*, 1993). The domain is covered by the N-terminal β-propeller fold and the active site is positioned near the center of each monomeric molecule, forming the base of a large cavity of ~30–45 Å in width. The α/β-hydrolase domain consists of a central eight-stranded β-sheet (six parallel and two antiparallel strands) sandwiched by four α-helices on one side and two on the other. The β-sheet is strongly twisted by approximately 120°, with the outermost strand (Gln731–Tyr735) participating in dimer-interface formation. The closest known structural homologue of the α/β-hydrolase domain is that of the related S9 endopeptidase prolyl oligopeptidase (POP; Gorrell *et al.*, 2001; Abbott *et al.*, 1999), which is covered by the propeller, forming a gating filter of substrates in POP (Fülöp *et al.*, 2000). The larger N-terminal β-propeller domain of DPP-IV comprises eight blades, each made up of four antiparallel β-strands. The first N-terminal blade 1 bends strongly outwards from the cavity, resulting in an ellipsoid shape of the interior of the propeller. Various structural elements from both the α/β-hydrolase and the β-propeller domain of DPP-IV contribute to the formation of a tight non-crystallographic dimer interface. These include the C-terminal $\alpha\beta\alpha$ structure as well as an antiparallel two-stranded β-sheet of the β-propeller, which

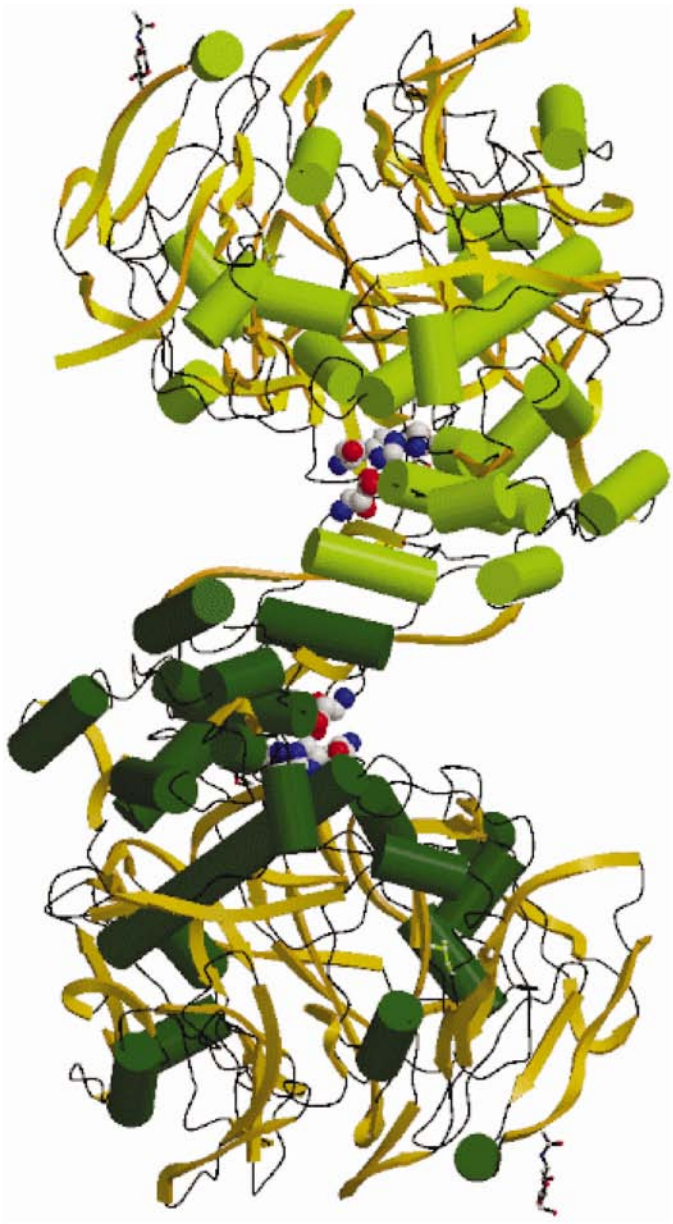
forms a small extended domain ranging from residues Pro234 to Pro255. The calculated solvent-accessible surface area of both the non-crystallographic dimer and the individual monomers shows that ~4400 Å^2 of the total surface area are buried in the dimer, of which ~2800 Å^2 has hydrophobic character. Approximately half of the total surface area is contributed by the small β-propeller domain.

