Crystal structure of human dipeptidyl peptidase IV in complex with a decapeptide reveals details on substrate specificity and tetrahedral intermediate formation

KATHLEEN AERTGEERTS, SHENG YE, MIKE G. TENNANT, MICHELLE L. KRAUS, JOE ROGERS, BI-CHING SANG, ROBERT J. SKENE, DAVID R. WEBB,[1] AND G. SRIDHAR PRASAD

Syrrx Inc., San Diego, California 92121, USA

(RECEIVED September 26, 2003; FINAL REVISION October 25, 2003; ACCEPTED October 27, 2003)

Abstract

Dipeptidyl peptidase IV (DPPIV) is a member of the prolyl oligopeptidase family of serine proteases. DPPIV removes dipeptides from the N terminus of substrates, including many chemokines, neuropeptides, and peptide hormones. Specific inhibition of DPPIV is being investigated in human trials for the treatment of type II diabetes. To understand better the molecular determinants that underlie enzyme catalysis and substrate specificity, we report the crystal structures of DPPIV in the free form and in complex with the first 10 residues of the physiological substrate, Neuropeptide Y (residues 1–10; tNPY). The crystal structure of the free form of the enzyme reveals two potential channels through which substrates could access the active site—a so-called propeller opening, and side opening. The crystal structure of the DPPIV/tNPY complex suggests that bioactive peptides utilize the side opening unique to DPPIV to access the active site. Other structural features in the active site such as the presence of a Glu motif, a well-defined hydrophobic S1 subsite, and minimal long-range interactions explain the substrate recognition and binding properties of DPPIV. Moreover, in the DPPIV/tNPY complex structure, the peptide is not cleaved but trapped in a tetrahedral intermediate that occurs during catalysis. Conformational changes of S630 and H740 between DPPIV in its free form and in complex with tNPY were observed and contribute to the stabilization of the tetrahedral intermediate. Our results facilitate the design of potent, selective small molecule inhibitors of DPPIV that may yield compounds for the development of novel drugs to treat type II diabetes.

Keywords: Dipeptidyl peptidase IV; DPPIV; CD26; crystal structure; adenosine deaminase binding protein; serine protease; tetrahedral intermediate

The type II transmembrane serine protease, DPPIV, also known as CD26, or adenosine deaminase binding protein (ADAbp), is highly expressed on endothelial cells, differentiated epithelial cells and lymphocytes (Hegen et al. 1997; De Meester et al. 1999; Kahne et al. 1999). A soluble form of the enzyme was also found in plasma (Iwaki-Egawa et al. 1998; Durinx et al. 2000). As a dipeptidyl peptidase, DPPIV plays a major role in the regulation of physiological processes including immune, inflammatory, CNS, and endocrine functions. For example, DPPIV plays an important role in maintaining glucose homeostasis (Deacon et al. 1998; Balkan et al. 1999; Pauly et al. 1999; Drucker 2003). These studies reveal that DPPIV helps regulate plasma glucose levels by controlling the activity of the incretins glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP). Inhibition of DPPIV in wild-

Reprint requests to: G. Sridhar Prasad, Syrrx Inc., 10410 Science Center Drive, San Diego, CA 92121, USA; e-mail: Sridhar.Prasad@syrrx.com; fax: (858) 550-0526.

[1]Present address: Celgene Corp., San Diego, CA 92121, USA.

Abbreviations: DPPIV, dipeptidyl peptidase IV; NPY, Neuropeptide Y; tNPY, N-terminal decapeptide (residues 1–10) of Neuropeptide Y.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.03460604.

type and diabetic mice leads to increased levels of unprocessed GLP-1 and GIP in the circulation, enhanced insulin secretion, and improved glucose tolerance. Selective inhibitors of DPPIV improve plasma glucose levels in human type II diabetics (Ahren et al. 2002). Independent of its dipeptidyl peptidase activity, DPPIV also binds adenosine deaminase (ADA; Morrison et al. 1993). This interaction has been shown to modulate immune function (Franco et al. 1998; Morimoto and Schlossman 1998).

