⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 28, pp. 21487–21495, July 9, 2010
Printed in the U.S.A.

Induced-fit Mechanism for Prolyl Endopeptidase*[S]

Received for publication, December 7, 2009, and in revised form, April 7, 2010 Published, JBC Papers in Press, May 5, 2010, DOI 10.1074/jbc.M109.092692

Min Li[‡], Changqing Chen[§], David R. Davies[‡], and Thang K. Chiu[¶1]

From the [¶]Department of Biochemistry and Molecular Biology, Louisiana State University Health Sciences Center, New Orleans, Louisiana 70112, the [‡]Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, and the [§]Shanghai Research Center of Biotechnology, Chinese Academy of Sciences, Shanghai 200233, China

Prolyl peptidases cleave proteins at proline residues and are of importance for cancer, neurological function, and type II diabetes. Prolyl endopeptidase (PEP) cleaves neuropeptides and is a drug target for neuropsychiatric diseases such as post-traumatic stress disorder, depression, and schizophrenia. Previous structural analyses showing little differences between native and substrate-bound structures have suggested a lock-and-key catalytic mechanism. We now directly demonstrate from seven structures of *Aeromonus punctata* PEP that the mechanism is instead induced fit: the native enzyme exists in a conformationally flexible opened state with a large interdomain opening between the β-propeller and α/β-hydrolase domains; addition of substrate to preformed native crystals induces a large scale conformational change into a closed state with induced-fit adjustments of the active site, and inhibition of this conformational change prevents substrate binding. Absolute sequence conservation among 28 orthologs of residues at the active site and critical residues at the interdomain interface indicates that this mechanism is conserved in all PEPs. This finding has immediate implications for the use of conformationally targeted drug design to improve specificity of inhibition against this family of proline-specific serine proteases.

Prolyl endopeptidase (PEP)[2] is a serine protease that recognizes and cleaves small peptides at internal proline residues. It is involved in the maturation and degradation of peptide hormones and neuropeptides. It is widely distributed in many tissues but is present at two to three times higher levels in the brain (1). Altered levels of serum PEP activity have been observed in patients with post-traumatic stress disorder, depression, mania, schizophrenia, anorexia, and bulimia nervosa (2–4). Administration of the PEP inhibitor S17092 to rats reverses scopolamine-induced amnesia (5, 6) and increases brain levels of the proline-containing peptides substance P, α-melanocyte-stimulating hormone, thyrotrophin-releasing hormone, and arginine-vasopressin (7, 8). Cleavage of bradykinin and angiotensin I-II by PEP suggests that it is involved in blood pressure regulation (9). PEP is thus a significant target for the treatment of numerous diseases and mental disorders.

PEP is structurally related to aminoacyl- (APP), dipeptidyl- (DPP) and tripeptidyl- (TPP) prolyl peptidases (PrPs). These enzymes cleave peptides at proline placed at the first, second and third position from the N terminus, respectively (10–14). DPP4 is the major enzyme that cleaves glucagon-like peptide 1 (GLP1), an incretin peptide that is involved in the regulation of glucose homeostasis. Increasing the level of GLP1 by inhibiting DPP4 activity is currently heavily investigated to treat type II diabetes (15–17), with the two compounds Januvia and Galvus already in clinical use (18, 19). DPP4 also cleaves chemokines and binds adenosine deaminase, suggesting an involvement in lymphocyte chemotaxis and T cell activation (20, 21). Fibroblast activation protein (FAP) α is a DPP that is only expressed in tumor fibroblasts. It has gelatinase and collagenase activities that suggest an involvement in cancer invasion, tumor formation, wound healing, and blood cell disorders (22–24). DPP8 and DPP9 are widely distributed (25, 26). Although definitive biological functions are yet to be identified, they have been implicated in a range of diseases that include diabetes, cancer, arthritis, and asthma (27, 28). There is evidence for their involvement in T cell activation, cell cycle progression, antigen presentation, and inhibition of these two enzymes results in severe toxicity and pathology in rats and dogs (29–32). DPP7 is involved in apoptosis of quiescent T-cells (33). A TPP encoded by *Porphyromonas gingivalis* (*pg*TPP), a bacterial pathogen associated with periondontitis, has the same structure as DPP4. It helps cleave type I collagen fragments containing the $(Gly\text{-}Xaa\text{-}Pro)_n$ sequence at the N terminus (14). These collagen peptide fragments are generated both by host collagenases, as well as by *P. gingivalis*-encoded collagenases, other prolyl-specific peptidases, and cysteine proteases (12).

