Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases

Shu Y. QI, Pierre J. RIVIERE, Jerzy TROJNAR, Jean-Louis JUNIEN and Karen O. AKINSANYA[1]

Molecular and Cellular Biology Department, Ferring Research Institute, 3550 General Atomics Court, San Diego, CA 92121-1122, U.S.A.

Two dipeptidyl peptidase IV (DPPIV, DPP4)-related proteins, DPP8 and DPP9, have been identified recently [Abbott, Yu, Woollatt, Sutherland, McCaughan, and Gorrell (2000) Eur. J. Biochem. **267**, 6140–6150; Olsen and Wagtmann (2002) Gene **299**, 185–193; Qi, Akinsanya, Riviere, and Junien (2002) Patent application WO0231134]. In the present study, we describe the cloning of DPP10, a novel 796-amino-acid protein, with significant sequence identity to DPP4 (32 %) and DPP6 (51 %) respectively. We propose that DPP10 is a new member of the S9B serine proteases subfamily. The DPP10 gene is located on the long arm of chromosome 2 (2q12.3-2q14.2), close to the DPP4 (2q24.3) and FAP (2q23) genes. The active-site serine residue is replaced by a glycine residue in DPP10, resulting in the loss of DPP activity. The serine residue is also replaced in DPP6, which lacks peptidase activity. DPP8 and DPP9 share an identical active site with DPP4 (Gly-Trp-Ser-Tyr-Gly). In contrast with the previous results suggesting that DPP9 is inactive, we show that DPP9 is a DPP, hydrolysing Ala-Pro-(7-amino-4-methylcoumarin) with similar pH-specificity and protease-inhibitor-sensitivity to those of DPP4 and DPP8. Northern-blot analysis shows that whereas DPP8 and DPP9 are widely expressed, DPP10 is expressed mainly in the brain and pancreas. DPP6, which has the highest amino acid identity with DPP10, has been shown previously [Nadal, Ozaita, Amarillo, de Miera, Ma, Mo, Goldberg, Misumi, Ikehara, Neubert et al. (2003) Neuron **37**, 449–461] to associate with A-type K^+ channel subunits, modulating their transport and function in somatodendritic compartments of neurons. It is possible that DPP10 is involved in similar functions in the brain. Elucidation of the physiological or pathophysiological role of DPP8, DPP9 and DPP10 and characterization of their structure–function relationships will add impetus to the development of inhibitor molecules for pharmacological or therapeutic use.

Key words: dipeptidyl peptidase IV (DPPIV, DPP4), DPP8, DPP9, DPP10, S9B subfamily.

INTRODUCTION

Proteases are involved in a number of physiological processes, such as inflammation, infection, fertilization, coagulation, allergic reactions, cell growth and death, tumour growth and bone remodelling. Pathophysiological events involving proteases include membrane-bound and extracellular polypeptide proteolysis, tumour invasion, tissue remodelling and cell migration. As a result, there are a number of successful drugs available in the market, including HIV protease inhibitors, angiotensin-converting-enzyme inhibitors and a few others which modulate protease functions [1]. Of the 400 known human proteases, 60 are under investigation as potential drug targets [2]. Amongst these hundreds of identified proteases, only a few proline-specific proteases have been described. Proline adjacent bonds are relatively resistant to breakdown by general oligopeptidases, and general amino-, di- or tri-peptidases due to the unique conformational constraints imposed by proline motifs in a peptide chain. Proline residues are conserved in many biologically active proteins and peptides (e.g. cytokines, growth factors, G-protein-coupled receptors, V3 loops of the HIV envelope glycoprotein gp 120 and neuroendocrine and vasoactive peptides).

Specialized proteases that cleave proline adjacent bonds include metallopeptidases activated by Mn^{2+} (aminopeptidase P, caboxypeptidase P and prolidase) and three serine proteases [dipeptidyl peptidase (DPP) II, DPPIV (EC 3.4.14.5) and prolyl oligopeptidase (POP; EC 3.4.21.26)]. DPPIV (DPP4) is an ectoenzyme that liberates Xaa-Pro and less efficiently Xaa-Ala dipeptides from the N-terminus of a variety of small proteins. Thus the proteolytic function of DPP4 is of great physiological importance for the activation and inactivation of extracellular endocrine, paracrine and autocrine peptides [3]. DPP4 is widely distributed in mammalian tissues and is expressed highly in the kidney cortex, intestinal epithelium, liver hepatocytes, pancreatic epithelial cells and placenta [4,5]. In the human immune system, DPP4 is synonymous with the cell-surface antigen CD26, expressed by activated lymphocytes (T-, B- and natural killer cells) and is involved in the regulation of chemokine function, and may play an important role in HIV infection [6].

The mechanistic classes of proteases (serine, cysteine, aspartate and metalloproteases) are divided into a series of subfamilies and clans [3]. Peptidases and proteases that exploit a serine residue in their catalytic activity are ubiquitous, being found in viruses, bacteria and eukaryotes. Over 20 subfamilies (denoted S1–S27) of serine protease genes have been identified and grouped into six clans (SA, SB, SC, SE, SF and SG) on the basis of structural similarity, evolutionary origins and other functional evidence [4–6]. DPP4 belongs to the POPs, a relatively new

Abbreviations used: AMC, 7-amino-4-methylcoumarin; DPP, dipeptidyl peptidase; DASH, DPPIV activity and/or structure homologues; DPRP, DPPIV-related protein; EST, expressed sequence tag; FAP, fibroblast activation protein; ORF, open reading frame; PLEES, peptidases, lipases, esterases, epoxide hydrolases or serine hydrolases; pNA, *p*-nitroanilide; POP, prolyl oligopeptidase; RACE, rapid amplification of cDNA ends.

[1] To whom correspondence should be addressed (e-mail karen.akinsanya@ferring.com).

The nucleotide sequence data reported will appear in GenBank® Nucleotide Sequence Database under the accession numbers AY172659, AY172660 and AY172661.

Table 1 List of primers used to clone DPP8, DPP9 and DPP10 and generate expression vectors

Primer name	Primer sense	Oligonucleotide sequence	Targeted DNA	Product
DPP8FL-5′	Sense	5′-CGGTACCATGGCAGCAGCAATGGAAACAG-3′	DPP8	Full length
DPP8FL-3′	Antisense	5′-GGAGCTCGCGGCCGCTCATATCACTTTTAGAGCAGCAAT-3′	DPP8	Full length
DPP9FL-5′	Sense	5′-CAAGCTTACCATGGCCACCACCGGGAC-3′	DPP9	Full length
DPP9FL-3′	Antisense	5′-CGGATCCGCGGCCGCTCAGAGGTATTCCTGTAGAAAG-3′	DPP9	Full length
DPP10FL-5′	Sense	5′-AGAATTCACCATGAACCAAACTGCCA-3′	DPP10	Partial clone
DPP10FL-3′	Antisense	5′-ACTCGAGTTATTCATCTTCTTCTGGTTCCTGT-3′	DPP10	Partial clone
DPP10-5′ RACE-1	Outer	5′-ATCACATATGAAGCAGTATACGAATAATGA-3′	DPP10	5′-sequence
DPP10-5′ RACE-2	Inner	5′-AATCTGGTGAAACTGAATGTCTTGATGCT-3′	DPP10	5′-sequence
pGEM7-DPP8f-2	Antisense	5′-CAAGCTTTATCACTTTTAGAGCAGCAA-3′	DPP8	DPP8 fusion precursor
pGEM7-DPP9f-2	Antisense	5′-CGGATCCAGGTATTCCTGTAGAAAGTG-3′	DPP9	DPP9 fusion precursor
pGEM7-DPP10f-1	Sense	5′-AGCTAGCACCATGAACCAAACTGCCA-3′	DPP10	DPP10 fusion precursor
pGEM7-DPP10f-2	Antisense	5′-AGAATTCATCTTCTTCTGGTTCCTGT-3′	DPP10	DPP10 fusion precursor
Exp-DPP8-1	Sense	5′-CACATTCTTGCTGCATCAGTCA-3′	DPP8	DPP8QPCR
Exp-DPP8-2	Antisense	5′-TTGGGTCATCTTCAGGACTTGA-3′	DPP8	DPP8QPCR
Exp-DPP9-1	Sense	5′-TACGCCGTGGTTGTGATTGA-3′	DPP9	DPP9QPCR
Exp-DPP9-2	Antisense	5′-CCATACTTCTCGGCCACGAA-3′	DPP9	DPP9QPCR
Exp-DPP10-1	Sense	5′-TGCTATTGCTCTGCTGGTGATT-3′	DPP10	DPP10QPCR
Exp-DPP10-2	Antisense	5′-CTGGGCTTAAGAGGATGACTGAC-3′	DPP10	DPP10QPCR

class of serine proteases (family S9B in clan SC), which have evolved by independent convergent evolution [7,8]. A number of DPP4-like proteins, identified and characterized either by molecular cloning and functional studies of expressed proteins or as biochemical activities in tissue extracts, are reported in the literature. Referred to as DPPIV activity and/or structure homologue (DASH) proteins [9], they share similarities with DPP4 enzyme activities or are structurally similar to DPP4 but inactive enzymically. DASH genes with sequence homology to DPP4 include fibroblast activation protein (FAP) [10–14], DPP6/DDPX [15] and the recently cloned DPP8 [16]. Cloned DASH proteins with DPP4-like substrate specificity include DPP7 (human quiescent cell proline dipeptidase, the probable orthologue of DPPII [17,18]) and N-acetylated α-linked acidic dipeptidases I, II and L [19,20]. Purified DASH proteins, with enzymic activities similar to DPP4, include DPP4β [21–23] and attractin/mahogany protein [24]. The cloning, characterization and appropriate classification of additional members of the family of POPs would enable the design of specific inhibitors for use as tools for functional validation and as potential therapeutic agents for pharmaceutical applications. In the present study, we report on the cloning and comparative characterization of three new DASH proteins with sequence similarity to DPP4.

