the FEBS Journal

The structural comparison of the bacterial PepX and human DPP-IV reveals sites for the design of inhibitors of PepX activity

Pascal Rigolet[1], Xu Guang Xi[1], Stephane Rety[1] and Jean-François Chich[2]

1 Laboratoire de Biotechnologies et Pharmacologie Génétique Appliquée CNRS, Ecole Normale Supérieure (ENS) Cachan, France
2 Virologie et Immunologie Moléculaires, INRA, Jouy-en-Josas, France

Keywords
docking; DPP-IV inhibitors; PepX; X-prolyl dipeptidyl aminopeptidase

Correspondence
P. Rigolet, Laboratoire de Biotechnologies et Pharmacologie Génétique Appliquée CNRS, Ecole Normale Supérieure (ENS) Cachan, 61 avenue du Président Wilson, 94235 Cachan cedex, France
Fax: +33 1 47 40 76 71
Tel: +33 1 47 40 68 76
E-mail: Pascal.Rigolet@lbpa.ens-cachan.fr
Website: http://www.lbpa.ens-cachan.fr

(Received 21 January 2005, accepted 25 February 2005)

doi:10.1111/j.1742-4658.2005.04631.x

X-prolyl dipeptidyl aminopeptidases (X-PDAP) are enzymes catalysing the release of dipeptides from the amino termini of polypeptides containing a proline or an alanine at the penultimate position. Involved in various mammalian regulation processes, as well as in chronic human diseases, they have been proposed to play a role in pathogenicity for *Streptococci*. We compared the structure of X-PDAP from *Lactococcus lactis* (PepX) with its human counterpart DPP-IV. Despite very different overall folds, the residues most implicated for X-PDAP activity are conserved in the same positions and orientations in both enzymes, thus defining a structural signature for the X-PDAP specificity that crosses the species frontiers of evolution. Starting from this observation, we tested some inhibitors of DPP-IV on PepX activity, for which no specific inhibitor is known. We thus found that PepX was highly sensitive to valine-pyrrolidide with a K_I of 9.3 μM, close to that reported in DPP-IV inhibition. We finally used the structure of PepX from *L. lactis* as a template for computer-based homology modeling of PepX from the pathogenic *Streptococcus gordonii*. Docking simulations of valine-pyrrolidide into the active site of PepX led to the identification of key residues for a rational drug design against PepX from *Streptococci*. These results could have applications in human health giving new perspectives to the struggle against pathogens.

The X-prolyl dipeptidyl aminopeptidases (X-PDAP) are enzymes that remove X-Pro dipeptides from the N-terminus of peptides containing a proline or an alanine at the penultimate position. They are involved in various mammalian regulation processes as well as in serious human diseases. Present in lactic acid bacteria, which are used for various useful human activities, they are also found in *Streptococci* where they have been proposed to play a role in pathogenicity. Enzymes with such specificity are classified into the clan SC [1] in two distinct families, S9 and S15, according to structural and functional characteristics [1].

The type example of the S9 family is DPP-IV. Localized in the membrane of several cell types such as epithelial or endothelial cells [2], it is identical to the T-cell activation antigen CD26. DPP-IV is a multifunctional enzyme of 766 amino acids. The 3D structure of the cytoplasmic part of the human enzyme was recently solved at 1.8 Å resolution [3]. It has been shown to play a role in the activation or degradation of biological peptides, peptides hormones and neuropeptides [4,5] and to interact as a receptor or ligand with various proteins playing a role in the immune response. The enzyme has been described as deaminase binding protein [6] and appears to be associated with CD45 [7].

Abbreviations
DPP-IV, dipeptidyl peptidase IV; DPP-II, dipeptidyl peptidase II; LSQ, least squares; pNA, *para*-nitro-anilide; POP, prolyl oligopeptidase; SI, selectivity index; X-PDAP, X-prolyl dipeptidyl aminopeptidase.

Involved in various processes such as obesity, tumor growth, graft rejection and allergic phenomena, DPP-IV also contributes to maintain glucose homeostasis by activating insulin [8]. Recently, inhibitors of the enzyme have been used to treat diabetes mellitus [8].

X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis* (PepX) is the type example of the S15 family [1]. The structure of this dimeric enzyme was solved at 2.2 Å [9]. This enzyme is also found in pathogenic *Streptococci*, such as *Streptococcus gordonii*, responsible for bacterial endocarditis [10], or *Streptococcus agalactiae*, the leading cause of neonatal sepsis and meningitis, where PepX was identified as a virulence factor [11,12]. *Streptococcus pneumoniae* and *Streptococcus pyogenes*, implicated in several serious diseases, also possess a PepX-encoding gene, as detected by systematic genomic sequencing [13,14]. While PepX from beneficial bacteria are probably involved in the degradation of milk caseins, proline-rich proteins, the *in vivo* function of this enzyme in *Lactococci* and *Lactobacilli* is not fully understood [15,16]. As PepX has been proposed to play an important role in the virulence [10–12] of pathogens, it could be of great interest to selectively inhibit this enzyme, to stop or at least slow down the infectious process of some *Streptococci*.

We compared the sequences and the 3D structures of prokaryotic PepX and eukaryotic DPP-IV. This study is the first evaluating the effects of evolution on these distant enzymes sharing the same activity. Based on this comparison, we tested the effects of some inhibitors of DPP-IV on the activity of PepX, for which no specific inhibitor is known. We thus found that PepX was highly sensitive to valine-pyrrolidide, with a K_I close to that reported for DPP-IV inhibition. We then used the structure of PepX from *L. lactis* as a template for computer-based homology modeling of PepX from the pathogenic *S. gordonii*. Finally, we proceeded to docking simulations of valine-pyrrolidide into the active site of PepX from *L. lactis* and PepX from *S. gordonii*. The identification of key-residues gives new insights toward a rational drug design against PepX from *Streptococci*. These results could have applications in human health, giving new perspectives to struggle against pathogens.