Previous crystallographic studies of porcine PEP (*p*PEP) have shown that the native and substrate- and inhibitor-bound structures are all identical. The α/β-hydrolase and β-propeller domains are packed together in a closed conformation with

* The use of the Advanced Photon Source was supported by United States Department of Energy, Office of Basic Energy Sciences, Contract W-31-109-Eng-38.

⌘ *Author's Choice*—Final version full access.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables IS and IIS, Figs. 1S and 2S, and Movie 1S.

The atomic coordinates and structure factors (codes 3IUJ, 3IUL, 3IUM, 3IUN, 3IUQ, 3IUR, and 3IVM) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, Louisiana State University Health Sciences Center, 1901 Perdido St., New Orleans, LA 70112. Tel.: 504-568-2024; Fax: 504-568-2093; E-mail: tchiu@lsuhsc.edu or thangchiu@yahoo.com.

[2] The abbreviations used are: PEP, prolyl endopeptidase; *p*PEP, porcine PEP; *ap*PEP, *Aeromonas punctata* PEP; *mx*PEP, *Myoxococcus xanthus* PEP; *sc*PEP, *Sphingomonas capsulata* PEP; PrP, prolyl peptidase; APP, aminoacyl prolyl peptidase; DPP, dipeptidyl prolyl peptidase; TPP, tripeptidyl prolyl peptidase; *pg*TPP, *Porphyromonas gingivalis* TPP; GLP1, glucagon-like peptide 1; FAP, fibroblast activation protein; zPP, benzyloxy-carbonyl-prolyl-prolinal; WT, wild-type; HP35, 35-residue villin headpiece subdomain; H2H3, peptide spanning helices 2 and 3 of HP35; MES, 4-morpholineethanesulfonic acid; PEG, polyethylene glycol; H-bond, hydrogen bond; *E*S, enzyme substrate; *E*A, activated enzyme substrate; x, cross-linked.

only a ~4-Å hole on the surface of the β-propeller that has been proposed to be the route of substrate entry into the ~8500-Å^3 internal cavity where the active site is located (34). The β-propeller fold is observed in many proteins with diverse functions and consists of a four-stranded antiparallel β-sheet repeated in a circular manner four to eight times: seven in PEP and eight in DPP4 (35–38), FAP (39), DPP6 (40), and *pg*TPP (14). The α/β-hydrolase domain is also found in other PrPs for which structures are known (DPP4, FAP, DPP6, *pg*TPP, *Lactococcus lactis* DPP (41), and APP (42), as well as in lipase, acetylcholinesterase, dienelactone hydrolase, and carboxypeptidase II.

The structures of DPP4, DPP6, and FAP determined to date are also in a closed state, with few differences between substrate-free and substrate-bound structures. However, this closed state is different from *p*PEP in that there is an additional side opening between the two domains and the hole at the center of the β-propeller is larger because of the additional repeat. Besides recognition of the positively charged N terminus in DPPs, TPP and APP, the only side chain of the substrate that is recognized by these enzymes is the proline ring. This low sequence specificity makes the design of substrate-specific inhibitors very difficult and is compounded by each enzyme having many potential biological substrates, many of which can be cleaved by multiple members. An approach that has been used in the protein kinase field to improve inhibitor specificity is to target the different inactive conformations of the kinase and its homologs, rather than the conserved ATP binding pocket that all kinases share (43). However, all of the evidence so far for any conformational differences in PrPs has been inferred in PEP: enzyme kinetics studies suggest that a physical step involving conformational change is the rate-limiting step (and not chemical catalysis) (44–47); molecular dynamics simulations suggest that the two domains move in a concerted manner relative to each other (48), keeping the enzyme in the permanently closed state by disulfide cross-linking the two domains across the interdomain interface, which inactivates the enzyme (49); and another study compares the inhibitor-bound structure of *Myoxococcus xanthus* PEP (*mx*PEP-zPP) in the closed state with the substrate-free structure of *Sphingomonas capsulata* PEP (*sc*PEP) in an opened state with a cleft between the two domains (50).