EXPERIMENTAL

Bioinformatics

DPP4-like gene sequences were identified following a tBLASTn [25] search of the National Center for Biotechnology Information GenBank® databases [26] with the full-length sequences of human FAP (GenBank® accession no. U09278) and DPP4 (GenBank® accession no. X60708) using the BLAST family of programs (http://www.ncbi.nlm.nih.gov/BLAST/). A tBLASTn analysis of the GenBank® database using the full-length sequence of human DPP4 and FAP revealed almost 100 hits. After subtraction of the high similarity ($< 1 \times 10^{-20}$) FAP and DPP4 entries, one expressed sequence tag (EST) with high similarity, but not identical with DPP4 and FAP, was identified from three independent clones (AL043338 from testis; AA417787 and AA278625 from germinal centre B-cell). The EST sequences identified were extended initially using the mouse and human high throughput genome sequence databases (http://www.ncbi.nlm.nih.gov/HTGS/). Based on looped pairwise sequence comparison procedures using DPP4 family proteins, several novel sequences were retrieved, including three new genes DPPIV-related proteins-1 (DPRP-1), DPRP-2 and DPRP-3. Alignments among multiple paralogues were performed using dCLUSTALW with comparable outcomes [27]. A simple modular architecture research tool (SMART), which utilizes a hidden Markov model (TMHMM), was used to predict transmembrane-spanning domains [28]. Conserved domains and motifs were identified using the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/) [29]. To construct phylogenetic trees, the human DPP4 protein sequence was used to blast all non-redundant GenBank® coding sequence translations, Protein Data Bank, SwissProt, Protein Information Resource and Protein Research Foundation databases, and generate a list of DPP, DPP4 or aminopeptidase-like sequences. The tree was constructed using a sequence distance method and utilizes the Neighbour-Joining algorithm [30].

Isolation of cDNA clones and sequencing

A DNA fragment encoding a partial DPP10 sequence was amplified from a human hypothalamus cDNA library using 5′ and 3′ sequences of PCR oligonucleotide primers (Table 1). The PCR was performed using Pfu polymerase (Stratagene, San Diego, CA, U.S.A.) for 30 cycles at 94 °C for 30 s, 55 °C for 30 s and 72 °C for 300 s in 5% (v/v) DMSO. To determine the full-length coding sequence of the DPP10, 5′-RACE (rapid amplification of cDNA ends) was performed on human hypothalamus total RNA (ClonTech Laboratories, Palo Alto, CA, U.S.A.) using a 5′-RACE kit (Invitrogen, Carlsbad, CA, U.S.A.). Nested primers specific to AB040925 were used according to the manfacturer's instruction. 5′-RACE PCR products were cloned into pGEMT-easy (Promega, Madison, WI, U.S.A.) and ten independent clones were sequenced. The consensus sequence of the ten reactions was used to design the primer flanking the initiating methionine residue. The full-length DPP10 was subsequently amplified using the primers flanking the initiating methionine residue and the stop codon. DPP8 and DPP9 gene sequences were used to design oligonucleotide primers (Table 1)

for amplification of the full-length cDNA from a commercially available human testis and colon cDNA libraries (ClonTech Laboratories) respectively. Amplified full-length sequences were isolated from a 0.7 % agarose gel using a commercially available kit (GFX PCR DNA and Gel Band Purification Kit; Amersham Pharmacia Biotech, Piscataway, NJ, U.S.A.), and ligated into cloning vector pGEM-7Zf(−) (Promega) and sequenced using an ABI 377 automatic sequencer (Applied Biosystems, Foster City, CA, U.S.A.). The corresponding cloning constructs were designated as pGEM7-DPP8, pGEM7-DPP9 and pGEM7-DPP10 respectively. Mammalian expression constructs were made by restriction enzyme digests of pGEM7-DPP8 (*Kpn*I and *Not*I), pGEM7-DPP9 (*Hin*dIII and *Bam*HI) and pGEM7-DPP10 (*Eco*RI and *Xho*I) to release the full-length DPP8, DPP9 and DPP10 genes respectively. After gel purification, the DNA fragments were inserted into expression vector pcDNA3 (Invitrogen) to make the native expression constructs pcDNA-DPP8, pcDNA-DPP9 and pcDNA-DPP10.

C-terminal Myc-His-tagged DPP8, DPP9 and DPP10 constructs

To enable purification and detection of recombinant DPP8, DPP9 and DPP10 proteins, C-terminal fusion constructs were created. The DNA sequences encoding the truncated DPP8, DPP9 or DPP10 were amplified using pGEM7-DPP8, pGEM7-DPP9 or pGEM7-DPP10 respectively as a template and PCR oligonucleotide primers (Table 1). The amplified sequences were gel-purified and subcloned into pGEM-7Zf(−).

The resulting constructs were designated pGEM7-DPP8f, pGEM7-DPP9f and pGEM7-DPP10f. pGEM7-DPP8f was digested with the restriction enzymes *Xba*I and *Hin*dIII to release the truncated DPP8f gene. The DNA fragment carrying the DPP8f gene was gel-purified and inserted into expression vector pcDNA3.1(−)/myc-His A (Invitrogen) to make the tagged DPP8 expression construct pcDNA-DPP8-MycHis. This vector contains a 63 bp c-Myc tag (EQKLISEEDL) and a His′ tag (HHHHHH) as open reading frames (ORFs) at the 3′-terminus of DPP8 coding region. Similar tagged constructs were made for DPP9 and DPP10. An *Eco*RI and *Bam*HI digest of pGEM7-DPP9f released the truncated DPP9f gene, which was gel-purified and inserted into expression vector pcDNA3.1(−)/myc-His B to make the tagged DPP9 expression construct and designated pcDNA-DPP9-MycHis. pGEM7-DPP10f was digested with the restriction enzymes *Nhe*I and *Eco*RI to release the truncated DPP10f gene, which was gel-purified and inserted into expression vector pcDNA3.1(−)/myc-His B to make the tagged DPP10 expression construct pcDNA-DPP10-MycHis.

Northern-blot and PCR analyses of DPP8, DPP9 and DPP10 mRNA expression

Northern-blot analysis was performed using a human multiple-tissue Northern blot (ClonTech Laboratories) with 2 μg of polyadenylated [poly(A)$^+$] RNA isolated from eight different tissues. Blots were probed with 851-bp product DPP8 (nucleotides 1–851), 9150-bp product DPP9 (nucleotides 1–951) and 1154-bp product DPP10 (nucleotides 883–2037) probes, which were radioactively labelled by random priming in the presence of [^{32}P]dCTP. Hybridization was performed at 68 °C overnight in ExpressHyb™ hybridization solution (ClonTech Laboratories). The blots were first washed in 2 × SSC buffer (0.15 M NaCl/0.015 M sodium citrate, pH 7.0) and 0.05 % SDS at room temperature (22 °C), and subsequently washed in 0.1 × SSC and 0.1 % SDS at 60 °C (DPP8 and DPP9) and 50 °C (DPP10). Northern blots were developed using a Konica automatic medical film processor (QX-130A plus). Northern-blot lane loading was quality-controlled by the manufacturer and verified by using β-actin probes. Quantitative PCR analysis was performed to examine the levels of expression of the mRNAs for DPP8, DPP9, DPP10 and DPP4 in cDNA from normalized human tissues [human multiple-tissue cDNA (MTC™) panels I and II; ClonTech Laboratories] and cell lines. Oligonucleotide primers listed in Table 1 (Exp-DPP8-1, Exp-DPP8-2, Exp-DPP9-1, Exp-DPP9D-2 and Exp-DPP10-1, Exp-DPP-2) were used for quantitative PCR analysis. Each cDNA (0.5 ng) was used in a 25 μl PCR, with each primer at a final concentration of 300 nM. For human tissues, the PCR was performed using a SYBR Green PCR Core Reagents kit (Applied Biosystems). The quantitative PCR analysis was performed at 50 °C for 2 min, 95 °C for 10 min followed by 40 PCR cycles at 95 °C for 15 s and 60 °C for 1 min. The amplified products were analysed by electrophoresis or detected with a GeneAmp 5700 sequence detection system (Applied Biosystems).

Mammalian expression and purification of DPP proteins

C-terminal fusion plasmid DNA (10 μg) of pcDNA-MycHis-DPP8, pcDNA-MycHis-DPP9 and pcDNA-MycHis-DPP10 were transfected into 10 cm dishes of PEAK (Edge Bio Systems, Gaithersburg, MD, U.S.A.) or COS-1 cells (CRL-1650; A.T.C.C.) with 24 μl of LIPOFECTAMINE™ (LipofectAmine, Gaithersburg, MD, U.S.A.). Transfected cells were maintained in Dulbecco's modified Eagle's medium with 5 % (w/v) foetal bovine serum at 37 °C with 5 % CO_2 for 48 h. Cell pellets (2×10^7) were harvested 48 h after transfection, homogenized on ice and then spun at 18000 ***g*** for 40 min. The supernatant was collected as cytosolic fractions, and cell pellets were washed once and collected as membrane fractions. Solubilized membrane fractions (0.25 % Tween 20) or cytosolic fractions were loaded on to TALON affinity spin columns (ClonTech Laboratories), and His-tagged proteins were eluted with 50 mM PBS and 150 mM imidazole (pH 7.0). Proteins in the eluted fractions were resolved by SDS/PAGE. Samples (10 μl) were run on 4–15 % (w/v) SDS/polyacrylamide gels, which were then stained with BioSafe™ Coomassie G25 (Bio-Rad Laboratories, Hercules, CA, U.S.A.). After electrophoretic transfer to Trans-Blot™ nitrocellulose membranes (Bio-Rad Laboratories), recombinant proteins were detected by Western blotting using anti-myc antibody. Protein bands were visualized using an enzyme-linked immunodetection system with anti-IgG secondary antibodies conjugated with alkaline phosphatase (ProtoBlot II AP system; Promega).