Results

Comparison of the sequences and structures of the prokaryotic PepX and the eukaryotic DPP-IV

PepX from *L. lactis* is a homodimeric enzyme composed of an α/β hydrolase [17] catalytic domain

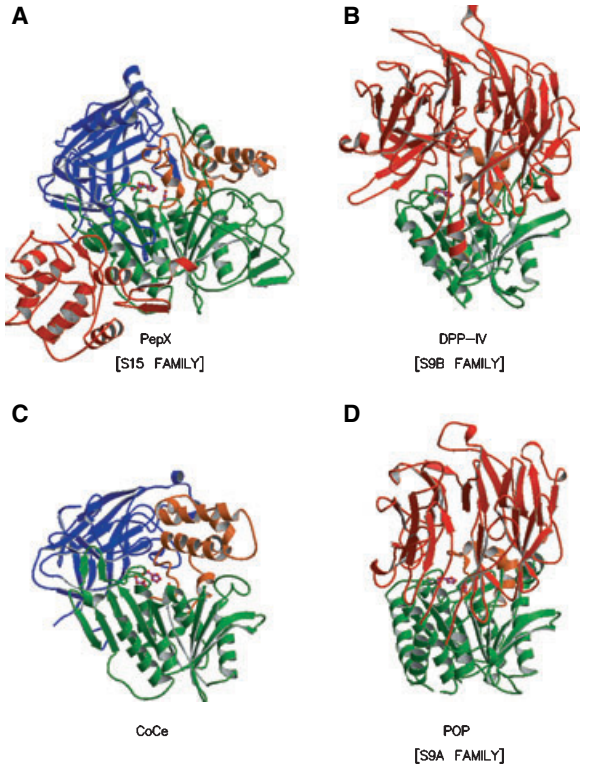

Fig. 1. Structures of enzymes of the clan SC. The N-terminal domain of each enzyme is colored red, the catalytic domain is green, the helical and smallest domain is orange, the C-terminal domain (only present in PepX) is blue, and the catalytic triad is magenta with a ball-and-stick representation. (A) PepX from *L. lactis* ssp. *Cremoris* (1LNS), type example of the S15 family. This is the only enzyme of the clan SC presenting four domains including a large C-terminal domain. (B) Soluble secreted form of human DPP-IV (1N1M), type example of the S9B subfamily. (C) Bacterial cocaine esterase (1JU3). (D) Muscle POP, porcine prolyl oligopeptidase (1QFS), type example of the S9A subfamily.

(Fig. 1A, green) covalently bound to an N-terminus and a C-terminal β-sandwiched domain (respectively in red and blue in Fig. 1A) [9], whereas DPP-IV consists of an eight-blade β propeller domain and a C-terminal α/β hydrolase domain (Fig. 1B) [3,18–22] which forms the catalytic domain with two small α-helices of the N-terminal sequence. When compared, the structures of PepX and DPP-IV do not show any homology except for their catalytic portion, where it is possible to superimpose only 140 Cα atoms (around 18% of PepX and DPP-IV Cα positions) with a root mean square deviation (rmsd) less than 1.0 Å.

The small helical domain, present in all enzymes of the clan SC and carrying residues important for the enzyme specificity, seems also to be very specific to each enzyme as it is clearly of a different size,

sequence and frankly different orientations with respect to the common α/β hydrolase domain (Fig. 1, orange). Although the two X-PDAP enzymes are of the same size (763 residues for PepX and 766 residues for DPP-IV), the closest structure to DPP-IV is prolyl oligopeptidase (POP; Fig. 1D) [23], an endopeptidase specifically cleaving after proline residues and belonging to the same family S9 but having less than 20% residues in common, whereas the closest structure to PepX is the bacterial cocaine esterase (Fig. 1C) [24], which is only composed of two domains. The funnel in the centre of the β8-propeller domain responsible for substrate selection in DPP-IV [3,18–22] has no equivalent in PepX.

These differences are reflected in the low homology between the sequences of the two X-PDAP enzymes even when the comparison is restricted only to their catalytic domains (17.8% identity). Only three conserved sequences can notably be distinguished between the sequences of PepX and DPP-IV: a sequence NxxxAxxGxSYxG around the active serine; a sequence LxxHGxxDxNVxxxxQxxxxxKAL around the active aspartic acid; and a sequence AxAxx-SxWxxY in the helical domain (where 'x' represents any amino acid).

Comparison of the active and specificity sites in the two X-PDAP enzymes

As emphasized, the differences in the overall structures and substrate selection between the bacterial PepX and its mammalian counterpart DPP-IV are both numerous and important. Nevertheless, about two-thirds of the residues involved in catalysis superimpose with good accuracy in both structures while the rest are replaced by more or less equivalent residues positioned in the same locations of the specificity sites (Fig. 2A, Table 1).

Beside the two catalytic triads, which correspond perfectly, the residues Tyr662 in DPP-IV and Tyr380 in PepX (Fig. 2A, Pos1) that ensure the stacking with the Pro residue of the substrate in the N^{-1} position [3,9,18–22], superimpose almost exactly. This is also the case for the main chain of the tyrosine residues, Tyr349 in PepX and Tyr631 in DPP-IV (Fig. 2A, Oxa1 site), positioned immediately after the catalytic serine in both enzymes, and involved in the oxyanion hole. Their aromatic rings are held perpendicular to each other with the OH group pointing in the same direction. Moreover, one of the very important glutamic acid residues, Glu206, responsible with Glu205 for