The catalytic triad of DPPIV is composed of residues S630, D708, and H740, which are located within the last 140 residues of the C-terminal region (Ogata et al. 1992). The enzyme specifically removes dipeptides from the N terminus of peptide substrates that contain on average 30 residues and have a Pro or Ala in the penultimate position. In addition, a slow release has been observed for dipeptides composed of X-Ser or X-Gly (Bongers et al. 1992; De Meester et al. 1999; Hinke et al. 2000; Lambeir et al. 2002). Physiological peptides recognized by DPPIV that contain this specificity profile at their cleavage site include neuropeptides like neuropeptide Y, circulating peptide hormones like peptide YY, glucagon-like peptides (GLP)-1 and -2, gastric inhibitory peptides, as well as paracrine chemokines like RANTES (De Meester et al. 1999; Mentlein 1999). Catalytic efficiencies for the cleavage by DPPIV of different physiological substrates were determined by mass spectrometry-based protease assays (Lambeir et al. 2001a,b; Zhu et al. 2003). These studies demonstrated that residues surrounding the scissile bond mainly determine the substrate selectivity of DPPIV. However, there is supporting kinetic evidence that nonconserved residues along the entire length of the peptide are involved in long-range interactions that play a role in substrate binding and catalysis (Lambeir et al. 2001a,b, 2002; Zhu et al. 2003).

Crystal structures of DPPIV in complex with several small molecule inhibitors and substrates have been published (Engel et al. 2003; Hiramatsu et al. 2003; Oefner et al. 2003; Rasmussen et al. 2003; Thoma et al. 2003). However, the exact molecular determinants that contribute to the substrate specificity of DPPIV and how substrate peptides access the active site remains unclear. To help understand the function of DPPIV, we crystallized and solved the X-ray crystal structure of the enzyme in both its free form and in the presence of the first 10 residues of Neuropeptide Y. Neuropeptide Y is a physiological substrate of DPPIV widely distributed in the nervous system (Mentlein 1999), and involved in cardiovascular homeostasis and the regulation of insulin release (Ahren 2000; Ghersi et al. 2001). The catalytic efficiciency for N-terminal dipeptide cleavage of Neuropeptide Y by DPPIV is 3.0×10^6 $M^{-1}sec^{-1}$ (Mentlein et al. 1993). The DPPIV/tNPY structure provides direct evidence that the decapeptide accesses the active site through a side opening, unique to DPPIV, and not through the β-propeller opening. The latter mechanism was suggested for the closely related enzyme prolyl oligopeptidase (POP; Fulop et al. 1998, 2000). Our work also provides a detailed understanding of the molecular determinants that contribute to the substrate specificity of DPPIV. Moreover, in the DPPIV/tNPY crystal structure the peptide was trapped in a tetrahedral intermediate, and gives new insight into DPPIV enzyme catalysis. Earlier studies provided evidence for the existence of a tetrahedral intermediate, which was based on structural studies on complexes with small molecule "transition-state analog" inhibitors; ab initio quantum mechanics (QM), molecular mechanics (MM), and molecular dynamics (MD) simulations or combined time-resolved/pH jump crystallographic studies (Wilmouth et al. 2001; Topf et al. 2002a,b). Until now, no direct structural evidence of a single discrete intermediate formed between a physiological substrate and a serine protease has been published.

Results

Structure and domain organization of DPPIV

The crystal structure of the extracellular domain (residues 39–766) of DPPIV was solved to a resolution of 2.1 Å. The structure consists of two domains: an N-terminal 8-bladed β-propeller domain (residues 61–495) and a C-terminal α/β hydrolase domain (Nardini and Dijkstra 1999; residues 39–55 and 497–766; Fig. 1). The propeller domain packs against the hydrolase domain, and the catalytic triad (S630, H740, and D708) is at the interface of the two domains. In vitro catalytic activity of recombinant DPPIV was measured. The catalytic efficiency for the cleavage of the fluorogenic substrate H-Ala-Pro-7-amido-4-trifluromethylcoumarin (Ala-Pro-AFC) by DPPIV is 5.2×10^6 $M^{-1}sec^{-1}$.

The asymmetric unit is composed of two homodimers, the monomers of which are related by a twofold dyad axis (Fig. 1). This dimeric structure correlates with the biologically active form of DPPIV (Bednarczyk et al. 1991; De Meester et al. 1992; Gorrell et al. 2001; Ajami et al. 2003). The overall structures of the monomers are similar with root-mean-square deviations (RMSDs) from 0.64 Å to 0.98 Å for all heavy atoms and from 0.28 Å to 0.56 Å for the Cα atoms. The dimer interface buries a total of 2188 $Å^2$ accessible surface area per monomer and comprises: (1) the last β-strand (β8) of the peptidase central β-sheet, (2) the last two α-helices (αG and αH), (3) the loop between β6 and αE, and (4) the antiparallel β-strand subdomain (β1* and β2*; Fig. 1). β8 mainly contains hydrophobic residues forming hydrophobic interactions at the center of the dimer interface. α-Helix H forms hydrogen bonds with the loop between αG and β8 in the other monomer. The antiparallel β-strand arm formed by β1* and β2* interacts with its related arm, αG, and the loop between β6 and αE in the other monomer.