Enzymic activity of recombinant DPP proteins

The kinetic properties of purified recombinant DPP8, DPP9, DPP10 and DPP4 (5 m-units/μl; Sigma) were determined in a continuous fluorimetric assay measuring accumulation of 7-amino-4-methylcoumarin (AMC) liberated from Ala-Pro-AMC. First, buffer (PBS, Hepes and Tris)-, pH (5–9)- and temperature (25 and 37 °C)-dependent assays were performed using 15.6 μM–1 mM Ala-Pro-AMC (Enzyme Systems Products, Livermore, CA, U.S.A.) to find the optimal reaction conditions. Subsequently, enzyme assays were performed in 50 mM Hepes buffer (pH 8) with 50 μl (50 μg/ml) of purified enzyme mixed with 1 μl of various concentrations of Ala-Pro-AMC (100 mM) in a 96-well plate. Plates were incubated at 37 °C for 30 min, and fluorescence was detected using a Wallac 1420 Fluorimeter (PerkinElmer, Norwalk, CT, U.S.A.) at an excitation wavelength of 355 nm and an emission wavelength of 535 nm. Fluorescence intensity was recorded over 60 readings during a

60 min period. In chromogenic assays, *p*-nitroaniline liberated from Ala-Pro *p*-nitroanilide (-pNA) (1 mM) was measured by absorbance A at 410 nm. Initial-velocity data for substrate hydrolysis were determined by assaying the enzyme activity at increasing (15.6 μM to 2 mM) concentrations of substrate. K_m and V_{max} values were obtained by fitting data to the Michaelis–Menten equation ($V = V_{max}[S]/K_m + [S]$). Additional substrates such as Gly-Pro-AMC, Z-Gly-Pro-AMC, Ala-Pro-pNA, Pro-pNA, Gly-Arg-pNA, Lys-Ala-pNA and Ala-Phe-Pro-pNA (Bachem, Torrance, CA, U.S.A.) were used to determine whether the newly identified DPP proteins had di-, tri- or endopeptidase activities. A series of enzyme inhibitors were used to confirm the classification of the protease activity (Table 4). These include reversible and irreversible DPP4 inhibitors (10–200 μM) di-isopropyl fluorophosphate and diprotin A (Calbiochem, La Jolla, CA, U.S.A.). DPPIV inhibitor II ['H-Glu-(NHO-Bz)-Pyr-HCl'; Calbiochem], an acylating inhibitor of the proline-specific peptidases [31] (DPP4, DPP2 and prolyl endopeptidase), was also tested for DPP8 and DPP9 enzyme activity inhibition. Other inhibitors tested include serine protease inhibitors 4-(2-aminoethyl)benzenesulphonyl fluoride hydrochloride (5 mM) and aprotinin (30–120 μg/ml); the aspartic protease inhibitor, pepstatin; a metalloproteinase inhibitor, EDTA and leupeptin, which inhibits serine and cysteine proteases such as trypsin, papain and cathepsin (Sigma–Aldrich, St. Louis, MO, U.S.A.). The recombinant enzymes were preincubated for 30 min at room temperature with inhibitors before the addition of the substrate.

RESULTS

Isolation of DPP8, DPP9 and DPP10 cDNA clones and protein sequence analyses

We found that the human DPP4 gene subfamily contains at least six members: DPP4, DPP6, FAP/seprase and three additional sequences (GenBank® accession nos. AY172659, AY172660 and AY172661), designated originally as DPRP-1, DPRP-2 and DPRP-3 [32]. The deduced amino acid sequence of DPRP-1 is identical with the recently cloned DPP4-like protease DPP8 (GenBank® accession no. AF221634) [16]. The cloning of DPP9, a protein that is identical with DPRP-2, has been reported very recently by Olsen and Wagtmann [33]. The third DPP4-like protein, DPP10 (DPRP-3), was derived from pairwise sequence extension using EST R52809 and the human genomic sequences (AC016721 and AC016874). A partial 712-amino-acid ORF, derived from AB040925, was extended by RACE PCR enabling identification of the 5′-region. The full-length DPP10 protein sequence of 796 amino acids was encoded by a 2583-bp cDNA sequence (Figure 1). The DPP10 cDNA sequence aligned with 99% identity to part of an unannotated human foetal brain cDNA (NM_020868) [34]. DPP10 contains ten potential N-linked glycosylation sites. Comparison of the identified DPP10 cDNA sequence with the genome sequence deposited in the Human Genome Project Database suggested that DPP10 belongs to a gene with at least 23 exons and two transcripts, and maps to chromosome 2 (NT_005445) at position 2q12.3-2q14.2. Two mouse sequences (XM_193952 and XM_129475) identified *in silico* during the sequencing of the mouse genome have similarity to DPP10 and KIAA1492 cDNAs. Both sequences were aligned with a National Center for Biotechnology Information mouse contig (NW_000154) derived from assembled genomic sequence data. Alignment with the human DPP10 amino acid sequence showed that the two mouse sequences XM_193952 and XM_129475 shared more than 90% identity with the N-terminal (1–152) and C-terminal (322–796) regions respectively.

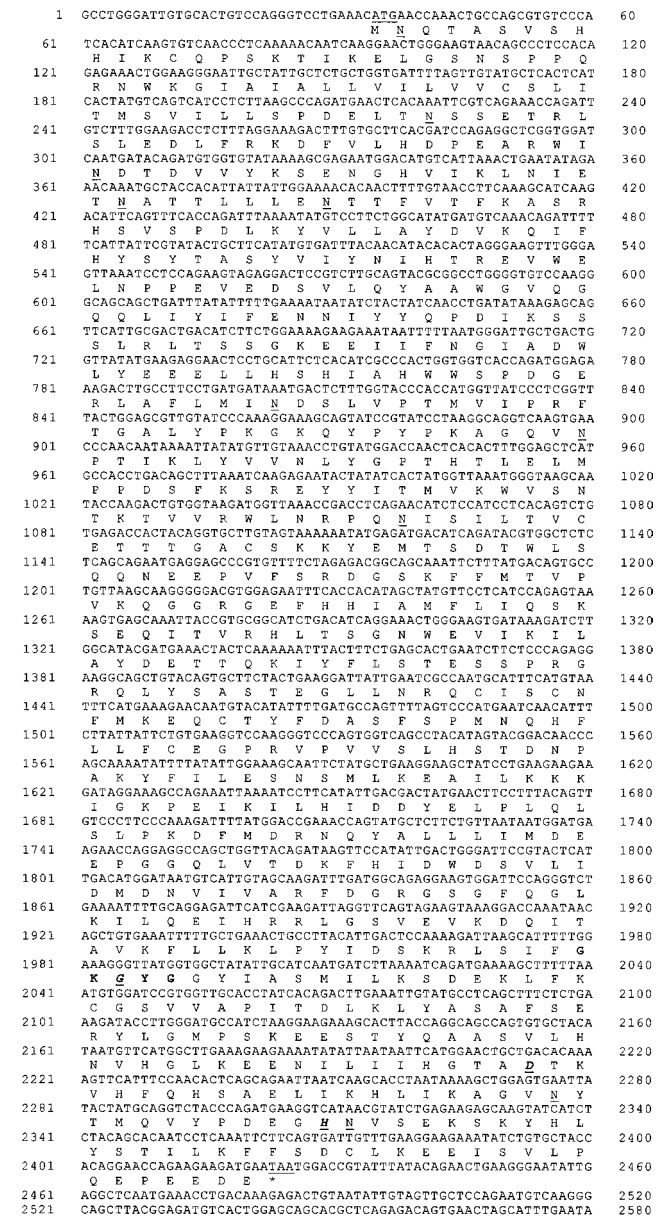

Figure 1 Nucleotide sequence of the DPP10 cDNA numbered 5′ to 3′ (GenBank® accession no. AY172661) and the deduced 796-amino-acid sequence

Start and stop codons in the nucleotide sequence are underlined. The serine residue is highlighted in boldface and catalytic triad residues are indicated in underlined italic boldface letters. Potential N-linked glycosylation sites are underlined.

DPP10 has the highest sequence identity, 51 and 48% respectively, with short (DPP6-S) and long (DPP6-L) isoforms of DPP6 (also known as DPPX-L and DPPX-S). Significant sequence identity with DPP4 (32%) and FAP (29%) suggests that DPP10 is closely related to the POPs in family S9B of clan SC (Figure 2). A search in the conserved domain database confirmed high similarity between DPP10 and the domains found in DPP4 and the existing S9B proteases. DPP10 had the highest score of the three new proteases for the DPPIV N-terminal region (PF00930) and the POP family domain (PF00326) and was close to FAP and DPP6 (Table 2). Analysis of the primary amino acid sequence of DPP10 reveals that the protein lacks the active-site serine