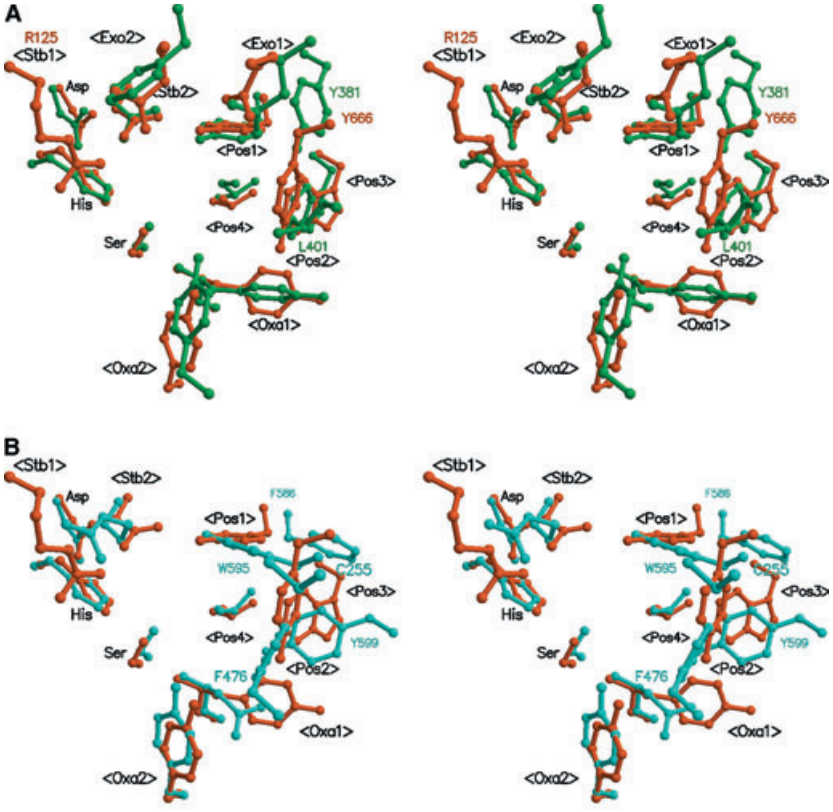

Fig. 2. The X-PDAP signature. This figure shows the LSQ superimposition of the most similar residues involved in the specificity of the enzymes compared. According to the descriptions of the active sites of DPP-IV [3,18–22] and PepX [9], the residues involved in the positioning of the substrate proline are labelled <Pos1> to <Pos4>, the residues involved in the oxyanion hole are labelled <Oxa1> and <Oxa2>, and residues responsible for the exopeptidase activity are labelled <Exo1> and <Exo2>. Finally, the residues postulated to stabilize the substrate when it is positioned in the specificity pocket are labelled <Stb1> and <Stb2>. (A) X-PDAP signature resulting from the comparison of the active and specificity sites of the two X-PDAP enzymes, the bacterial PepX (green) and the human DPP-IV (orange). (B) Superposition of the active and specificity sites of the S9 enzymes porcine POP (cyan) and human DPP-IV (orange). The stacking with the substrate proline is ensured by Trp595 in POP instead of Tyr380 in PepX or TYR 666 in DPP-IV.

Table 1. Equivalent residues in compared enzymes of the clan SC. Labels used are the same as in Fig. 2. Apart from the conserved catalytic triad, a group of six to 11 residues, depending on the enzyme, are involved in the specificity. These can be divided into four subgroups: the residues involved in positioning the substrate in the active site (labelled Pos1 to Pos4); residues involved in the stabilization of the substrate in the active site (Stb1 and Stb2); those forming the oxyanion hole (Oxa1 and Oxa2); and those responsible for the exopeptidase specificity (Exo1 and Exo2). For each subsite, residues with main chains and side chains superimposing well (rms fit of around 1 Å) are shown in bold, while those present in the same region with main chain or side chain superimposing relatively well are in normal font. Residues that do not superpose at the subsite but are present in the same position are underlined. Absence of a residue is indicated '–'. Glu204 and Glu232, responsible for the SPAP exopeptidase activity, are not superimposable with Exo1 or Exo2 subsites.

Label	PepX (family S15)	DPP-IV (family S9B)	POP (family S9A)	SPAP (family S33)	CBPY (family S10)
Catalytic triad					
Ser	**S348**	**S630**	**S554**	**S113**	**S146**
Asp	**D468**	**D708**	**D641**	**D268**	**D338**
His	**H498**	**H740**	**H680**	**H296**	**H397**
Residues implicated in positioning of the substrate proline in the active site					
Pos1	**Y380**	**Y662**	W595	–	–
Pos2	L401	Y666	F476	E232	W312
Pos3	**W377**	**W659**	–	L141	–
Pos4	**I374**	**V656**	V580	F139[a]	L178
Residues stabilizing the binding of the substrate in the specificity pocket					
Stb1	–	R125	R643	–	Y256
Stb2	**N470**	**N710**	R643	A270	**I340**
Oxyanion hole					
Oxa1	Y349	Y631	N555	W114	Y147
Oxa2	Y210	Y547	Y473	–	–
Residues responsible for the exopeptidase activity					
Exo1	F393	E205	–	–	–
Exo2	E396	E206	–	–	–
Other residues postulated to play a role in enzyme specificity					
Other	**V471**	**V711**	**V644**	C271	C341

[a] F139 of the SPAP enzyme occupies Pos4 but plays the same role as the Pos1 subsite (stacking with the substrate proline).

exopeptidase specificity but also contributing to the precise positioning of the substrate proline in DPP-IV [3,18–22], superimposes with good agreement on Glu396 of PepX (Fig. 2A, Exo1 subsite), which is also supposed to play a role in specificity [9]. For Glu205, crucial for inhibition of DPP-IV [3], the superimposition involves a phenylalanine, Phe393, in the same region of PepX (Exo2 subsite). As in the case of Glu205 and Glu206 in DPP-IV, their homologues in PepX (Phe393 and Glu396, respectively) are close in the sequence, belonging to a loop detached from the rest of the molecule. Nevertheless, these structures come from entirely different domains: the helical domain in the case of PepX (Fig. 1A, orange) but a small helix in the N-terminal propeller domain in DPP-IV (Fig. 1B, red).