The monomeric DPP-IV molecule comprises five disulfide bridges, four of which are located in the propeller domain. These are Cys328–Cys339, Cys385–Cys394, Cys444–Cys447 and Cys454–Cys472, which are formed within blades 5, 6, 7 and 8, respectively, as well as Cys649–Cys762, which is present in the α/β-hydrolase domain and links the C-terminal α-helix with the β-strand formed by residues Cys649–Val653. N-Glycosylation of the endo-F-treated protein has been found on two of the nine possible residues in each monomer. These occur in the propeller domain within blade 1 and blade 4 and concern Asn85 and Asn229; the glycosylation has been confirmed by electron density consistent with a single covalently bound *N*-acetyl-glucosamine (NAG) moiety.

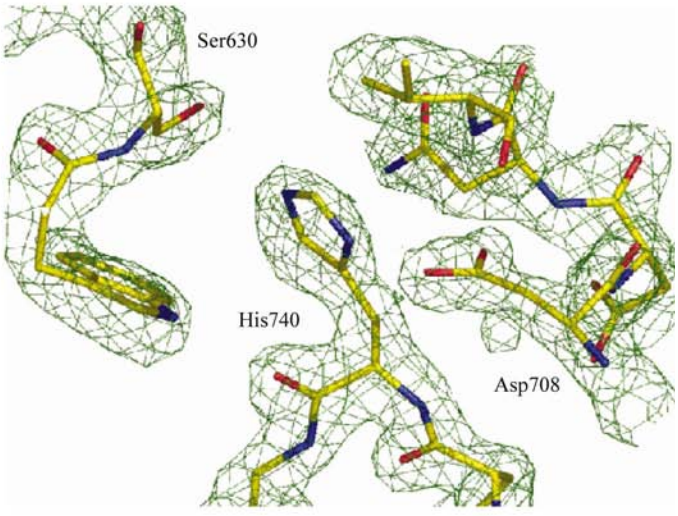
3.2. Active-site architecture

The active-site area of the post-proline and post-alanine cleaving enzyme is composed of fragments from both domains, whereas the catalytic residues Ser630, Asp708 and His740 exclusively belong to the α/β-hydrolase region. Comparing this part of the structure of the non-inhibited enzyme with that of the binary complex, it becomes obvious that there are only minor changes which correspond to some side-chain rearrangements, *e.g.* Arg125 becomes ordered, with the guanidinium group pointing towards the catalytic residues. The major rearrangement is observed for the β-indole of Trp639. In the apo form it is found in an up conformation, shielding the catalytic serine residue from access. Within the inhibited enzyme, however, it adopts a down conformation, changing the dihedral angle χ_2 from 79 to -54° (Fig. 3). This conformation is also observed for the complex with the substrate-analogue inhibitor valine-pyrrolidide (Rasmussen *et al.*, 2003). In all known structures, the indole nitrogen is hydrogen bonded to the carbonyl O atom of His740, providing a hinge for the side-chain rearrangement.