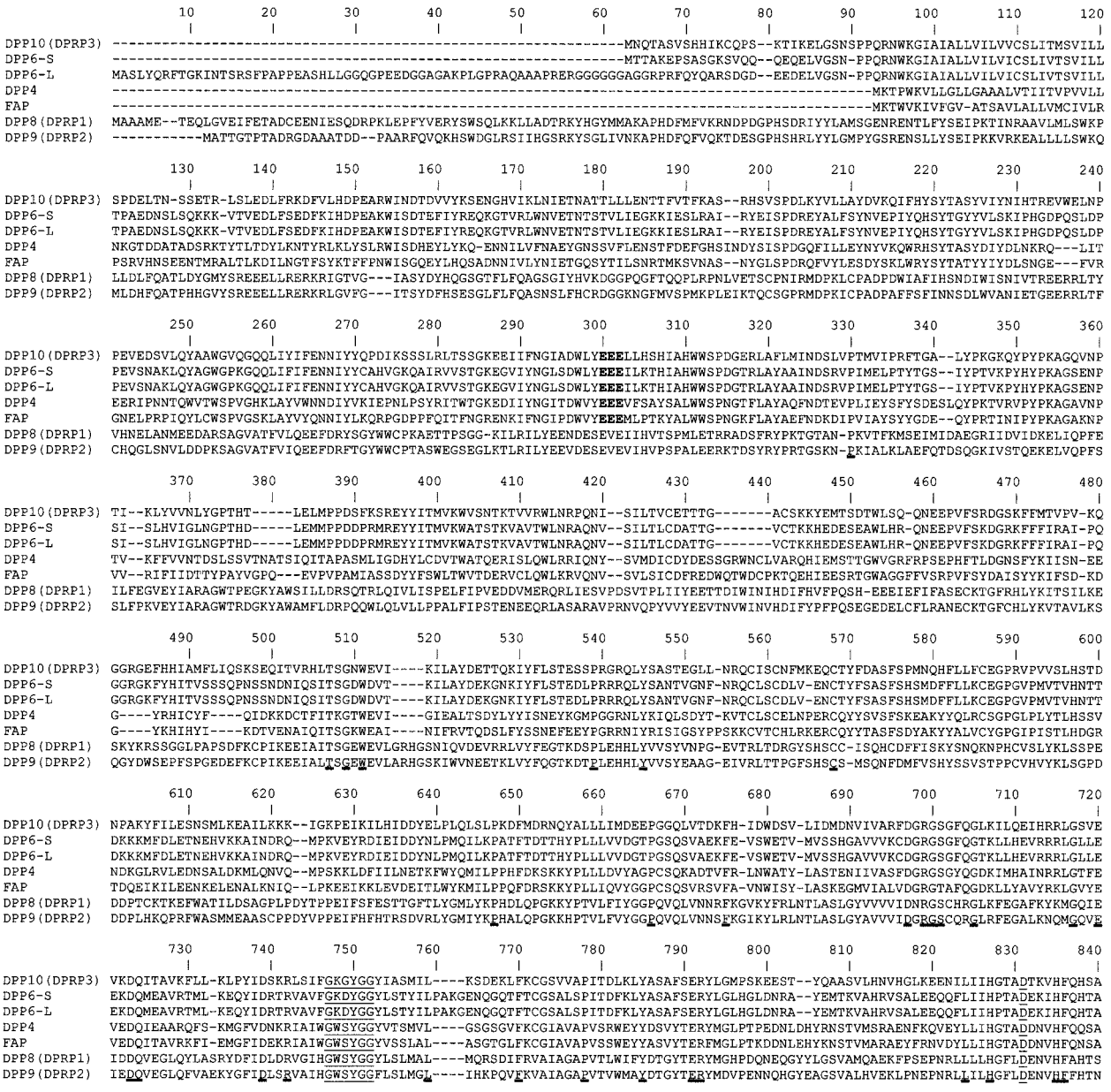

Figure 2 Multiple-sequence alignment of deduced DPP10 and existing members of family S9B

DPP6 short isoform (DPP6-S, NP_001927.2), DPP6 long isoform (DPP6-L, NP_570629.1), DPPIV (DPP4, NP_001926.2), FAP (AAB49652.1), DPP8 (AAG29766.1) and DPP9 (NP_631898.1). Amino acids conserved in all six proteins are underlined in the DPP9 sequence. The serine, asparagine and histidine residues that comprise the catalytic triad are underlined in all six protein sequences. The S9B consensus sequence DW(V/L)YEEE glutamate residues are shown in boldface.

residue, which has been substituted with a glycine residue (Figure 2). The other active-site residues (aspartate and histidine) are retained. Three glutamate residues and the overall consensus sequence in a conserved set of residues DW(V/L)YEEE found in the N-terminal region of DPP4, FAP and DPP6, as described by Abbott et al. [35], are also found in DPP10 (Figure 2). Several of the conserved amino acids found in six conserved domains characteristic of the peptidases, lipases, esterases, epoxide hydrolases or serine hydrolases (PLEES) superfamily [36] are also conserved in DPP10 as well as in DPP8 and DPP9 (Figure 2).

In agreement with a previous study [16], the isolated nucleic acid for DPP8 contains an ORF encoding a protein of length 882 amino acids and shares 20 % identity and 32 % similarity over the entire human DPP4 protein sequence. The DPP9 nucleic acid sequence contains an ORF encoding for a protein of 863 amino acids, which is 19 % identical and 31 % similar to the entire DPP4

Table 2 Search results from the Conserved Domain Database and compilation of multiple-sequence alignments representing protein domains conserved in molecular evolution

Comparisons use reverse-position-specific BLAST, a variant of the widely used PSI-BLAST algorithm. DPP8, DPP9 and DPP10 sequences were compared against databases of position-specific score matrices derived from alignments in Conserved Domain Database [29]. PFAM, protein families database of alignments and heavy meromyosins; DLH, dienelactone hydrolase.

	PFAM	hDPP4	hFAP	hDPP6	hDPP10	hDPP8	hDPP9
DPPIV	PF00930	1×10^{-170}	3×10^{-169}	2×10^{-155}	2×10^{-115}	3×10^{-30}	1×10^{-26}
POP	PF00326	2×10^{-26}	6×10^{-17}	2×10^{-12}	1×10^{-13}	3×10^{-8}	6×10^{-8}
DLH family	PF01738	1×10^{-4}	7×10^{-4}	–	1×10^{-3}	9×10^{-3}	2×10^{-3}

amino acid sequence. DPP8 and DPP9 both have the same active-site consensus sequence (Gly-Trp-Ser-Tyr-Gly) found in DPP4. The greatest similarity between DPP8, DPP9 and DPP4 is seen in the C-terminal sequences. DPP10 is approx. 20% identical with DPP8 and DPP9. In contrast with DPP8 and DPP9, which do not contain transmembrane regions, DPP10 was found to have a 23-amino-acid hydrophobic region between amino acids 34 and 56 (GIAIALLVILVVCSLITMSVILL). SMART analysis [28] predicted that this region represents a transmembrane domain. If this putative transmembrane region is confirmed, DPP10, like DPP4, FAP and DPP6, is a type II membrane protein (Table 3).

Relationships between the 34 DPP4-like sequences from bacteria, nematodes, insects and vertebrates, analysed by constructing a phylogenetic tree using the Neighbour-Joining algorithm on the basis of their deduced amino acid sequences, revealed three clusters of DPP4-like proteins. The analysis shows that DPP10 and DPP6, which share a mutation at the active-site serine residue, were clustered together in a separate subfamily to FAP and DPP4. DPP8 and DPP9 belong to an evolutionarily divergent group of genes separated from DPP10, vertebrate orthologues of DPP4, and other human DPP4-like serine proteases (Figure 3). DPP8 and DPP9 cluster with bacterial DPP4 proteins, including *Caulobacter crescentus* DPP4 and *Xanthomonas axonopodis* DPP4 and several *Drosophila melanogaster* and *Caenorhabditis elegans* DPP4-like proteins.

Expression of DPP10 in human tissues compared with DPP8 and DPP9

Northern-blot analysis showed expression of DPP10 mRNA in brain and pancreas (Figure 4). A major DPP10 mRNA transcript of approx. 3.5 kb was seen in both tissues. Two larger transcripts at 5.0 and 7.5 kb were seen in the brain, but not in the pancreas. ESTs, representing the DPP10 mRNA, are abundant in tissues derived from multiple-sclerosis lesions, retinoblastoma, hypothalamus, hippocampus and whole brain, with few transcripts being found in uterus, colon and various tumours. Mouse ESTs, which comprise the DPP10 mouse orthologue cDNA, are also expressed in several brain regions (corpora quadrigemina, cortex, spinal cord, cerebellum, head, medulla oblongata, subfornical organ, area postrema, hypothalamus, diencephalon and olfactory brain) and the retina.

The DPP10 expression pattern was much more discrete than those observed for DPP8 and DPP9, which display a somewhat ubiquitous pattern of expression in the present and previous studies. DPP8 transcripts are found in several tissues, with the most abundant signal being in testis, prostate, muscle and brain. Testis expressed three transcripts, approx. 7.5, 4.5 and 2.5 kb in length. The shorter mRNA species was very abundant in testis, but negligible in the other tissues tested. DPP9 mRNA was expressed ubiquitously in every tissue, with highest levels in liver and muscle, and a predominant transcript at 4.5 kb. A larger transcript, approx. 5.0 kb, was observed in liver, which was also expressed as a small 1.4 kb transcript. Quantitative PCR analysis confirmed the tissue distribution shown in the Northern blots (results not shown). ESTs, representing the DPP8 transcripts, were found in numerous tissues, including senescent fibroblasts, T-lymphocytes, germinal centre B-cells, germ-cell seminoma, testis, melanocytes, uterus, ovary, breast, multiple-sclerosis lesions, pancreas and placenta. DPP9 was expressed in a wide variety of tissues on examination of ESTs coverage (e.g. over 64 ESTs expressed in liver, spleen, muscle, melanocytes, heart, lung, placenta, skin, pancreas, stomach, brain and parathyroid gland).