As a prelude to conformationally targeted structure-based drug design for this family of enzymes, we report the crystal structures of *Aeromonas punctata* PEP (*ap*PEP) with the native enzyme in an opened state, and binary complex of the peptide-aldehyde transition-state inhibitor benzyloxy-carbonyl-prolyl-prolinal (zPP) with D622N (the *E*S structure) and WT (the *E*A structure) in the closed state. Addition of zPP to preformed native crystals directly induces large scale interdomain closure and reveals that the enzymatic mechanism is induced fit and not lock and key. Absolute sequence conservation of the active site and critical interdomain interface residues means this mechanism is conserved in all PEPs. The requirement for interdomain opening and closing has direct implications for drug design based on the screening of substrate size and structure. As proof of principle, we show that the highly stable 35-residue peptide HP35 (51) is resistant to PEP cleavage whereas a smaller 22-residue fragment (H2H3) is susceptible.

EXPERIMENTAL PROCEDURES

Protein Expression and Crystallization—The gene for *ap*PEP was PCR-amplified and inserted into vector pET15b between NdeI and BamHI restriction sites for expression as a $MGS_2H_6S_2GLVPRGSH$_*ap*PEP construct in BL21(DE3) cells (pET_PEP_WT). The α/β-hydrolase domain pET_PEP_CAT was generated from pET_PEP_WT using overlapping PCR primers to delete the nucleotides corresponding to the β-propeller domain. Mutants were generated by site-directed mutagenesis using the QuikChange kit (Stratagene). For expression, 2 liters of Luria broth with 0.1 mg/ml ampicillin was inoculated with a 20-ml starter culture and induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside for 4 h at 37 °C when the culture reached A_{600} ~0.6. Selenium-labeled protein was prepared using Novagen B834(DE3) cells grown in minimal medium containing selenomethionine. Cells were harvested by centrifugation at 4000 rpm, resuspended in 50 ml of lysis buffer (20 mM Hepes (pH 7.5), 550 mM NaCl, 5% w/v glycerol, 5 mM imidazole, and 0.2 mg/ml lysozyme), and sonicated. The lysate was clarified by centrifugation at 35,000 rpm for 1 h, and the supernatant was filtered and the enzyme purified by nickel-affinity chromatography, followed by thrombin cleavage, size-exclusion chromatography and concentrating with a Millipore Centriprep YM10 filtration unit to 10 mg/ml in storage buffer (20 mM Hepes (pH 7.5), 100 mM NaCl, 5% w/v glycerol, 1 mM EDTA, and 1 mM dithiothreitol).

After initially screening 960 crystallization conditions with 8 mg/ml enzyme using the Mosquito crystallization robot, large crystals were obtained by streak seeding into preequilibrated hanging drops at 14 °C from precipitants containing 20 mM MES (pH 6.5) and either 15–18% w/v PEG 10,000 or 11–14% w/v PEG 20,000. Native crystals were cryoprotected by transferring them into a 10-μl droplet of reservoir solution and equilibrating for 1–2 days at 4 °C over 1 ml of cryosolution. The cryosolution consists of reservoir solution plus 5% w/v more of PEG and 20% w/v glycerol. After equilibration, crystals were transferred into a small droplet of cryosolution for 30 s before plunging into liquid nitrogen. WT-zPP and D622N-zPP were obtained by soaking cryoequilibrated native crystals in a cryoequilibrated solution containing precipitant plus 10 mM zPP overnight at 4 °C, then cross-linking with glutaraldehyde to ensure that they remain intact during crystal handling. D622Nx-H2H3 was prepared by soaking cross-linked D622N crystals with H2H3 peptide. Cross-linking by vapor diffusion was performed by exposing at 4 °C cryoequilibrated crystals to a 10-μl droplet of 25% glutaraldehyde over 1 ml of cryosolution for 4 h (52).