In addition, Trp377 in PepX is exactly superimposable on Trp659 in the catalytic site of DPP-IV (Fig. 2A, Pos3 subsite). This residue seems to be very specific for the X-PDAP enzymes as it is not present in other members of clan SC such as POP (Fig. 2B). We hypothesize that this tryptophan residue could contribute to the positioning of the substrate. Good correspondence can also be found for the residues Val470 and Asn471 in PepX that superimpose with residues Val710 and Asn711, respectively, in DPP-IV (Stb1 subsite and the next residue). To a lesser extent this is also the case for Tyr210 in PepX and Tyr547 in DPP-IV present in the same region of the active site, with the hydroxyl pointing in the same direction (Fig. 2A, Oxa2).

Finally, Arg125 (Fig. 2A, Stb1 subsite), judged to be important for DPP-IV activity [3], has no equivalent in PepX, the space occupied by this bulky and charged residue being totally empty in the bacterial enzyme.

The residues and subdomains involved in the substrate specificities of other members of the clan SC whose structures are known show drastic differences (Table 1). Thus, comparing precisely the specificity sites of DPP-IV and POP (Fig. 2B), the enzyme that most resembles DPP-IV (Fig. 1), the similarities appear curiously lower than in the case of the superimposition of PepX and DPP-IV catalytic sites. Homologies essentially concern the Pos1, Oxa1 and Oxa2 subsites. The stacking with the substrate proline is ensured by Trp595 [23], superimposing only partially with the

tyrosine of the Pos1 site (Fig. 2B). The structural differences observed between the two peptidases could be connected with their different specificities.

In vitro inhibition of PepX from *L. lactis*

Contrary to DPP-IV, no specific inhibitor is known for the bacterial PepX. One of the immediate consequences of the structural similarities between the two active sites is to search for putative inhibitors of PepX, using the knowledge of DPP-IV inhibition as a starting place. We thus decided to test the effect of some efficient inhibitors of DPP-IV on PepX activity.

PepX was highly sensitive to the DPP-IV specific inhibitor valine-pyrrolidide, as shown by 50% residual activity when using a concentration of 30 μM (IC_{50}) and by a K_I of 9.3 μM (Table 2, Fig. 3A). Comparing these data with the inhibition experiments for DPP-IV, the concentration of valine-pyrrolidide inhibiting PepX is 7.5 times the concentration inhibiting DPP-IV ($IC_{50} = 4$ μM, Table 2), but the K_I values are close (Table 2). It should be noted that with a K_I of the micromolar range, more precisely 2 μM, the DPP-IV inhibitor valine-pyrrolidide represents an effective glucose-lowering compound *in vivo* [25].

The classical inhibitors of DPP-IV diprotin A (Ile-Pro-Ile) and diprotin B (Val-Pro-Leu) have also been tested on PepX activity. Both tripeptides had a lesser effect on inhibiting PepX activity. The IC_{50} was found to be 260 μM for diprotin A and 600 μM for diprotin B, diprotin A being thus nine times less efficient than valine-pyrrolidide (Table 2, Fig. 3B) whereas diprotin B is 20 times less efficient than valine-pyrrolidide.

The results obtained for diprotin A and diprotin B could be partly explained by the size of these compounds, fitting less well to some details of the structure of PepX than smaller ligands such as valine-pyrrolidide and requiring a more specific binding site that extends better. In consequence, as observed here, diprotin A or diprotin B inhibit PepX activity less efficiently.

Table 2. Inhibition experiments realized with PepX from *L. lactis*. All experiments were carried out under the same conditions of pH 8.5 and at 37 °C. IC_{50} and K_I values for PepX were obtained from a graphical analysis of the results of the experiments (Fig. 3). SI, selectivity index (PepX IC_{50}/DPP-IV IC_{50}).

Compound	PepX IC_{50} (μM)	K_I (μM)	DPP-IV IC_{50} (μM)	K_I (μM)	SI
Valine-pyrrolidide	30	9.3	4[a]	2[b]	7.5
Diprotin A	260	71.5	8[c]	4.6[d]	32.5
Diprotin B	600	118.45			

[a] [28], [b] [3], [c] [36], [d] [20].

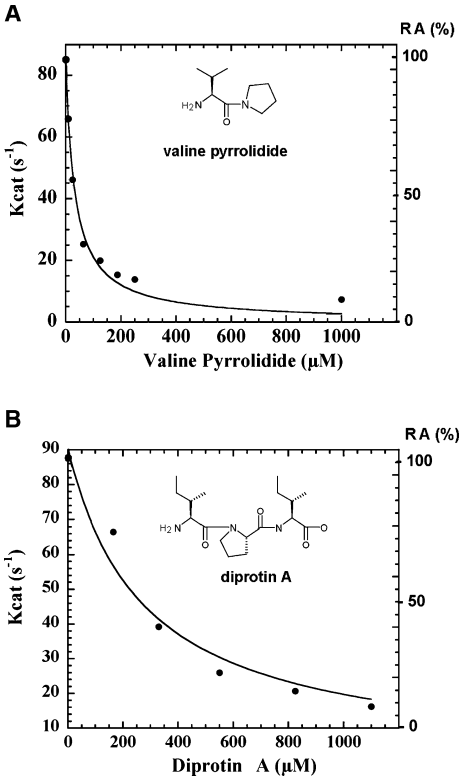

Fig. 3. Inhibition experiments. (A) Inhibition of PepX by the DPP-IV inhibitor valine-pyrrolidide. Values measured for k_{cat} are reported in s^{-1} for different concentrations of valine-pyrrolidide. The right axis shows the residual activity (RA, %). (B) Inhibition of PepX by the DPP-IV classical inhibitor diprotin A. Values measured for k_{cat} are reported in s^{-1} for different concentrations of diprotin A. The right axis shows the residual activity (RA, %). IC_{50} and K_I values for both inhibitors have been obtained from analysis of these graphics.

Docking simulations on X-PDAP enzymes

The results obtained for valine-pyrrolidide encouraged us to model the docking of the inhibitor into the PepX active site using the Lamarckian genetic algorithm of AUTODOCK 3.0 [26]. We started from the reference of the crystal structure of DPP-IV complexed with the valine-pyrrolidide inhibitor in order to make profitable comparisons.