Enzymic activity of the purified enzymes

Recombinant affinity-purified myc fusions of DPP8, DPP9 and DPP10 were separated on SDS/polyacrylamide gel (Figure 5A), and detected by Western-blot analysis using an anti-myc antibody at masses similar to the predicted molecular masses of approx. 100 kDa for DPP8 and DPP9 respectively, and 91 kDa for DPP10 (Figure 5B). DPP8 and DPP9 were found in cytosolic fractions, whereas DPP10 was found in the cell pellet representing the membrane fraction. No bands were detected in the untransfected PEAK cell extract. As predicted, based on amino acid alignments with other members of the S9B POP family, biochemical characterization reveals that DPP10 does not possess DPP activity (results not shown), whereas DPP8 and DPP9 have similar enzymic profiles to DPP4. Maximum DPP8 and DPP9 enzyme activity was observed between pH 7.5 and 8.5, with a decrease in

Table 3 Comparison of properties of DPP8, DPP9 and DPP10 with DPP4 and other related peptidases

QPP, quiescent cell proline dipeptidase.

Protease family	Protease name	EC number	Number of amino acids	Transmembrane region	Ser-Asp-His triad	Gene location	Optimal pH
Clan SC, family S9B	DPP4 (CD26)	3.4.14.5	766	Y	Y	2q24.3	7.5–8.0 [3,48]
	FAP (seprase)	–	760	Y	Y	2q23	7.5–8.0 [11,14]
	DPP6 (DPPX)	–	865	Y	Ser mutation	7q36.1-q36.2	–
	DPP8 (DPRP1)	–	882	N	Y	15q22.1-15q22.2	7.5–8.0 [16]
	DPP9 (DPRP2)	–	863	N	Y	19p13.3	7.5–8.0
	DPP10 (DPRP3)	–	796	Y	Ser mutation	2q12.3-2q14.2	–
Clan SC, family S28	QPP, DPP2 (DPP7)	3.4.14.2	492	N	N	19	4.5–7.5 [18,45]
Clan MA, family M49	DPP3	3.4.14.4	737	N	N	11q12-q13.1	8.5–8.8 [49]
Clan CA, family C1	DPP1 (cathepsin C)	3.4.14.1	463	N	N	11q14.1-q14.3	5–5.5 [50]

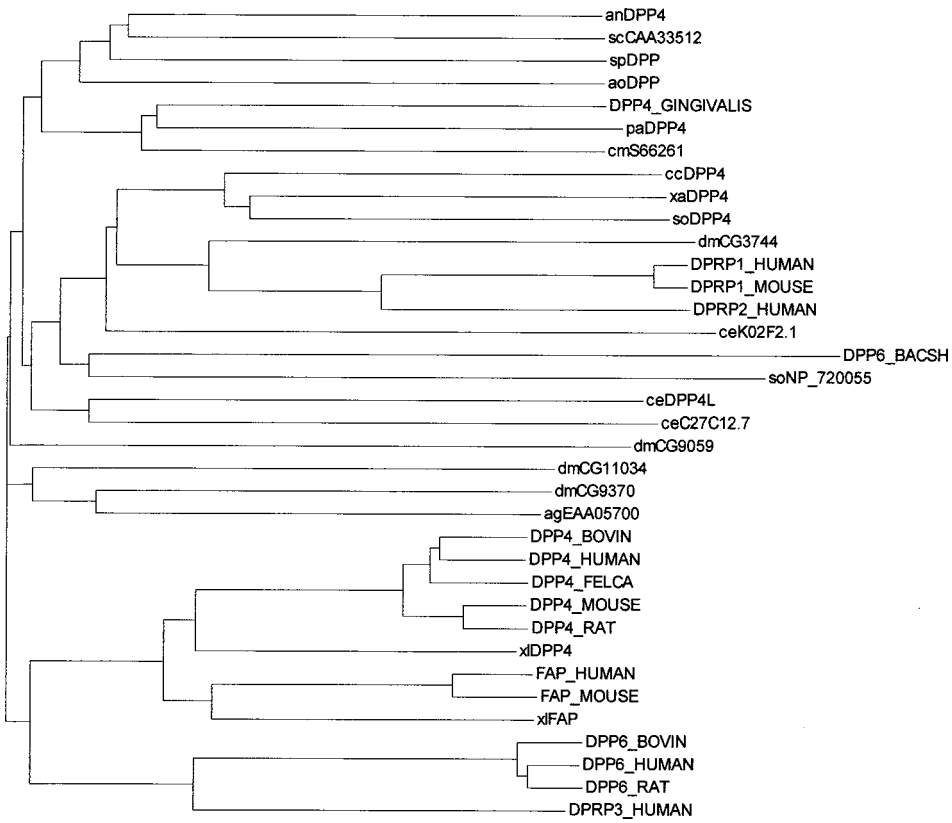

Figure 3 Phylogenetic relationship of DPP4-like serine proteases, including DPP8, DPP9 and DPP10 (reported in the present study), based on the primary amino acid sequences

The tree was constructed by the Neighbour-Joining method. The lengths of lines indicate the relative distances between nodes. Enzymes used for alignment are as follows [species name (code, GenBank® accession number)]. *Aspergillus fumigatus* (afDPP4, AAC34310), *Anopheles gambiae* (agEAA05700, EAA05700), *Aspergillus niger* (anDPP4, CAC41019), *Aspergillus oryzae* (aoDPP, CAA05343), *Bacillus sphaericus* (DPP6_BASCH, P39043), *Bos taurus* (btDPP4, P81425), *B. taurus* (btDPP6, P42659), *C. crescentus* (ccDPP4, NP_420957), *C. elegans* (ceC27C12.7, NP_510461), *C. elegans* (ceDPP4L, NP_506851), *C. elegans* (ceK02F2.1, T32919), *Chryseobacterium meningosepticum* (cmS66261,S66261), *D. melanogaster* (dmCG11034, NP_608961), *D. melanogaster* (dmCG3744, NP_733056), *D. melanogaster* (dmCG9059, NP_7280410), *Felis silvestris catus* (fcDPP4, Q9N2I7), *Homo sapiens* (hsDPP4, NP_001926), *H. sapiens* (hsDPP6, P42658), *H. sapiens* (hsFAP, NP_004451), *Mus musculus* (mmDPP4, NP_034204), *M. musculus* (mmDPP6, NP_034205), *M. musculus* (mmFAP, NP_032012), *P. albensis* (paDPP4, CAC42932), *Rattus norvegicus* (rnDPP4, NP_036921), *R. norvegicus* (rnDPP6, NP_074041), *R. norvegicus* (rnFAP, NP_620205), *Saccharomyces cerevisiae* (scCAA33512, CAA33512), *S. oneidensis* (soDPP4, NP_719520), *S. oneidensis* (soNP_720055, NP_720055), *Saccharomyces pombe* (SpDAP, NP_594920), *S. pombe* (SpDPP, NP_593970), *X. axonopodis* (xaDPP4, NP_644345), *X. campestris* (xcDPP4, NP_639300), *Xenopus laevis* (xlDPP4, CAA70136) and *X. laevis* (xlFAP, AAC59872).

activity being observed below pH 6.5 (Figure 6A). The K_m values of DPP8 and DPP9 were similar (208 and 161 μM respectively; Figure 6B).

Enzyme activities of DPP4, DPP8 and DPP9 were inhibited potently by di-isopropyl fluorophosphate and 4-(2-aminoethyl)benzenesulphonyl fluoride-HCl at pH 7.0 (Table 4). DPP8 and DPP9 displayed different sensitivities to DPPIV inhibitor II and diprotin A, when compared with DPP4. Aprotinin, which has specificity for serine proteases with trypsin-like tertiary structures in the vicinity of the active site, had little effect on the activity of the three enzymes. All three proteases were resistant to the other protease inhibitors tested (Table 4). DPP8 and DPP9 were also confirmed to be DPPs as a result of the inactivity of these enzymes against additional substrates (Ala-Pro-pNA, Pro-pNA, Gly-Arg-pNA, Lys-Ala-pNA Ala-Phe-Pro-pNA) (Table 5). These results support the structural and domain analysis prediction that these proteins belong to the POP family.

DISCUSSION

In the present study, we describe the cloning and comparative characterization of three proteases (DPP8, DPP9 and DPP10) with significant similarity to the existing serine proteases in the POP S9B family of the enzyme clan SC. The order of the catalytic triad residues is different in the POPs (Ser-Asp-His) compared with clan SA chymotrypsin-like (His-Asp-Ser) and clan SB subtilisin-like (Asp-His-Ser) enzymes. Despite the replacement of the serine residue in DPP10, it is clear that DPP10 is a member of the S9B family of proteins, due to the significant identity of the primary amino acid sequence with DPP4, FAP and DPP6. Substitution of the serine residue in DPP10, as described in the present study, is also found in a related protein DPP6. Originally isolated from bovine and rat brain libraries, DPP6 was found to be inactive as a DPP [37]. This was originally assumed to be due to an aspartate residue replacing the serine residue at the active site. However, DPP6 was found to have no DPP4-like enzyme activity after site-directed mutagenesis to replace the serine residue and complete the consensus active-site sequence for serine proteases. Both DPP6 and DPP10 also have a second amino acid substitution (Trp $\rightarrow$ Lys) in the active-site consensus sequence. In the present study, we have confirmed that DPP8 and DPP9 are both functionally active members of the POP family, with catalytic residues in the order Ser, Asp, His, and the active-site sequence (Gly-X-Ser-X-Gly) found in DPP4 [38] and other catalytically active members of the S9B subfamily of

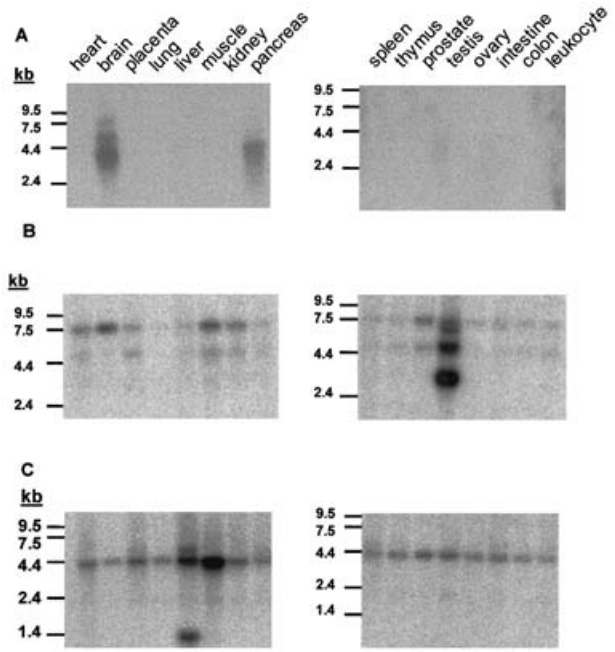

Figure 4 Northern-blot analysis of the localization of DPP8, DPP9 and DPP10 mRNA in various human tissues

Northern-blot analysis was performed with 2 μg of poly(A)$^+$ RNA. A human multiple-tissue Northern blot (ClonTech Laboratories) was probed with the following ^{32}P-labelled probes: (**A**) 1154-bp DPP10 fragment, (**B**) 851-bp (DPP8) N-terminal fragment and (**C**) 915-bp DPP9 fragment. Hybridization was performed at 68 °C overnight. The filters were then washed in 2 × SSC and 0.05 % SDS at room temperature, and then washed in 0.1 × SSC and 0.1 % SDS at 60 °C (DPP8 and DPP9) and 50 °C (DPP10) respectively.