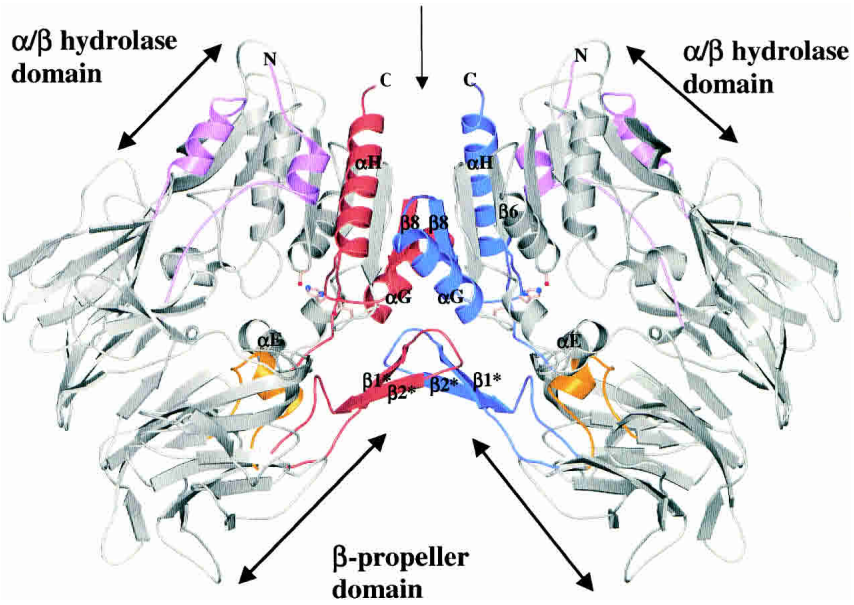

Figure 1. Ribbon diagram showing overall structure of the DPPIV homodimer, viewed perpendicular to the twofold dyad axis. Secondary structural elements that are involved in dimer formation are represented in red and in blue. The active site residues are shown as ball-and-stick representations. The α-helix comprising residues E205 and E206 is indicated in gold. The figure was made using the programs MOLSCRIPT (Kraulis 1991) and Raster3D (Merrit and Bacon 1997).

The primary structure of DPPIV contains nine potential N-linked glycosylation sites: N85, N92, N150, N219, N229, N281, N321, N520, and N685. The first N-acetylglucosamine (NAG) sugar moiety is observed with clear electron density in all the nine predicted sites. Detailed structural and biochemical analysis revealed that the glycosylation of DPPIV is not important for catalytic activity, homodimer formation and ADA binding (Aertgeerts et al. 2004).

Substrate access to the active site

Bioactive peptides recognized by DPPIV could theoretically access the active site in two possible ways: through an opening in the propeller domain or via a side opening formed at the interface of the β-propeller and hydrolase domains (Fig. 2). The propeller opening is formed by the β-propeller domain, which is composed of an unusual eightfold repeat of blades. Each blade is composed of a four-strand antiparallel β-sheet. The β-propeller domain defines a funnel shaped, solvent-filled tunnel that extends from the β-propeller's lower face to the active site. The lower face of the funnel, distal to the hydrolase domain, has a diameter of approximately 15 Å. The closing of the circle between the first and the last blade of propeller proteins has been termed "Velcro" (Neer and Smith 1996), and unlike most of the other known propeller proteins, the "Velcro" is not closed between the first and the last blades in the DPPIV structure. This is similar to the arrangement observed in POP (Fulop et al. 1998). The propeller opening connects to a larger side opening (~21 Å) formed at the interface of the β-propeller domain and the hydrolase domain. This oval-shaped cavity creates a second entrance to the active site (Fig. 2). To understand which entrance/exit pathway substrate peptides use to access the active site of DPPIV, we cocrystallized the enzyme with YPSKPDNPGE (tNPY), corresponding to the first 10 residues of the physiological substrate, Neuropeptide Y. (DPPIV used in the experiment contains a single mutation S716A. The catalytic efficiency of this mutant for cleavage of Ala-Pro-AFC is 41×10^6 M^{-1} sec^{-1}, which is similar to the value measured for wild-type DPPIV. We also obtained crystals of wild-type DPPIV in complex with tNPY, but the crystals using DPPIV-S716A/tNPY diffracted to a higher resolution.) Clear continuous electron density was observed for the first six of the 10 residues of the peptide. Four of the six residues make molecular interactions (see below), with residues lining the side opening of DPPIV. No clear electron density was observed for the last four residues because they are solvent exposed and therefore not ordered in the structure. In conclusion, the crystal structure of the DPPIV/tNPY complex suggests that physiological substrates may employ the side opening of DPPIV to access the active site.