We first carried out docking tests of valine-pyrrolidide in the DPP-IV active site to validate the efficiency of the method. As a result, the positions and conformations of the lowest binding energy docked solutions were found with a frequency of 94% and an rmsd of 1.36 Å, compared with the true crystal structures that served as references (Table 3). The low rmsd between

Table 3. Docking simulations of valine-pyrrolidide in the X-PDAP enzymes. For each simulation, 150 runs were carried out. The robustness is given by the frequency of observation of the cluster among all solutions.

Enzyme	Source	Cluster rank	Robustness	Mean docked energy (kCal·mol^{-1})	KI[b] (μM)	Reference rmsd[c] (Å)
DPP-IV[a]	Human	1	91%	−8.4	1.50	1.36
PepX	*L. lactis*	1	83%	−7.2	9.60	1.50
PepX	*S. gordonii*	2	80%	−7.5	7.10	1.70

[a] Tests carried out on DPP-IV to validate the computational protocol of AUTODOCK. [b] Mean of values calculated by AUTODOCK 3.0. [c] Mean of the rmsd calculated between the solutions of AUTODOCK 3.0 and the reference in the complex ligand-DPP-IV crystal structure.

the docked inhibitor and the reference indicated that the method was suitable to study real interactions between substrate and enzyme.

Valine-pyrrolidide was then docked into the active site of PepX from *L. lactis*. The results are presented in Table 3 and Fig. 4A. For the cluster of lowest energy, both the position and the conformation of the ligand are close to those observed in the 3D structure of the complex between DPP-IV and valine-pyrrolidide [3], serving as reference for the rmsd calculation. The rmsd was 1.5 Å for this robust cluster obtained with a frequency of 83% (Table 3). The other solutions, also found in docking calculations with DPP-IV, were too far from the original position to be considered as likely. The value of 9.6 μM computed by AUTODOCK for the K_I of the solution of the lowest binding energy (Table 3) is remarkably close to the experimental value of 9.3 μM found for K_I in inhibition tests concerning PepX (Table 2).

The interactions of the bound ligand with the residues of the specificity site (Fig. 4A) reveal good stacking between the pyrrolidine ring of the drug and the side chain of Tyr380 (Pos1 subsite), even slightly better than in the true crystallographic complex formed between the drug and DPP-IV and involving Tyr666. As expected from the resemblances between the two active sites, the OE1 atom of Glu396 forms a hydrogen bond with the amino terminus of valine-pyrrolidide, but the additional bond observed in DPP-IV between the N-terminus of valine-pyrrolidide and the Glu313 side chain is absent. The PepX active site lacks an equivalent residue to Arg125 from DPP-IV [3,18–22],

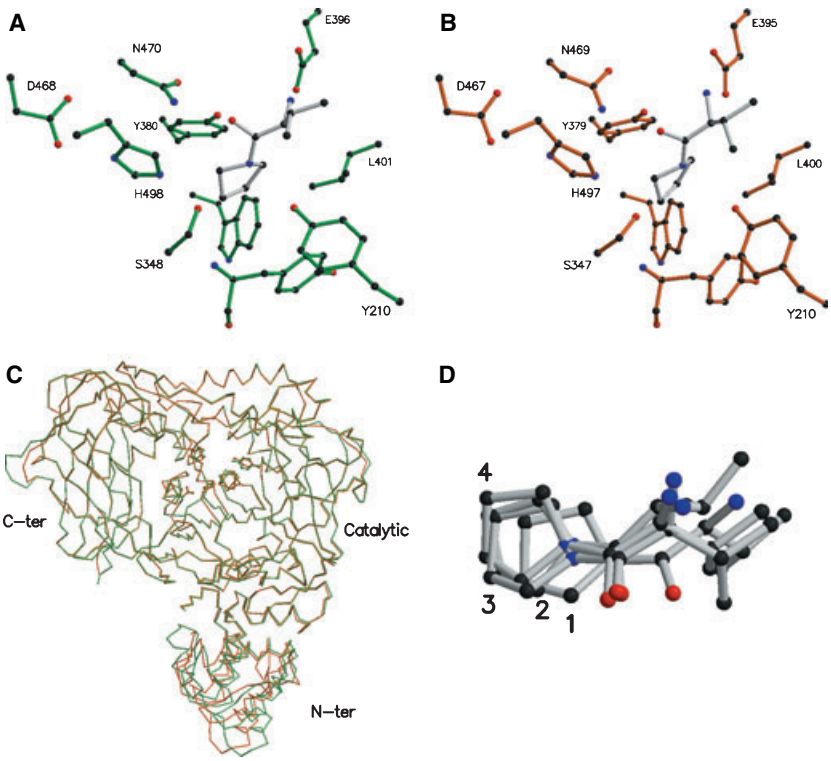

Fig. 4. Docking simulations. Valine-pyrrolidide docked into the active site of: (A) the crystalline structure of PepX from *L. lactis*; (B) the modeled structure of PepX from *S. gordonii*. (C) Structure of PepX from *S. gordonii* (orange) computed by MODELLER software, using homology modeling approach and superimposed with the crystalline structure of PepX from *L. lactis* (green) that served as the template structure in the modeling. The rmsd between the two structure is 0.45 Å (D) Compared positions and conformations of the ligand valine-pyrrolidide. **1**, Complexed with DPP-IV in the X-ray crystallographic structure (considered to be the reference); **2**, docked into the active site of the human DPP-IV; **3**, docked into the active site of the crystalline structure of PepX from *L. lactis*; and **4**, docked into the active site of the modeled structure of PepX from *S. gordonii*.

but this does not seem to be essential for the binding of valine-pyrrolidide in the PepX active site. Here a unique bond is observed between the carboxyl group of the ligand and the NH_2 of Asn470 (equivalent to Asn710 in DPP-IV). Another potential bond can be proposed between the OH atom of Tyr380 and the carbonyl group of the ligand. Finally, the valine side chain of the ligand, pointing towards the active site cavity, shows only low hydrophobic interaction with the side chain of Leu401. These data thus reveal that PepX and DPP-IV enzymes bind small drugs with comparable interactions.