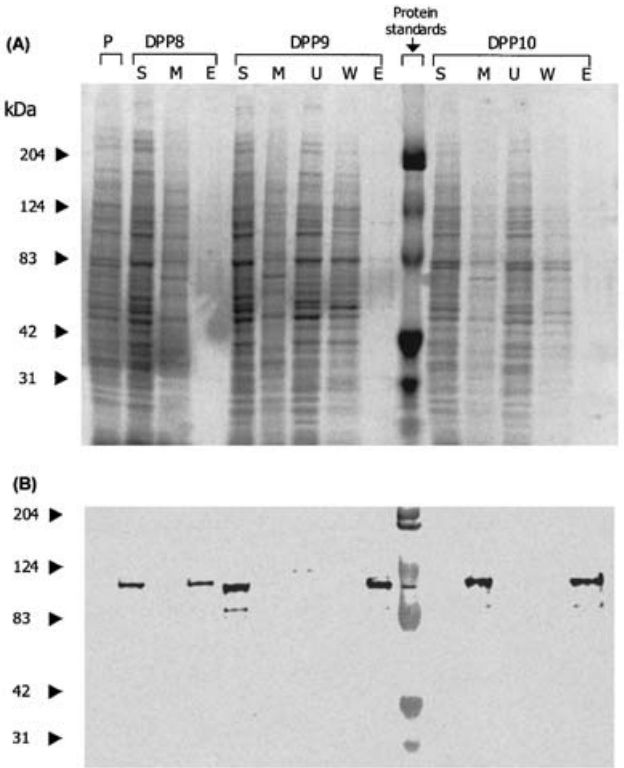

Figure 5 SDS/PAGE of recombinant DPP8, DPP9 and DPP10 expressed in PEAK cells

Samples (10 μl) were analysed by electrophoresis on a 4–15 % gradient gel. (**A**) Coomassie Blue-stained gel. Lane 1, untransfected PEAK, whole-cell lysate (P); lane 2, DPP8 transfected PEAK cell-soluble fraction (S); lane 3, DPP8 transfected PEAK cell-membrane fraction (M); lane 4, DPP8 transfected PEAK cell from column elution (E); lane 5, DPP9 transfected PEAK cell-soluble fraction (S); lane 6, DPP9 transfected PEAK cell-membrane fraction (M); lane 7, DPP9 transfected PEAK cell-unbound fraction from the TALON column (U); lane 8, DPP9 transfected PEAK cell column wash fractions (W); lane 9, DPP9 transfected PEAK cell from column elution (E); lane 10, marker; lane 11, DPP10 transfected PEAK cell-soluble fraction (S); lane 12, DPP10 transfected PEAK cell-membrane fraction (M); lane 13, DPP10 transfected PEAK cell-unbound fraction from the TALON column (U); lane 14, DPP10 transfected PEAK cell column wash fractions (W); lane 15, DPP10 transfected PEAK cell from column elution (E). (**B**) Recombinant proteins (lane arrangement as detailed above) were then detected, following electrophoretic transfer to nitrocellulose membranes, by Western blotting with anti-myc antibody. Protein bands were visualized using an enzyme-linked immunodetection system with anti-IgG secondary antibodies conjugated with alkaline phosphatase (ProtoBlot II AP system).

serine peptidases. In a previous study describing the isolation of DPP9, no significant DPP4-like enzyme activity was observed [33]. The lack of activity may have been due to the use of a translation system *in vitro* to express the recombinant DPP9 protein. Expression of recombinant enzymically active DPP4, FAP [12] and DPP8 [16] proteins in mammalian expression systems has been demonstrated previously. Thus it appears that these proteases and DPP9, as described in the present study, are produced in their active form by mammalian recombinant expression. Furthermore, enzymic activity is not dependent on pretreatment as is required for partial proteolytic activation of zymogens found in other serine protease subfamilies. It is possible that important post-translational modifications absent from the translation system *in vitro* were responsible for the lack of activity seen in the previous DPP9 study. For example, a difference in post-translational modifications and final protein characteristics following biosynthesis of DPP6 in both *in vitro* and *in vivo* systems has been demonstrated previously [37]. In a cell-free translation system, short (DPP6-S) and long (DPP6-L) isoforms of DPP6 (also known as DPPX-L and DPPX-S) were synthesized as 93 and 97 kDa forms respectively, in good agreement with the molecular masses estimated from their primary structure. In COS-1 cells transfected with DPP6-S and DPP6-L cDNAs, modified with high-mannose-type oligosaccharides, resulted in 115 and 120 kDa forms respectively [37]. Although the role of glycosylation was not tested in the present study, it is possible that post-translational modifications are required to produce an enzymically active protein for DPP9. Although DPP8 does not have any N- or O-linked glycosylation sites, DPP9 may be glycosylated at amino acids 205 and 653. Based on sequence similarity and functional activity demonstrated in the present and previous studies, DPP8, DPP9 and DPP10 can all be grouped as S9B family members and with the DASH family of proteins [9].

It is interesting to note that DPP10 was found in the same high-throughput genome clone (AC016721) that contains part of the DPP4 gene sequence. DPP10 is located on the long arm of chromosome 2 (2q12.3-2q14.1), where DPP4 (2q24.3) and FAP (2q23) genes are also found. DPP8 and DPP9 were found on chromosomes 15 and 19 respectively. Previous analysis of the gene structure of DPP8 and DPP9 shows that the catalytic serine, asparagine and histidine triad residues are encoded by a single exon, whereas in DPP4 and FAP the triad residues are divided between two separate exons. This difference in exon structure of DPP8 and DPP9 is assumed to be due to the divergence of DPP8 and DPP9 from the remaining DPP4-like genes earlier in evolution. The gene structure of DPP10 was not analysed in detail in the present study. However, the divergence of DPP8 and DPP9 from DPP10 and other DPP4-like sequences was confirmed by phylogenetic analysis of the DPP proteins as well as eukaryotic, prokaryotic and vertebrate DPP4 proteins and DPP4-like proteins.

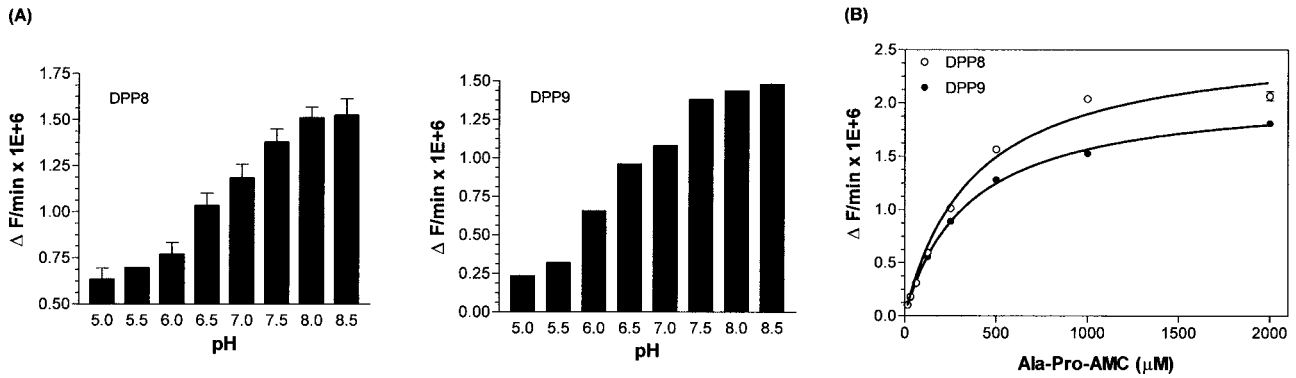

Figure 6 Recombinant DPP8 and DPP9 Ala-Pro-AMC cleaving activity

Assays were conducted using 50 μl (50 μg/ml) of purified enzymes in 50 mM Hepes (pH 8.0). Purified enzymes were added to 1 μl of various concentrations of Ala-Pro-AMC and incubated at 37 °C for 30 min. Fluorescence was detected at $\lambda_{excitation}$ 355 and $\lambda_{emission}$ 535. (**A**) pH dependence of DPP8 and DPP9, and (**B**) Michaelis–Menten plots for DPP8 and DPP9 hydrolysis of Ala-Pro-AMC.