The iodinated variant of the inhibitor NVP-DPP728, which has been synthesized according to Villhauer (1998) by repla-



(*a*)



(*b*)

Figure 1
(*a*) Overall structure of the dimeric native deglycosylated human DPP-IV. α-Helices are shown as green cylinders; yellow arrows denote β-strands. Residues of the catalytic site are presented as cpk models. The view is approximately along the non-crystallographic twofold axes. The figure was generated with the programs *MOLSCRIPT* and *Raster3D* (Kraulis, 1991; Merritt & Bacon, 1997). (*b*) shows the active-site residues Ser630, Asp708 and His740 superimposed with the final $2F_o - F_c$ electron density calculated with phases from the refined structure and contoured at 1.6σ. The figure was generated with *PyMOL* (DeLano, 2001).

cing 5-cyano-2-chloropyridine with 5-iodo-2-chloropyridine, is covalently bound to the active-site serine by the cyanopyrrolidine part of the small molecule. It has been shown that the potency of NVP-DPP728 is derived from a slow-binding inhibition mechanism that is dependent upon the nitrile functionality of the inhibitor (Hughes *et al.*, 1999). It forms a transient imidate intermediate, which is in agreement with the proposed model of inhibition of DPP-IV by NVP-DPP728 described by Hughes and coworkers. The nitrogen of the imidate intermediate is within hydrogen-bonding distance of the hydroxyl group of Tyr547 and the pyrrolidide moiety is buried in a narrow hydrophobic S1 pocket next to the active serine residue as observed for the substrate-analogue inhibitor. The iodinated molecule is further stabilized through the hydrophilic interactions of OE2 and the main-chain carbonyl O atom of Glu205 with the N atoms of the ethylenediamine spacer as shown in Fig. 3. The hydrogen-bonding distances are 3.1 and 2.4 Å, respectively. A further hydrogen bond is observed between the side chain of Ser209 and the pyridine nitrogen. A very weak interaction has been observed between the side chains of Arg125 and Asn710 and the peptide-bond carbonyl of the enzyme; however, there is no direct involvement of Glu206 in inhibitor binding. The planar 5-iodopyridine moiety extends towards the S2 subsite and is involved in hydrophobic interactions with the side chain of Phe357. A boundary of this subsite is formed by the hydrophobic face of the guanidinium group of Arg358. Approximately 40% of the total solvent-accessible surface area of 1150 Å^2 of the inhibitor is in direct contact with the enzyme, mainly owing to the recognition of one face of the small molecule in its extended conformation. Nearly half of the observed interaction surface is of hydrophobic origin, of which a large portion can be attributed to the S1 subsite of the enzyme. This subsite is formed by residues Tyr631, Val656, Trp659, Tyr662, Tyr666 and Val711, which form the hydrophobic recognition pocket for a proline residue preceding the scissile bond.