Discussion

Comparing the evolutionarily distant bacterial PepX and human DPP-IV led to the conclusion that most of the residues implicated in X-PDAP activity are conserved in the same position in both specificity sites (Fig. 2A), despite very different domains flanking the catalytic domain, different dimer organization and substrate selectivity processes. These resemblances are characteristic of X-PDAP activity, as a broader comparison involving all of the known structures of enzymes belonging to the clan SC has shown that most of these key residues are only present in X-PDAP enzymes (Table 1, Fig. 2). This study reveals the existence of a structural signature of X-PDAP specificity, crossing the subdivisions among peptidase families. This particular spatial arrangement is probably the result of divergent evolution that retained an efficient site for the release with high specificity of a dipeptide containing a proline from a polypeptide.

The residues postulated to stabilize the substrate in the active site of DPP-IV are Arg125 and Asn710. The arginine residue is not necessary for catalysis in PepX, where it is replaced by an empty space that constitutes the most important difference between the bacterial and the mammalian active sites. Another difference concerns the residues responsible for the exopeptidase specificity, namely the two glutamic acids Glu205 and Glu206 in DPP-IV but Glu396 and Phe393 in PepX. This observation emphasizes the important role of Glu396 as already described [9], and reveals a function for Phe393 in PepX. These positions have been maintained in the signature despite structural rearrangements throughout evolution.

The existence of the X-PDAP signature led us to test inhibitors of DPP-IV on PepX activity for which no specific inhibitor is known. The inhibitors of DPP-IV chosen were those of the smallest size. The resemblances between the two active sites are well confirmed by the results of inhibition tests with valine-pyrrolidide. This compound has been shown here to be an inhibitor for PepX and docking simulations revealed that the active site of PepX can accommodate such compounds.

PepX from *S. gordonii* is a potential virulence factor in bacterial endocarditis [10]. The amino acid alignment of PepX from *L. lactis* and PepX from *S. gordonii* gave a sequence identity of 48% and 65% homology. This enabled us to obtain a realistic model of PepX from *S. gordonii* (Fig. 4B,C), with the homology modelling approach of the MODELLER software [27], using the structure of PepX from *L. lactis* as protein template. As shown in Fig. 4C, no important differences have been observed between the two PepX structures. Moreover, the two active sites are highly conserved (Fig. 4A,B). Docking simulations done with this modelled PepX gave similar results to the computations done with the enzyme from *L. lactis* (Table 3, Fig. 4B). The ligand valine-pyrrolidide presents the same interesting interactions, which suggest that it could also inhibit the streptococcal PepX enzyme.

As DPP-IV is involved in a great variety of physiological processes it is important to avoid adverse reactions. Thus a nonspecific inhibition directed against the X-PDAP of pathogens would probably affect the activity of the mammalian enzymes, leading to harmful consequences for human health. Taking advantage of the structural differences and similarities between the mammalian and bacterial specificity sites, potential targets can be selected designing inhibitors acting specifically on PepX of pathogens but not on DPP-IV of the infected host.

As revealed by the results of our inhibition tests with diprotin A or diprotin B, we have found cases for which inhibitors are much more adapted to one enzyme (DPP-IV) than to the other one (PepX). Conversely, it should be possible to find drugs that inhibit PepX more efficiently than DPP-IV. Recently, compounds have been found to be more effective on DPP-II than on DPP-IV activity, with a high selectivity enabling differentiation between DPP-II and DPP-IV in biological systems [28].

To obtain compounds inhibiting PepX more efficiently than DPP-IV, advantage could be taken of the large empty space present in the PepX active site in place of the important Arg125 residue in DPP-IV (Fig. 2A). A pyrrolidide derivative that fills, at least partially, this free space in PepX with a close adaptation to the rest of the active site would probably be unable to enter into the DPP-IV active site due to steric hindrance or to the positive charge of Arg125. Another strategy would be to exploit the difference of occupancy between Phe393 in PepX and Glu205 in DPP-IV (Fig. 2A, Exo2 site). The Tyr381-Glu396 loop

of the helical domain, only present in PepX, could also be a target to specifically block the bacterial enzyme.

The work presented here could be usefully employed in the research of treatments of particularly severe diseases involving *Streptococci* (*S. gordonii, Streptococcus pneumoniae*, *Streptococcus agalactatiae*, *Bacillus anthracis*), avoiding the damage caused by PepX to host tissues during pathology.

Experimental procedures

Multiple sequence alignment and superimposition of structures

The sequences and structures of the following enzymes were compared: CBPY, carboxypeptidase Y from *Saccharomyces cerevisiae* (1CPY) [29], belonging to the S10 family; TricornF1, tricorn-interacting factor F1 from *Thermoplasma acidophilum* (1 MU0) [30], belonging to the S33 family; PIP, prolyl iminopeptidase from *Xanthomonas campestris* (1AZW) [31], belonging to the S33 family; SPAP, prolyl iminopeptidase from *Serratia marcescens* (1QTR) [32], belonging to the S33 family; POP, muscle porcine prolyl oligopeptidase (1QFS) [23], enzyme of the S9A subfamily; h_DPP-IV, the secreted part of human DPP-IV (1 N1M) [3], enzyme of the S9B subfamily and PepX from *L. lactis* ssp. *cremoris* (1LNS) [9], enzyme of the S15 family. Each enzyme represents one of the functions associated with families of the clan SC [1] for which at least one member has a known structure. The sequences, extracted from the PDB (Protein Data Bank) files, have been aligned based on their three dimensional structure information using the STAMP software [33]. Only catalytic domains of each enzyme have been considered as the members of the clan SC show very different domains flanking their catalytic domain.