Table 4 Inhibition of DPP8 and DPP9 enzyme activity by protease inhibitors

			Residual activity (% of control)		
Inhibitor	Inhibitor property	Concentration	DPP8	DPP9	DPP4
AEBSF*	Serine, irreversible	5 mM	29.6	23.9	21.1
Di-isopropyl fluorophosphate	DPP4, irreversible	100 μM	23.9	26.6	22.2
Diprotin A	DPP4, reversible	100 μM	73.0	47.1	34.9
DPPIV inhibitor II	DPP4, reversible	100 μM	87.4	79.3	43.8
Aprotinin	Serine, reversible	120 μg/ml	64.1	76.0	63.6
Pepstatin	Aspartic, reversible	2 μg/ml	97.3	95.0	93.5
EDTA	Metallo, reversible	2 mM	91.5	86.0	93.5
Leupeptin	Cysteine/serine, reversible	50 μg/ml	91.1	90.4	90.7

* Abbreviation: AEBSF, 4-(2-aminoethyl) benzenesulphonyl fluoride.

Table 5 Confirmation of DPP8 and DPP9 as DPPs based on the inactivity of these enzymes against additional substrates

	Amount of product after 30 min incubation (% increase)		
Substrate	DPP8	DPP9	DPP4
Ala-Pro-AMC	239.0	127.5	379.0
Gly-Pro-AMC	341.5	205.0	444.0
Ala-Pro-pNA	45.5	44.0	29.5
Pro-pNA	−1.0	−2.5	0.0
Gly-Arg-pNA	−4.5	−0.5	0.0
Lys-Ala-pNA	2.5	0.5	0.5
Ala-Phe-Pro-pNA	−4.0	−0.5	2.0

This analysis suggests that DPP10 is clustered with DPP6-L (51 % identity) and is closely related to the vertebrate DPP4 orthologues (cow, cat and mouse). DPP8 and DDP9 seem to be more closely related to ancient DPP4 enzymes found in bacteria, nematodes and arthropods. Furthermore, DPP10, DPP4, FAP and DPP6 all share an identical N-terminal consensus sequence DW(V/L)YEEE, proposed by Abbott et al. [35] to be characteristic of the subfamily S9B. This consensus sequence is present in the predicted β-propeller domain of DPP4-like genes, and Glu^{205} and Glu^{206} in this motif have been shown to be essential for dipeptidase activity of DPP4. DPP8 and DPP9 show several amino acid differences in this consensus sequence (Figure 2) that are more closely aligned with a *Drosophila* DPP4-like protein (NP_733056). Despite this, they do retain an aligned pair of glutamate residues in this motif. Indeed, all three recently cloned DPP protein sequences show several areas of conservation within the six domains or blocks of the superfamily of eukaryote and prokaryote PLEES enzymes [36] in addition to the three that are conserved around the amino acids which would form the catalytic triad (Figure 2).

The highly conserved patterns around the residues forming the active site are assumed to play an important role in the conformation of the enzyme. The α/β hydrolase fold found in POP proteases and all PLEES enzymes comprises a crucial structural element of the catalytic apparatus. Structural analysis of pig POP showed that the C-terminal domain is comprised of alternative α-helix and β-sheet segments, which are structurally related to those found in α/β hydrolases (lipases, cholinesterases, dehalogenases and serine carboxypeptidases) [39]. These enzymes have diverged from a common ancestor to preserve the arrangement of the catalytic triad, the components of which are borne on highly conserved loops [40]. A β-propeller domain, based on a 7-fold repeat of four-stranded antiparallel β-sheets, creates a central tunnel which, along with the hydrolase fold, forms the surface of the inner cavity of the enzymes in which catalysis takes place [39]. The three-dimensional structure of POPs predicts a narrow entrance, approx. 4 Å (1 Å = 10^{-10} m) into the hollow enzyme, with some degree of flexibility brought by partial separation of blades 1 and 7 due to a lack of covalent bonds usually seen in other propeller domains [41]. Limited access to the active site, located about 20 Å from the narrow entrance, is assumed to be a mechanism for substrate selection and prevention of uncontrolled protein degradation in which entrance and proteolysis of larger proteins and structured peptides are restricted, whereas oligopeptides of not more than 30 amino acids in length are potential substrates [41,42]. Thus in accordance with previous work by Abbott et al. [16], although the structures of DPP8, DPP9 and DPP10 have not yet been solved, we propose, on the basis of primary amino acid sequence identity, that these proteases are members of the S9B serine protease family that utilizes potentially the same mechanism of action.

Although the absence of the serine residue prevents DPP activity via an intact catalytic triad in DPP10 and DPP6, it is probable that functions mediated by other functional domains of the protein remain intact. DPP4 is a multifunctional molecule that exerts important functions in addition to its catalytic activity as a peptidase. For example, activated T-cell signalling is mediated by the association of DPP4 with CD45 tyrosine phosphatase [43].

Recently, DPP6 has been suggested to play a role in activities that do not appear to involve DPP activity. During the preparation of the present paper, DPP6 was reported to play a crucial role in the trafficking, membrane targeting and function of A-type K^+ channels in somatodendritic compartments of neurons [44]. These A-type K^+ channels mediate currents that are important in neuronal function (dopamine release) and dysfunction (e.g. Parkinson disease). DPP6 is expressed highly in the hippocampus, thalamus, hypothalamus and striatum and may be involved in the regulation of neuronal plasticity. DPP10 is also expressed preferentially in the brain in both rodents and humans. Further experiments are required to determine the role that DPP10 plays in neuronal function. The expression of DPP10 ESTs in numerous multiple-sclerosis cDNA libraries and optical centres is intriguing. Several of the brain regions expressing DPP10 mRNA in human beings and mice are associated with optical reflexes and the correlation centres for auditory reflexes. Further immunohistochemical or *in situ* hybridization work is required to localize the DPP10 protein in normal central nervous system tissues and in samples from pathological conditions such as multiple sclerosis.

DPP10, like DPP4, FAP/seprase and DPP6, is predicted to be a type II membrane protein and may be cleaved to produce a soluble form. DPP8 and DPP9 do not have transmembrane domains and are, therefore, predicted to be cytosolic proteases. Moreover, the neutral pH at which they display optimal enzyme activity supports a cytosolic localization as opposed to the lysosomal localization of DPP1 and DPP3. Despite the fact that DPP8 and DPP9 appear to lack transmembrane domains, it is possible that these intracellular serine proteases are secreted on cellular activation. DPP2/DPP7 (also known as quiescent-cell proline dipeptidase) is targeted against intracellular vesicles that are distinct from lysosomes [45]. This hypothesis expands the potential site(s) and scope of DPP8 and DPP9 involvement in mechanisms for post-translational regulation of chemokines, cytokines, peptides and polypeptides.

Proline-specific peptidases, along with broad-specificity peptidases, are needed for the efficient degradation of peptides containing internal cyclic proline residues. The POPs, first described in 1991, comprise a group of peptidases [POP, DPPIV (DPP4) acylaminoacyl peptidase (EC 3.4.19.1) and oligopeptidase B (EC 3.4.21.83)]. Catalytic activity of these peptidases is provided by a charge-relay system similar to that of the trypsin family of serine proteases, but which evolved by independent convergent evolution [40]. DPP4, perhaps the most well-characterized among the POP proteases, is well documented for its role in regulating neuroendocrine and immunomodulatory peptide hormone activation and inactivation. For example, DPP4 inhibitors prevent DPP4-mediated cleavage of His-Ala or Tyr-Ala dipeptides and inactivation of circulating incretin hormones such as glucagon-like peptides-1 and -2 and gastric inhibitory polypeptide [46]. The enhanced insulin secretion and improved glucose tolerance in several animal models for type II diabetes has led to the development of specific DPP4 inhibitor drugs for type II diabetes [46,47]. As the only well-described, circulating and/or membrane-bound DPP enzyme [48], DPP4 clearly remains a valuable target for the modulation of circulating peptide hormone levels. However, the discovery of molecules, particularly non-peptide compounds capable of penetrating the cell membrane, which specifically inhibit DPP4-like proteins will provide valuable biochemical and pharmacological tools. These and other pharmacological tools are required to improve our understanding of the role that these cytosolic- and membrane-bound enzymes play in the regulation of protein and peptide production and other potential functions mediated by the other domains in these proteins. Previously [16], functional studies, structure–activity analysis and drug design for target proteins were not readily transferable to other potentially related targets. However, the sequencing of the human genome enables paralogous and orthologous proteins to be identified and grouped into families based on sequence similarity, enabling knowledge and expertise to be applied across multiple family members. An improved understanding of the structure–function relationships of the DPP4-like enzymes can improve significantly the future design and specificity of inhibitor molecules targeting these proteins.

We thank Dr Teddy Hsu and Professor Aaron Hsueh for their advice and skilled assistance with the bioinformatics portions of this study.