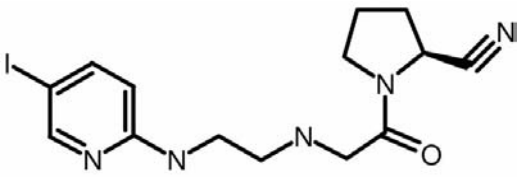
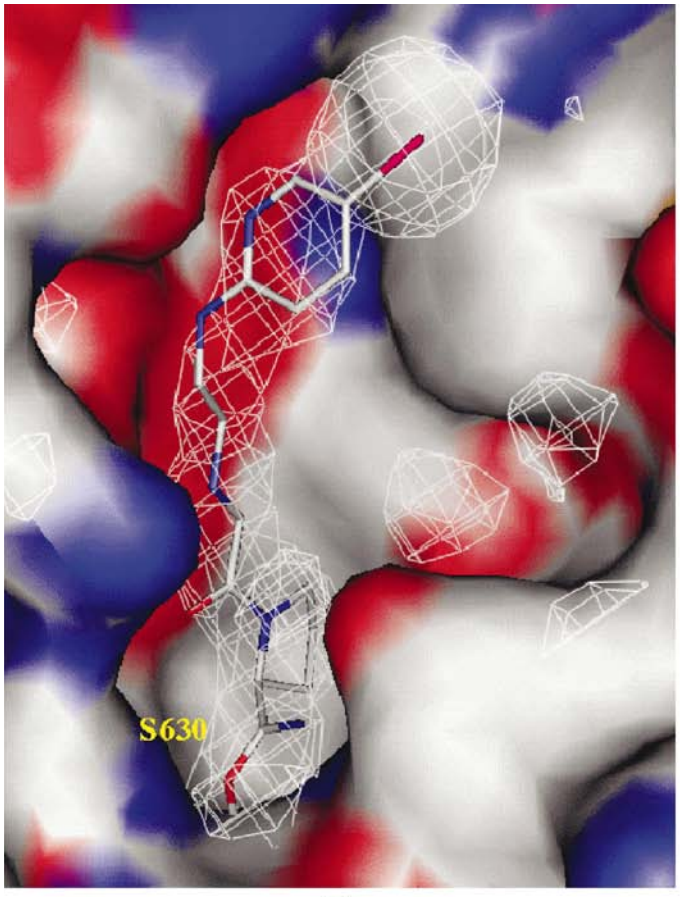
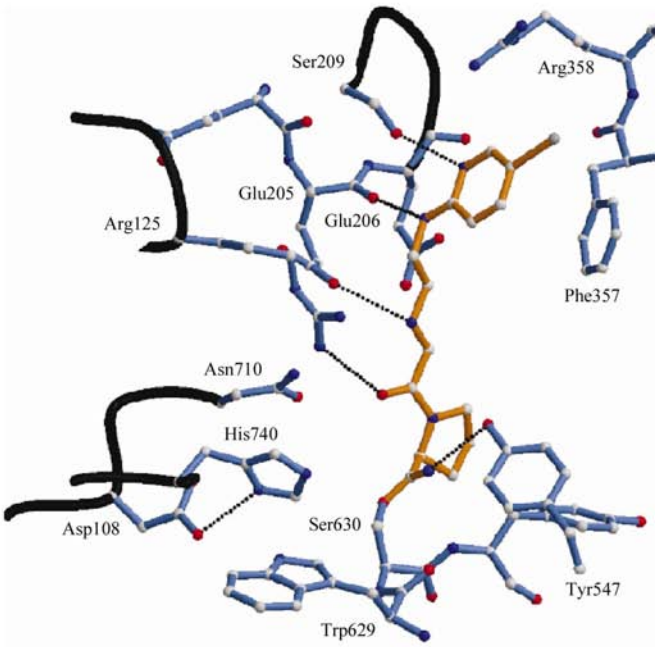


Figure 2
Structure of the competitive iodinated inhibitor 1-[({2-[(5-iodopyridin-2-yl)amino]-ethyl}amino)-acetyl]-2-cyano-(*S*)-pyrrolidine, which is a derivative of the Novartis compound NVP-DPP728.



(*a*)



(*b*)

Figure 3
Inhibitor binding to DPP-IV. (*a*) Active-site surface area superimposed with the $2F_o - F_c$ electron-density map of the inhibitor at 2.55 Å resolution, calculated with phases from the refined model. The map contoured at 1.6σ is shown in white and the active-site serine is indicated. Residual density belongs to water molecules, which are not shown. The molecular surface is coloured by the electrostatic potential: blue, positive; red, negative. The figure was generated using *PyMOL* (DeLano, 2001). (*b*) Intramolecular interactions between DPP-IV and the covalently bound inhibitor. Hydrogen bonds are indicated by dashed lines. Both pictures have approximately the same orientation. (*b*) was generated with the programs *MOLSCRIPT* and *Raster3D* (Kraulis, 1991; Merritt & Bacon, 1997).

Table 2
Crystal contacts.

Contacts between dimers	Total buried surface area (Å^2)			
	Deglycosylated material		Glycosylated material	
	Including sugar	Excluding sugar	Including sugar	Excluding sugar
$x, y, z \leftrightarrow x+1, y, z$	2 × 1592	2 × 1444	2 × 1011	2 × 838
$x, y, z \leftrightarrow \frac{1}{2}-x, -y, z+\frac{1}{2}$	2 × 736	2 × 736	2 × 1912	2 × 1537
$x, y, z \leftrightarrow -x, y+\frac{1}{2}, \frac{1}{2}-z$	2 × 1225	2 × 1097	2 × 894	2 × 739
$x, y, z \leftrightarrow \frac{1}{2}+x, \frac{1}{2}-y, -z$	2 × 935	2 × 935		
$x, y, z \leftrightarrow -x-1, y+\frac{1}{2}, \frac{1}{2}-z$			2 × 751	2 × 447