The program O [34] was used for precise graphical display and structural least squares (LSQ) calculation superimpositions. An LSQ calculation starting from the catalytic triads led to the superimposition of the common residues defining the active and specificity sites of DPP-IV and PepX, including some residues around them. As a result, 140 atoms superimposed with a mean rms fit of 1.0 Å. The α/β hydrolase folds of the compared enzymes were also in good agreement. The same procedure was repeated for the coupling of the structures of DPP-IV and POP, leading to 96 atoms superimposing with an rms fit of 0.8 Å. These calculations allowed the construction of Figs 1 and 2 and Table 1.

Enzyme purification specific activity and inhibition studies

H-Ala-Pro-*para*-nitroanilide (Ala-Pro-pNA) was purchased from Sigma (St Louis, MO, USA) and valine-pyrrolidide was synthesized by Neosystem (Strasbourg, France). Diprotin A and diprotin B were from Bachem (Bubendorf, Switzerland). All products were of the best analytical grade. PepX from *L. lactis* was produced starting from the published protocol [35]; 10 mg of highly purified enzyme was obtained after two steps of HPLC chromatography. The specific activity measured from this preparation was 1.4 μmol of *para*-nitroanilide released per second and per mg of enzyme; pNA was detected at 410 nm ($\varepsilon = 9600\ \text{M}^{-1}$).The kinetic parameters for PepX were determined at 37 °C using the substrate Ala-Pro-pNA at concentrations ranging from 50 μM to 1 mM with an enzyme concentration of 0.00566 μM, in 50 mM Tris/HCl pH 8.5 (pH optimum). An Eadie–Hofstee plot analysis gave 80 μM for K_M and 155 s^{-1} for k_{cat}. Inhibition tests were done with three classical compounds known to inhibit the prokaryotic DPP-IV, diprotin A, diprotin B and valine-pyrrolidide, at concentrations varying from 10 μM to 2.5 mM. The experiments were carried out under the same conditions for all inhibitors tested. Reactions occurred in 1 mL volumes with 50 mM Tris/HCl pH 8.5 and temperature was 37 °C. The hydrolysis was stopped with 30% (v/v) acetic acid and the absorbance of the solution was measured at 410 nm to evaluate the residual activity for each tested drug.

Molecular modelling of PepX from *S. gordonii* and docking simulations

The sequence of PepX from *S. gordonii* was obtained from the Swiss-Prot database and its 3D structure was constructed using the homology modelling method of the MODELLER software [27] starting from the atomic coordinates of PepX from *L. lactis* (1LNS) as template protein and the CLUSTALW alignment between the two enzyme sequences.

The AUTODOCK 3.0 package [26] was used to perform the automated molecular docking of the ligand valine-pyrrolidide into DPP-IV and PepX active sites. The structure of DPP-IV complexed with valine-pyrrolidide [3] and the structure of PepX [9] were downloaded from the PDB. The enzymes and ligand coordinates were saved as separate PDB files. For each enzyme structure polar hydrogens were added, Kollman united-atom charges were assigned and atomic solvation parameters were added. Hydrogen atoms were also added and Kollman united-atom charges assigned for the ligands before the nonpolar hydrogens were removed and their partial charges added to the bonded carbon atom. The internal degrees of freedom and torsions were finally set for each inhibitor. All preparations were done with ADT, the AUTODOCK tool graphical interface [26]. Interaction grids of $20 \times 20 \times 20$ Å centred in the active site and separated by 0.375 Å for all types of atom were then prepared with the AUTOGRID 3 utility for PepX and DPP-IV molecules. Docking runs were performed using the Lamarckian generic algorithm, described as being the most efficient [26]. The population size was set to 100 to ensure that the conformational space was exhaustively searched

and a total of 150 docking runs were performed for each simulation. A cluster analysis was finally carried out on the results using the crystallographic coordinates of the inhibitor as the reference structure and a tolerance of 1.0 Å rmsd. The complexes of lowest interacting energy solutions were selected as the best docked structures.

To verify that the docking protocol was suitable for such enzymes and to apply the correct parameters, calculations were first realized with DPP-IV and then performed with PepX using the same parameters.

Acknowledgements

We warmly thank Dr Babette Weksler for critical reading and comments on this manuscript.