REFERENCES

1 Deadman, J. (2000) Proteinase inhibitors and activators strategic targets for therapeutic intervention. J. Pept. Sci. **6**, 421–431
2 Southan, C. (2001) A genomic perspective on human proteases as drug targets. Drug Discov. Today **6**, 681–688
3 Mentlein, R. (1999) Dipeptidyl-peptidase IV (CD26) – role in the inactivation of regulatory peptides. Regul. Pept. **85**, 9–24
4 Yaron, A. and Naider, F. (1993) Proline-dependent structural and biological properties of peptides and proteins. Crit. Rev. Biochem. Mol. Biol. **28**, 31–81
5 McCaughan, G. W., Gorrell, M. D., Bishop, G. A., Abbott, C. A., Shackel, N. A., McGuinness, P. H., Levy, M. T., Sharland, A. F., Bowen, D. G., Yu, D. et al. (2000) Molecular pathogenesis of liver disease: an approach to hepatic inflammation, cirrhosis and liver transplant tolerance. Immunol. Rev. **174**, 172–191
6 Dang, N. H. and Morimoto, C. (2002) CD26: an expanding role in immune regulation and cancer. Histol. Histopathol. **17**, 1213–1226
7 Rawlings, N. D., Polgar, L. and Barrett, A. J. (1991) A new family of serine-type peptidases related to prolyl oligopeptidase. Biochem. J. **279**, 907–908
8 Polgar, L. (2002) The prolyl oligopeptidase family. Cell. Mol. Life Sci. **59**, 349–362
9 Sedo, A. and Malik, R. (2001) Dipeptidyl peptidase IV-like molecules: homologous proteins or homologous activities? Biochim. Biophys. Acta **1550**, 107–116
10 Niedermeyer, J., Scanlan, M. J., Garin-Chesa, P., Daiber, C., Fiebig, H. H., Old, L.J., Rettig, W. J. and Schnapp, A. (1997) Mouse fibroblast activation protein: molecular cloning, alternative splicing and expression in the reactive stroma of epithelial cancers. Int. J. Cancer **71**, 383–389
11 Niedermeyer, J., Enenkel, B., Park, J. E., Lenter, M., Rettig, W. J., Damm, K. and Schnapp, A. (1998) Mouse fibroblast-activation protein – conserved Fap gene organization and biochemical function as a serine protease. Eur. J. Biochem. **254**, 650–654
12 Scanlan, M. J., Raj, B. K., Calvo, B., Garin-Chesa, P., Sanz-Moncasi, M. P., Healey, J. H., Old, L. J. and Rettig, W. J. (1994) Molecular cloning of fibroblast activation protein α, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers. Proc. Natl. Acad. Sci. U.S.A. **91**, 5657–5661
13 Mathew, S., Scanlan, M. J., Mohan Raj, B. K., Murty, V. V., Garin-Chesa, P., Old, L. J., Rettig, W. J. and Chaganti, R. S. (1995) The gene for fibroblast activation protein α (FAP), a putative cell surface-bound serine protease expressed in cancer stroma and wound healing, maps to chromosome band 2q23. Genomics **25**, 335–337
14 Levy, M. T., McCaughan, G. W., Abbott, C. A., Park, J. E., Cunningham, A. M., Muller, E., Rettig, W. J. and Gorrell, M. D. (1999) Fibroblast activation protein: a cell surface dipeptidyl peptidase and gelatinase expressed by stellate cells at the tissue remodelling interface in human cirrhosis. Hepatology **29**, 1768–1778
15 Wada, K., Yokotani, N., Hunter, C., Doi, K., Wenthold, R. J. and Shimasaki, S. (1992) Differential expression of two distinct forms of mRNA encoding members of a dipeptidyl aminopeptidase family. Proc. Natl. Acad. Sci. U.S.A. **89**, 197–201
16 Abbott, C. A., Yu, D. M., Woollatt, E., Sutherland, G. R., McCaughan, G. W. and Gorrell, M. D. (2000) Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8. Eur. J. Biochem. **267**, 6140–6150
17 Underwood, R., Chiravuri, M., Lee, H., Schmitz, T., Kabcenell, A. K., Yardley, K. and Huber, B. T. (1999) Sequence, purification, and cloning of an intracellular serine protease, quiescent cell proline dipeptidase. J. Biol. Chem. **274**, 34053–34058
18 Fukasawa, K. M., Fukasawa, K., Higaki, K., Shiina, N., Ohno, M., Ito, S., Otogoto, J. and Ota, N. (2001) Cloning and functional expression of rat kidney dipeptidyl peptidase II. Biochem. J. **353**, 283–290
19 Shneider, B. L., Thevananther, S., Moyer, M. S., Walters, H. C., Rinaldo, P., Devarajan, P., Sun, A. Q., Dawson, P. A. and Ananthanarayanan, M. (1997) Cloning and characterization of a novel peptidase from rat and human ileum. J. Biol. Chem. **272**, 31006–31015

20 Pangalos, M. N., Neefs, J. M., Somers, M., Verhasselt, P., Bekkers, M., van der Helm, L., Fraiponts, E., Ashton, D. and Gordon, R. D. (1999) Isolation and expression of novel human glutamate carboxypeptidases with N-acetylated α-linked acidic dipeptidase and dipeptidyl peptidase IV activity. J. Biol. Chem. **274**, 8470–8483

21 Jacotot, E., Callebaut, C., Blanco, J., Krust, B., Neubert, K., Barth, A. and Hovanessian, A. G. (1996) Dipeptidyl-peptidase IV-β, a novel form of cell-surface-expressed protein with dipeptidyl-peptidase IV activity. Eur. J. Biochem. **239**, 248–258

22 Blanco, J., Jacotot, E., Callebaut, C., Krust, B. and Hovanessian, A. G. (1997) Further characterization of DPP IV-β, a novel cell surface expressed protein with dipeptidyl peptidase activity. Adv. Exp. Med. Biol. **421**, 193–199

23 Blanco, J., Nguyen, C., Callebaut, C., Jacotot, E., Krust, B., Mazaleyrat, J. P., Wakselman, M. and Hovanessian, A.G. (1998) Dipeptidyl-peptidase IV-β – further characterization and comparison to dipeptidyl-peptidase IV activity of CD26. Eur. J. Biochem. **256**, 369–378

24 Duke-Cohan, J. S., Tang, W. and Schlossman, S. Z. (2000) Attractin: a cub-family protease involved in T cell-monocyte/macrophage interactions. Adv. Exp. Med. Biol. **477**, 173–185

25 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402

26 Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A. and Wheeler, D. L. (2000) GenBank. Nucleic Acids Res. **28**, 15–18

27 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673–4680

28 Sonnhammer, E. L., von Heijne, G. and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6**, 175–182

29 Marchler-Bauer, A., Panchenko, A. R., Shoemaker, B. A., Thiessen, P. A., Geer, L. Y. and Bryant, S. H. (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res. **30**, 281–283

30 Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**, 406–425

31 Demuth, H. U., Schlenzig, D., Schierhorn, A., Grosche, G., Chapot-Chartier, M. P. and Gripon, J. C. (1993) Design of (ω-*N*-(*O*-acyl)hydroxy amid) aminodicarboxylic acid pyrrolidides as potent inhibitors of proline-specific peptidases. FEBS Lett. **320**, 23–27

32 Qi, S. Y., Akinsanya, K. O., Riviere, P. J. and Junien, J. L. (2002) Novel serine protease genes related to DPPIV. Patent application WO0231134

33 Olsen, C. and Wagtmann, N. (2002) Identification and characterization of human DPP9, a novel homologue of dipeptidyl peptidase IV. Gene **299**, 185–193

34 Nagase, T., Kikuno, R., Ishikawa, K., Hirosawa, M. and Ohara, O. (2000) Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. DNA Res. **7**, 143–150

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

36 Puente, X. S. and Lopez-Otin, C. (1997) The PLEES proteins: a family of structurally related enzymes widely distributed from bacteria to humans. Biochem. J. **322**, 947–949

37 Kin, Y., Misumi, Y. and Ikehara, Y. (2001) Biosynthesis and characterization of the brain-specific membrane protein DPPX, a dipeptidyl peptidase IV-related protein. J. Biochem. (Tokyo) **129**, 289–295

38 Ogata, S., Misumi, Y. and Ikehara, Y. (1989) Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA and identification of the NH_2-terminal signal sequence as the membrane-anchoring domain. J. Biol. Chem. **264**, 3596–3601

39 Holmquist, M. (2000) α/β-hydrolase fold enzymes: structures, functions and mechanisms. Curr. Protein Pept. Sci. **1**, 209–235

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

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

42 Gorrell, M. D., Abbott, C. A., Kahne, T., Levy, M. T., Church, W. B. and McCaughan, G. W. (2000) Relating structure to function in the β-propellar domain of dipeptidyl peptidase IV. Point mutations that influence adenosine deaminase binding, antibody binding and enzyme activity. Adv. Exp. Med. Biol. **477**, 89–95

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

44 Nadal, M. S., Ozaita, A., Amarillo, Y., de Miera, E. V., Ma, Y., Mo, W., Goldberg, E. M., Misumi, Y., Ikehara, Y., Neubert, T. A. et al. (2003) The CD26-related dipeptidyl aminopeptidase-like protein DPPX is a critical component of neuronal A-type K^+ channels. Neuron **37**, 449–461

45 Chiravuri, M., Agarraberes, F., Mathieu, S. L., Lee, H. and Huber, B. T. (2000) Vesicular localization and characterization of a novel post-proline-cleaving aminodipeptidase, quiescent cell proline dipeptidase. J. Immunol. **165**, 5695–5702

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

47 Sudre, B., Broqua, P., White, R. B., Ashworth, D., Evans, D. M., Haigh, R., Junien, J. L. and Aubert, M. L. (2002) Chronic inhibition of circulating dipeptidyl peptidase IV by FE 999011 delays the occurrence of diabetes in male zucker diabetic fatty rats. Diabetes **51**, 1461–1469

48 Durinx, C., Lambeir, A. M., Bosmans, E., Falmagne, J. B., Berghmans, R., Haemers, A., Scharpe, S. and De, M. I. (2000) Molecular characterization of dipeptidyl peptidase activity in serum: soluble CD26/dipeptidyl peptidase IV is responsible for the release of X-Pro dipeptides. Eur. J. Biochem. **267**, 5608–5613

49 Fukasawa, K., Fukasawa, K. M., Kanai, M., Fujii, S., Hirose, J. and Harada, M. (1998) Dipeptidyl peptidase III is a zinc metallo-exopeptidase. Molecular cloning and expression. Biochem. J. **329**, 275–282

50 Gossrau, R. (1991) Histochemical and biochemical studies of dipeptidyl peptidase I (DPP I) in laboratory rodents. Acta Histochem. **91**, 85–100

Received 10 December 2002/19 March 2003; accepted 28 March 2003
Published as BJ Immediate Publication 28 March 2003, DOI 10.1042/BJ20021914