X-ray Data Collection and Structure Determination—Diffraction data were collected from crystals frozen at 100 K at station SERCAT-22ID of the Argonne National Laboratory and processed with DENZO and SCALEPACK (53). The initial SeMAD structure was phased with SHELX c/d/e, traced with ARP/wARP, and refined with CNS (54–56). D622N mutant and inhibitor-bound structures were determined by molecular replacement using EPMR with separate domains of the SeMAD structure as search models. Models were improved via iterative rounds of manual refitting in O and torsion angle dynamics simulated annealing, positional, and temperature-factor refine-

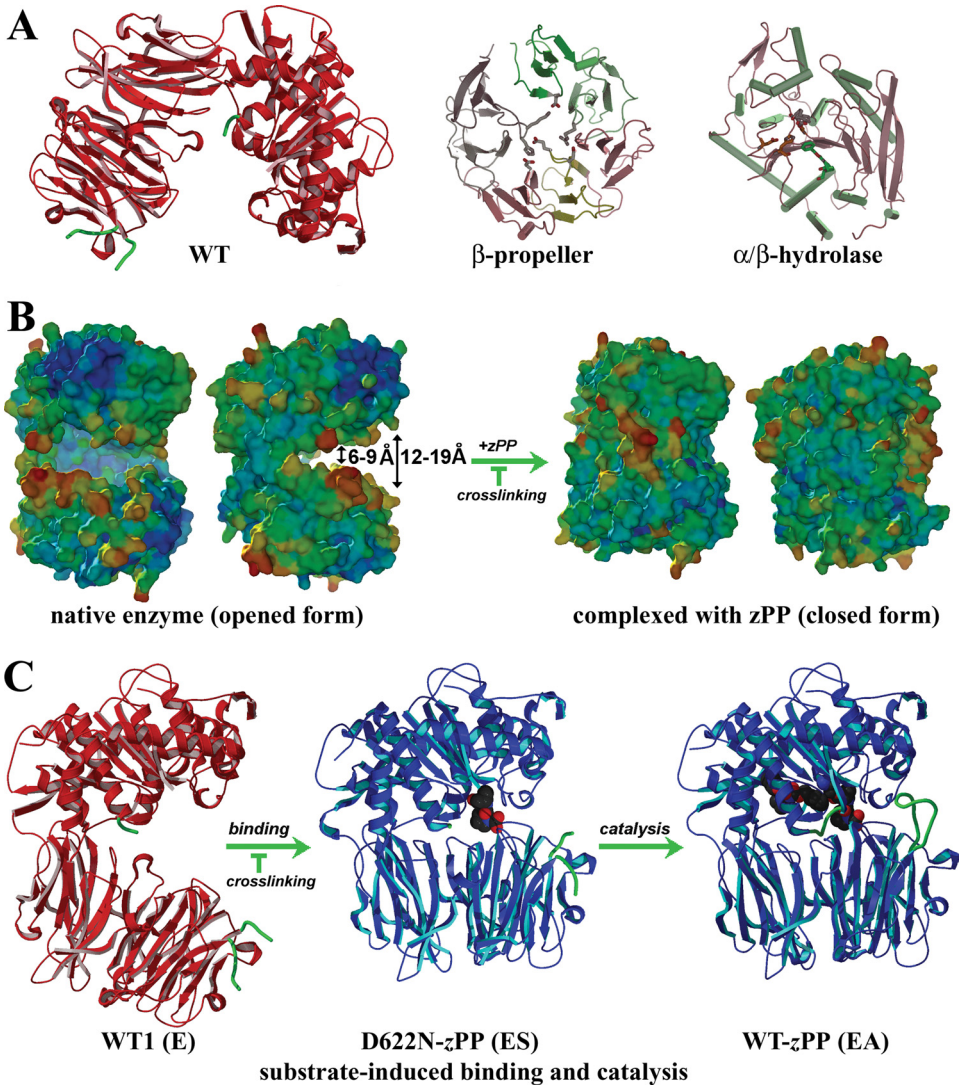


FIGURE 1. **Structures of *ap*PEP in different conformations.** *A*, structure of WT1 (WT2 is similar and not shown) with the two loops (residues 189–206 and 652–662) shown in *green*. The two domains are shown separately on the *right*: salt bridges lining the small 4 Å central pore, the three catalytic residues, and inhibitor molecules bound at the primary (*blue*) and secondary (*green*) sites are *highlighted*. *B*, orthogonal views of the surface of WT2 (similarly for WT1) and WT-zPP colored by temperature factor (*red*, *orange*, *yellow*, *green*, and *blue* representing high to low values, respectively). The opened state is thermally flexible, with interdomain distances varying from crystal to crystal. The most labile patches are located at the interdomain interface and correspond to the two loops that are shown in *green* in *A* and *C*. Inhibition of interdomain closure by glutaraldehyde cross-linking prior to the addition of substrate abolishes inhibitor binding. *C*, addition of the inhibitor zPP to preformed native crystals inducing a conformational change from opened (*red*) to closed (*blue*) states, with disordered residues (195–201 and 654–660) of the flexible loops becoming completely ordered in WT-zPP. Except for these two loops, catalytic and active site residues, the conformational changes involve mostly rigid body rotation with few changes in the secondary structure.

ment in CNS (57, 58). Inhibitor and solvent molecules were selected by examination of $2F_o - F_c$ and $F_o - F_c$ maps contoured at 1.5σ and 3.5σ, respectively. Statistics for the final models are given in supplemental Table IS.