3.3. Crystal contacts

In the orthorhombic crystal form of human endo-F-treated DPP-IV, approximately 8% (8976 Å^2) of the total accessible surface of the dimeric protein is involved in crystal contacts. This involves four types of contacts which occur twice and contribute to total buried surface areas of 1592, 736, 1225 and 935 Å^2, respectively. The solvent-accessible surface areas are listed in Table 2 for the various crystal contacts. If the packing contribution of the four single NAG moieties are excluded, the contact area decreases only slightly, by approximately 500 Å^2 to 8424 Å^2. Therefore, the packing of the deglycosylated material is mainly protein-dependent, a situation which differs from that of the glycosylated material used in the structure determination of the complex with the substrate analogue valine-pyrrolidide, which crystallized in the same space group with a different unit cell (Rasmussen *et al.*, 2003). Here, the glycosylation is estimated to be 11% and is contributed by seven of the nine possible N-glycosylation positions in both subunits, giving rise to 30 identified sugar units. Although the total buried surface area of this protein is slightly higher at 9136 Å^2, it decreases by 22% in the absence of the sugar residues to 7122 Å^2. The crystal packing of the glycosylated enzyme is therefore influenced by the extended and branched glycoside units to a much larger extent compared with the deglycosylated crystallized material. The deglycosylation of the material increases the size of the total buried surface area by approximately 1300 Å^2, which could explain the superior diffraction quality of the crystallized material. This observation coincides with a 7% decrease in the solvent content to 45.5%.

4. Conclusions

NVP-DPP728 has been identified as a potent and selective DPP-IV inhibitor for use in the treatment of diabetes mellitus. The compound inhibits human and rat plasma DPP-IV with IC_{50} values in the range 5–10 n*M*, with >15 000-fold selectivity relative to DPP-II and a range of proline-cleaving proteases (Hughes *et al.*, 1999). The structure of the complex of DPP-IV with the iodinated variant of the inhibitor NVP-DPP728 indicates that the mechanism of inhibition involves the formation of an imidate intermediate, which is different from the simple reversible competitive binding of dipeptide-like pyrrolidide compounds, *e.g.* valine pyrrolidide. This is achieved through a slow-binding mechanism which depends upon the L chirality of the pyrrolidine nitrile functionality (Hughes *et al.*, 1999) as confirmed by the complex structure. The high-affinity complex, which is dependent upon the nitrile functionality, is further stabilized by a number of hydrophilic interactions located on the concave surface of the active centre provided by residues of the S1 and S2 subsite. The comparison of the known complexed structures of DPP-IV with that of the apo enzyme clearly reveals a moderate flexibility of the peptidase upon small-molecule binding. Major structural alterations occur on the protein surface and are most likely a consequence of the different crystal contacts observed for the fully glycosylated and deglycosylated protein. Packing analysis indicates that a major fraction of the crystal contacts observed for the fully glycosylated protein involve glycoside units. This differs from the packing observed for the endo-F-treated protein, resulting in better diffracting protein crystals.