Substrate specificity

DPPIV cleaves the amide bond after the penultimate N-terminal residue (P1, according to Berger and Schechter

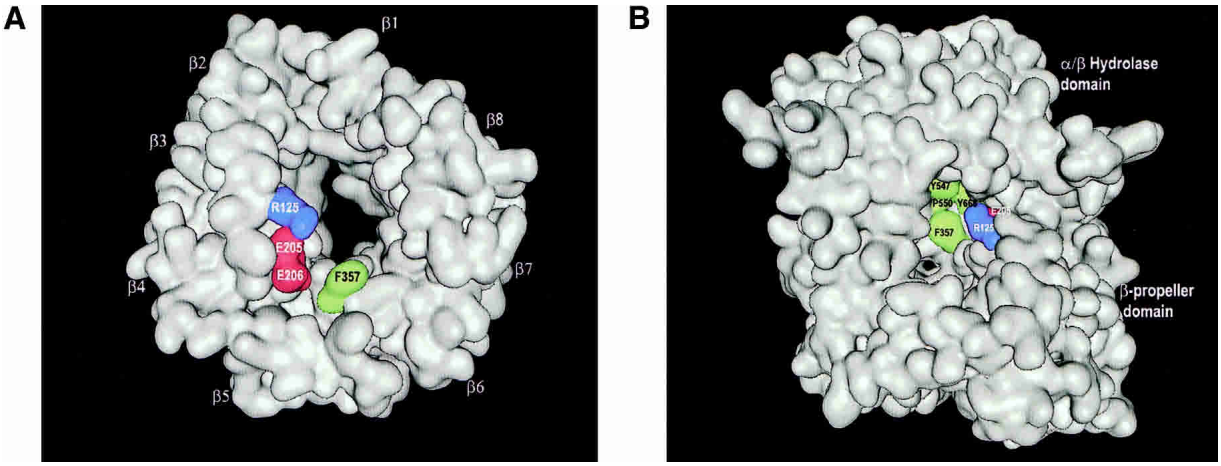

Figure 2. (*A*) Surface representation of the β-propeller domain only, showing the propeller opening to the active site. The view was taken from the interface with the α/β-hydrolase domain and down the pseudo-eightfold axis. The four-strand antiparallel β-sheets of the eight blades are indicated (β1–β8). (*B*) Surface representation of whole DPPIV molecule, showing the side opening to the active site. Residues of DPPIV that make direct molecular interactions with tNPY are colored in both panels. Hydrophobic negatively charged and positively charged residues are shown in green, in red, and in blue, respectively. The figures were made with the program MOE (MOE, Chemical Computing Group).

1970) of physiological peptides. Oligopeptide N termini are recognized by the negatively charged active site residues E205 and E206, and are anchored by hydrogen bond formation with the side chains of the two glutamates (Fig. 3). E205 and E206 reside on a short α-helix insertion (residues 200–206) protruding from the β-propeller domain and pointing toward the active site. The two glutamic acid residues are conformationally restrained by salt bridge formation and hydrogen bond interactions with residues R125, Y662, D663, and N710.