References

1 Rawlings ND, O'Brien EA & Barrett AJ (2002) MEROPS: the protease database. *Nucleic Acids Res* **30**, 343–346.
2 Mentzel S, Dijkman HBPM, Van Son JPHF, Koene RAP & Assmann KJM (1996) Organ distribution of aminopeptidase A and dipeptidyl peptidase IV in normal mice. *J Histochem Cytochem* **44**, 445–461.
3 Rasmussen HB, Branner S, Wiberg FC & Wagtmann N (2003) Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. *Nat Struct Biol* **10**, 19–25.
4 De Meester I, Korom S, Van Damme J & Sharpe S (1999) CD26, let it cut or cut it down. *Immunol Today* **20**, 367–375.
5 Lambeir AM, Proost P, Durinx C, Bal G, Senten K, Augustyns K, Scharpe S, Van Damme J & De Meester I (2001) Kinetic investigation of chemokine truncation by CD26/dipeptidyl peptidase IV reveals a striking selectivity within the chemokine family. *J Biol Chem* **276**, 29839–29845.
6 Kameoka J, Tanaka T, Nojima Y, Schlossman SF & Morimoto C (1993) Direct association of adenosine deaminase with a T cell activation antigen, CD26. *Science* **261**, 466–469.
7 Ishii T, Ohnuma K, Murakami A, Takasawa N, Kobayashi S, Dang NH, Schlossman SF & Morimoto C (2001) CD26-mediated signaling for T cell activation occurs in lipid rafts through its association with CD45RO. *Proc Natl Acad Sci USA* **98**, 12138–12143.
8 Drucker DJ (2003) Therapeutic potential of dipeptidyl peptidase IV inhibitors for the treatment of type 2 diabetes. *Expert Opin Invest Drugs* **12**, 87–100.
9 Rigolet P, Mechin I, Delage MM & Chich JF (2002) The structural basis for catalysis and specificity of the X-Prolyl dipeptidyl aminopeptidase from *Lactococcus lactis*. *Structure* **10**, 1383–1394.
10 Goldstein JM, Banbula A, Kordula T, Mayo JA & Travis J (2001) Novel extracellular x-prolyl dipeptidyl-peptidase (DPP) from *Streptococcus gordonii* FSS2: an emerging subfamily of viridans streptococcal x-prolyl DPPs. *Infect Immun* **69**, 5494–5501.
11 Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, Msadek T, Zouine M, Couve E, Lalioui L *et al.* (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol Microbiol* **45**, 1499–1513.
12 Jones AL, Knoll KM & Rubens CE (2000) Identification of *Streptococcus agalactiae* virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. *Mol Microbiol* **37**, 1444–1455.
13 Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, Peterson S & Fraser CM (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**, 498–506.
14 Ferretti JJ, McShan WM, Ajdic D, Savic DJ, Savic G, Lyon K & McLaughlin R (2001) Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci USA* **98**, 4658–4663.
15 Mierau I, Kunji ERS, Jeenhouts KJ, Hellendoorn MA, Handrikman AJ, Poolman B, Konings AN, Venema G & Kok J (1996) Multiple-peptidase mutants of *Lactococcus lactis* are severely impaired in their ability to grow in milk. *J Bacteriol* **178**, 2794–2803.
16 Byun T, Kofod L & Blinkovsky A (2001) Synergistic action of an X-prolyl dipeptidyl aminopeptidase and a non-specific aminopeptidase in protein hydrolysis. *J Agric Food Chem* **49**, 2061–2063.
17 Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow FM, Franken S, Harel M, Remington SJ, Silman I, Schrag J. *et al.* (1992) The α/β hydrolase fold. *Prot Engineer* **5**, 197–211.
18 Engel M, Hoffmann T, Wagner L, Wermann M, Heiser U, Kiefersauer R, Huber R, Bode W, Demuth HU & Brandstetter H (2003) The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. *Proc Natl Acad Sci USA* **100**, 5063–5068.
19 Hajime H, Kiyoshi K, Yutaka H, Chiaki F, Hideaki S, Shigeru S, Koji I, Atsushi Y & Ryo S (2003) The structure and function of human dipeptidyl peptidase IV, possessing a unique eight-bladed β-propeller fold. *Biochem Biophys Res Com* **302**, 849–854.
20 Thoma R, Loffler B, Stihle M, Huber W, Ruf A & Hennig M (2003) Structural basis of proline-specific exopeptidase activity as observed in human dipeptidyl peptidase-IV. *Structure (Camb)* **11**, 947–959.
21 Oefner C, D'Arcy A, MacSweeney A, Pierau S, Gardiner R & Dale GE (2003) 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. *Acta Cryst D Biol Crystallogr* **59**, 1206–1212.
22 Aertgeerts KYeS, Tennant MG, Kraus ML, Rogers J, Sang BC, Skene RJ, Webb DR & Prasad GS (2004) Crystal structure of human dipeptidyl peptidase IV in complex with a decapeptide reveals details on substrate specificity and tetrahedral intermediate formation. *Protein Sci* **13**, 412–421.
23 Fülöp V, Böcskei Z & Polgar L (1998) Prolyl oligopeptidase: an unusual β-propeller domain regulates proteolysis. *Cell* **94**, 161–170.
24 Larsen NA, Turner JM, Stevens J, Rosser SJ, Basran A, Lerner RA, Bruce NC & Wilson IA (2002) Crystal structure of a bacterial cocaine esterase. *Nat Struct Biol* **9**, 17–21.
25 Drucker DJ (2001) Development of glucagon-like peptide-1-Based pharmaceuticals as therapeutic agents for the treatment of diabetes. *Current Pharmaceut Design* **7**, 1399–1412.
26 Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK & Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and and empirical binding free energy function. *J Computational Chem* **19**, 1639–1662.
27 Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F & Sali A (2000) Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**, 291–325.
28 Senten K, Van der Veken P, Bal G, De Meester I, Lambeir AM, Scharpe S, Bauvois B, Haemers A & Augustyns K (2002) Development of potent and selective dipeptidyl peptidase II inhibitors. *Bioorg Med Chem Lett* **12**, 2825–2828.
29 Endrizzi JA, Breddam K & Remington SJ (1994) 2.8 Angstrom structure of yeast serine carboxypeptidase. *Biochem* **33**, 11106.
30 Goettig P, Groll M, Kim JS, Huber R & Brandstetter H (2002) Structures of the tricorn-interacting aminopeptidase F1 with different ligands explain its catalytic mechanism. *EMBO J* **21**, 5343–5352.
31 Medrano FJ, Alonso J, Garcia JL, Romeo A, Bode W & Gomis-Rüth FX (1998) Structure of Proline Iminopeptidase from *Xanthomonas campestris pv:* a prototype for the prolyl oligopeptidase family. *EMBO J* **17**, 1–9.
32 Yoshimoto T, Kabashima T, Uchikawa K, Inoue T, Tanaka N, Nakamura KT, Tsuru MT & Ito K (1999) Crystal structure of the prolyl aminopeptidase from. *Serratia Marcescens J Biochem* **126**, 559–565.
33 Russel RB & Barton GJ (1992) Multiple protein sequence alignment from tertiary structure comparison. *Proteins Struct Funct Genet* **14**, 309–323.
34 Jones TA & Kjeldgaard M (1995) *O – the Manual.* Uppsala University, Uppsala.
35 Chich JF, Rigolet P, Nardi M, Gripon JC, Ribadeau-Dumas B & Brunie S (1995) Purification, crystallization, and preliminary X-ray analysis of PepX, an X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis*. *Proteins* **23**, 278–281.
36 Russell JS, Chi H, Lantry LE, Stephens RE & Ward PE (1996) Substance P and neurokinin A metabolism by cultured human skeletal muscle myocytes and fibroblasts. *Peptides* **17**, 1397–1403.