References

Abbott, C. A., McCaughan, G. W., Levy, M. T., Church, W. B. & Gorrell, M. D. (1999). *Eur. J. Biochem.* **266**, 798–810.
Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Chayen, N. E. (1998). *Acta Cryst.* D**54**, 8–15.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst* D**50**, 760–763.
D'Arcy, A., Mac Sweeney, A., Stihle, M. & Haber, A. (2003). *Acta Cryst.* D**59**, 396–399.
David, F., Bernard, A. M., Pierres, M. & Marguet, D. (1993). *J. Biol. Chem.* **268**, 17247–17252.
DeLano, W. L. (2001). *The PyMOL User's Manual.* DeLano Scientific, San Carlos, CA, USA.
De Meester, I., Belyaev, A., Lambeir, A. M., De Meyer, G. R., Van Osselaer, N., Haemers, A. & Scharpe, S. (1997). *Biochem. Pharmacol.* **54**, 173–179.
Drucker, D. J. (2003). *Expert. Opin. Investig. Drugs*, **12**, 87–100.
Durinx, C., Lambeir, A. M., Bosmans, E., Falmagne, J. B., Berghmans, R., Haemers, A., Scharpe, S. & De Meester, I. (2000). *Eur. J. Biochem.* **267**, 5608–5613.
Engh, R. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Flentke, G. R., Munoz, E., Huber, B. T., Plaut, A. G., Kettner, C. A. & Bachovchin, W. W. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 1556–1559.
Fülöp, V., Szeltner, Z. & Polgar, L. (2000). *EMBO Rep.* **1**, 277–281.
Gerber, P. (1992). *Biopolymers*, **32**, 1003–1017.
Gorrell, M. D., Gysbers, V. & McCaughan, G. W. (2001). *Scand. J. Immunol.* **54**, 249–264.
Grueninger-Leitch, F., D'Arcy, A., D'Arcy, B. & Chene, C. (1996). *Protein Sci.* **12**, 2617–2622.
Hong, W. J. & Doyle, D. (1990). *J. Cell. Biol.* **111**, 323–328.
Houghton, A. N., Albino, A. P., Cordon-Cardo, C., Davis, L. J. & Eisinger, M. (1998). *J. Exp. Med.* **167**, 197–212.
Hughes, T. E., Mone, M. D., Russell, M. E., Weldon, S. C. & Villhauer, E. B. (1999). *Biochemistry*, **38**, 11597–11603.
Iwaki-Egawa, S., Watanabe, Y., Kikuya, Y. & Fujimoto, Y. (1998). *J. Biochem.* **124**, 428–433.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Marguet, D., Baggio, L., Kobayashi, T., Bernard, A. M., Pierres, M., Nielsen, P. F., Ribel, U., Watanabe, T., Drucker, D. J. & Wagtmann, N. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 6874–6879.
Mentlein, R. (1999). *Regul. Pept.* **85**, 9–24.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Misumi, Y., Hayashi, Y., Arakawa, F. & Ikehara, Y. (1992). *Biochim. Biophys. Acta*, **1131**, 333–336.
Morimoto, C. & Schlossman, S. F. (1998). *Immunol. Rev.* **161**, 55–70.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Oefner, C., D'Arcy, A., Hennig, M., Winkler, F. K. & Dale, G. E. (2000). *J. Mol. Biol.* **296**, 341–349.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Rasmussen, H. B., Branner, S., Wiberg, F. C. & Wagtmann, N. (2003). *Nature Struct. Biol.* **10**, 19–25.
Schrader, W. P., West, C. A., Miczek, A. D. & Norton, E. K. (1990). *J. Biol. Chem.* **265**, 19312–19318.
Sheldrick, G. M., Dauter, Z., Wilson, K. S., Hope, H. & Seiker, L. C. (1993). *Acta Cryst.* D**49**, 18–23.
Stockel-Maschek, A., Stiebitz, B., Born, I., Faust, J., Mogelin, W. & Neubert, K. (2000). *Adv. Exp. Med. Biol.* **477**, 117–123.
Torimoto, Y., Dang, N. H., Tanaka, T., Prado, C., Schlossman, S. F. & Morimoto, C. (1992). *Mol Immunol.* **29**, 183–192.
Villhauer, E. B. (1998). World Patent Application WO 98/1998.
Villhauer, E. B., Brinkman, J. A., Naderi, G. B., Dunning, B. E., Mangold, B. L., Mone, M. D., Russell, M. E., Weldon, S. C. & Hughes, T. E. (2002). *J. Med. Chem.* **45**, 2362–2365.
Zulauf, M. & D'Arcy, A. (1992). *J. Cryst. Growth*, **122**, 102–106.