The best catalytic efficiencies for dipeptide cleavage by DPPIV was measured for peptides with a Pro or Ala at P1 (Lambeir et al. 2003; Leiting 2003). The well-defined hydrophobic S1 pocket lined by residues V656, Y631, Y662, W659, Y666, and V711 determines this specificity (Figs. 3, 4). The S2 pocket is hydrophobic and determined by the

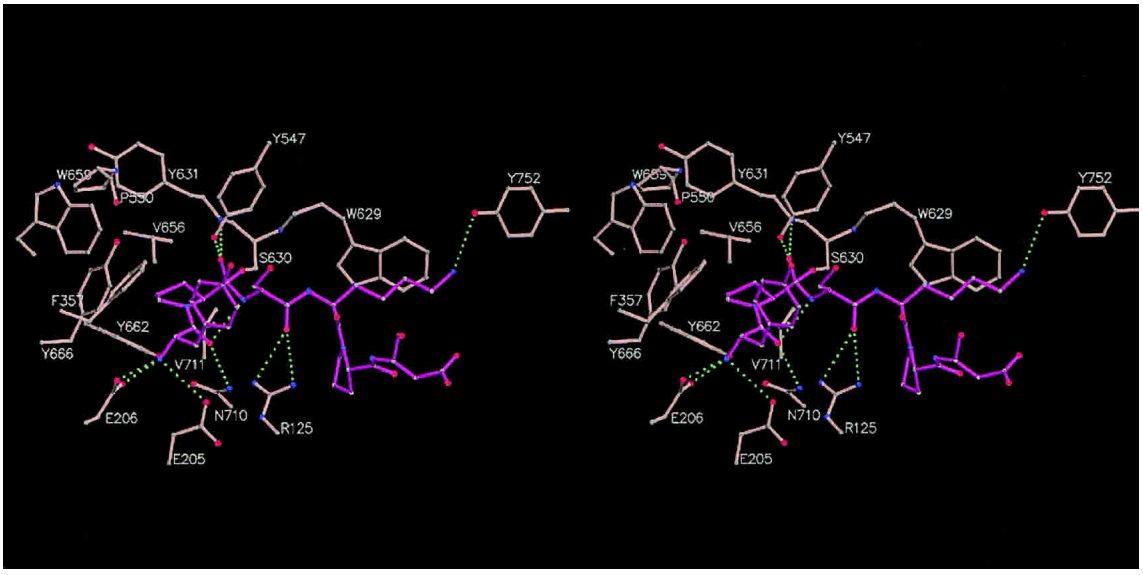

Figure 3. Stereo drawing of first six residues of Neuropeptide Y (magenta) and the underlying active site residues of DPPIV (pink) that make direct molecular interactions with the peptide. The peptide and selected DPPIV residues are shown as ball-and-stick representations. The peptide is not cleaved and trapped in a tetrahedral intermediate by which the carbonyl carbon is covalently linked to the active site S630. Hydrogen bonds are indicated as green dashed lines. The figure is made using the programs MOLSCRIPT (Kraulis 1991) and Raster3D (Merritt and Bacon 1997).

side chains of residues R125, F357, Y547, P550, Y631, and Y666. In this structure we observe two water molecules occupying this S2 pocket, and the P2 tyrosine is interacting with these waters and partially occupying the S2 site. The S1′ pocket is flat and not well defined, and the only interactions observed between the P1′ serine and the S1′ residues are nonspecific van der Waals interactions. The carbonyl oxygen of the P1′ serine makes a hydrogen bond with R125. The side chain of the P2′ lysine packs against the face of W629, completely occluding the tryptophane from solvent, and forms a hydrogen bond with the hydroxyl oxygen of Y752 (Fig. 3). Beyond P2′, no specific interactions are observed between the peptide and the underlying DPPIV residues.

Ser 630 is located on "the nucleophilic elbow" formed by residues Gly-Trp-Ser630-Tyr-Gly. This sequence is essential for DPPIV activity (Ogata et al. 1992) and conserved in the α/β hydrolase family (Gly-X-Ser-X-Gly). The orientation of S630 is maintained by hydrogen bonds between the carbonyl oxygen of S630 and the amide of Y634, and the amide of S630 and the carbonyl oxygen of V653.

Tetrahedral intermediate

In the crystal structure of the DPPIV/tNPY complex, we observed that the peptide was not cleaved, but trapped in a tetrahedral intermediate (Fig. 5). As expected for tetrahedral intermediate formation, the Oγ atom of S630 was found in close contact (between 1.6–1.8 Å) with the carbonyl carbon of the scissile bond. The electron density map contoured at 3σ was continuous between the two atoms, and the electron density map contoured at 1σ was discontinuous between the Oγ atom of S630 and Nδ2 of H740. Comparison of this structure with a 2.1 Å structure of the free form of DPPIV shows that the hydroxyl group of the active site serine (S630) has moved significantly to optimally interact with the carbonyl carbon of the scissile bond (Fig. 5B). In addition, the imidazole ring of H740 rotates by about 15° along the χ^2 torsion (Fig. 5B). The hydrogen bond distance between S630 and H740 in the native enzyme is 2.8 Å, whereas this distance changes to 3.2 Å in the transition state structure. The oxyanion is stabilized by hydrogen bond formation with the main chain amide of Y631 (~3.1 Å) and with the hydroxyl group of Y547 (~2.2 Å; Fig. 5B). Formation of such a short, very strong, low-barrier hydrogen bond is expected in transition states, which stabilizes intermediates in enzymatic reactions and lowers the energy of transition states.

To verify our conclusion that the decapeptide was trapped in a tetrahedral intermediate, the peptide was omitted from the model, the active site S630 was changed to an alanine, and the structure was again refined using REFMAC (CCP 1994). The resultant electron density maps showed unambiguous density for the decapeptide, Oγ atom of S630 and the continuous electron density between the Oγ atom of S630 and the carbonyl carbon of the scissile bond. The asymmetric unit is composed of four independent DPPIV/tNPY complexes, and in all four structures, the peptide is trapped in the tetrahedral intermediate.

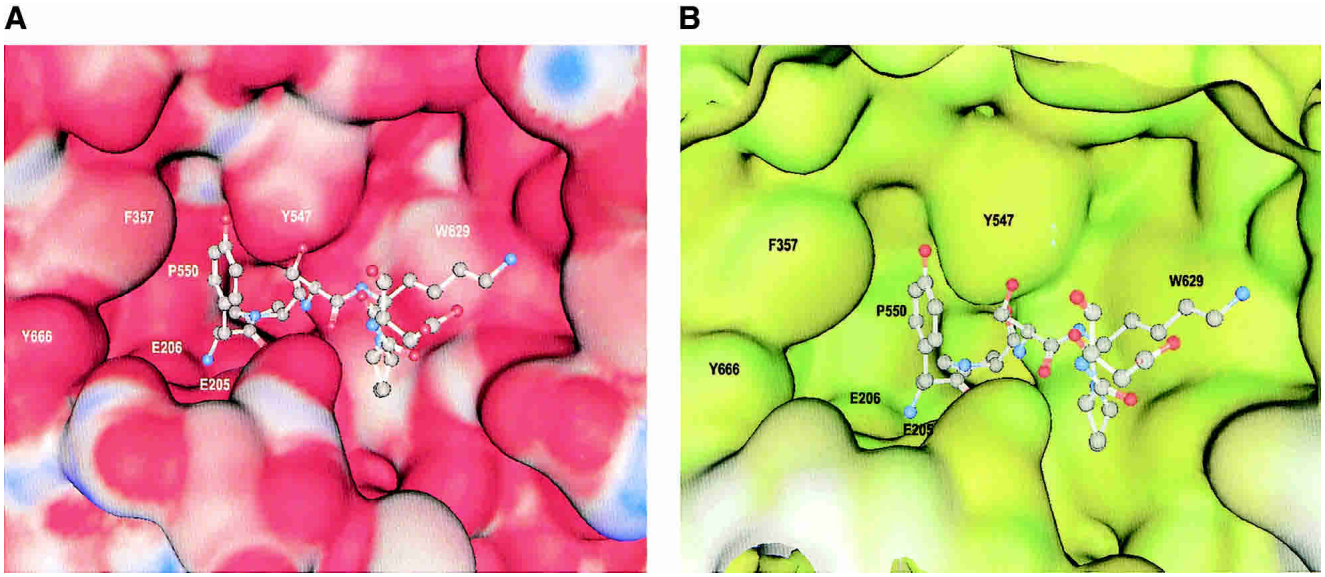

Figure 4. Molecular surface representations showing the interaction of tNPY with DPPIV. Residues of the peptide are shown in ball-and-stick representations and DPPIV is shown as a solid surface. (*A*) Colors represent positive and negative electrostatic potential from blue (electropositive; white, neutral) to red (electronegative). (*B*) Colors represent hydrophobicity (green, polar; yellow, hydrophobic; white, exposed). The figures were made with the program MOE (MOE, Chemical Computing Group).

DOCKET ALARM

Explore Litigation Insights

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

Real-Time Litigation Alerts

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

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

Advanced Docket Research

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

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

Analytics At Your Fingertips

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

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

API

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

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

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

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

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