Enzyme Assay—The K_m value of the enzyme for substrate Z-Gly-Pro-*p*NA over a concentration range of 31–500 μM was determined by Lineweaver-Burk plot of $1/V_o$ *versus* 1/[substrate]. The initial velocity V_o was calculated from the fluorescence signal of the *p*NA group that is liberated from the substrate by the enzyme during the initial stage of the reaction. The fluorescence signal was monitored at 410 nm with excitation at 335 nm upon addition of 0.1 ml of 40 nM enzyme to 0.9 ml of substrate in 20 mM Tris-HCl (pH 8) and 100 mM NaCl prewarmed at 34 °C and kept at this temperature for the duration of the measurements. k_{cat}/K_m values were determined from initial rates using the formula $k_{cat}/K_m = V_o/S_oE_o$, which is valid when S_o is much less than K_m. This determination was performed at a minimum of two different concentrations to ensure that the employed substrate concentration (16–62 μM) was far below the K_m. The K_m, k_{cat}, and k_{cat}/K_m of WT *ap*PEP for Z-Gly-Pro-*p*NA were 0.81 mM, 505 s^{-1}, and 623 s^{-1} mM^{-1}, respectively. PEP specificity toward different peptide substrates was assessed in a competitive assay in which 125 μM or 250 μM peptide and different concentrations of Z-Gly-Pro-*p*NA were mixed and reacted with 40 nM *ap*PEP.

RESULTS

Structure Determination—The gene encoding full-length *ap*PEP was PCR-amplified and inserted into vector pET15b for expression and purification from bacterial cells. Crystals of the selenomethionine labeled wild-type native enzyme (WT1 and WT2) and unlabeled active site mutant (D622N) were obtained by streak seeding. They belong in space group $P2_12_12_1$ and have unit cell dimensions $a \approx 63$ Å, $b \approx 85$–95 Å, $c \approx 148$–164 Å, and one protein molecule in the asymmetric unit. The structure was determined by heavy atom phasing from two independent SeMAD datasets and yielded WT1 and WT2 with $R_{work} \approx 19\%$, $R_{free} \approx 22$, and 99% most favored or generously allowed ϕ-ψ values. Except for a few residues at the N terminus (GSHMSGKAR, where GSH is part of the affinity tag that remains after thrombin cleavage), C terminus (QP), and two internal loops (residues 194–201 and 654–661), the rest of the protein residues are well resolved.

Structure of Native Enzyme in an Opened State—In contrast to the closed *p*PEP native enzyme structure (Protein Data Bank entry 1H2W), the native enzyme structure of *ap*PEP is opened (Fig. 1*A*). The α/β-hydrolase (residues 1–75 + 412–690) and β-propeller (residues 76–411) domains share an interface that buries ~2100 $Å^2$ of solvent-accessible surface area. This state is conformationally flexible as judged by the crystal-to-crystal variation of the interdomain separation (6–9 Å and 12–19 Å

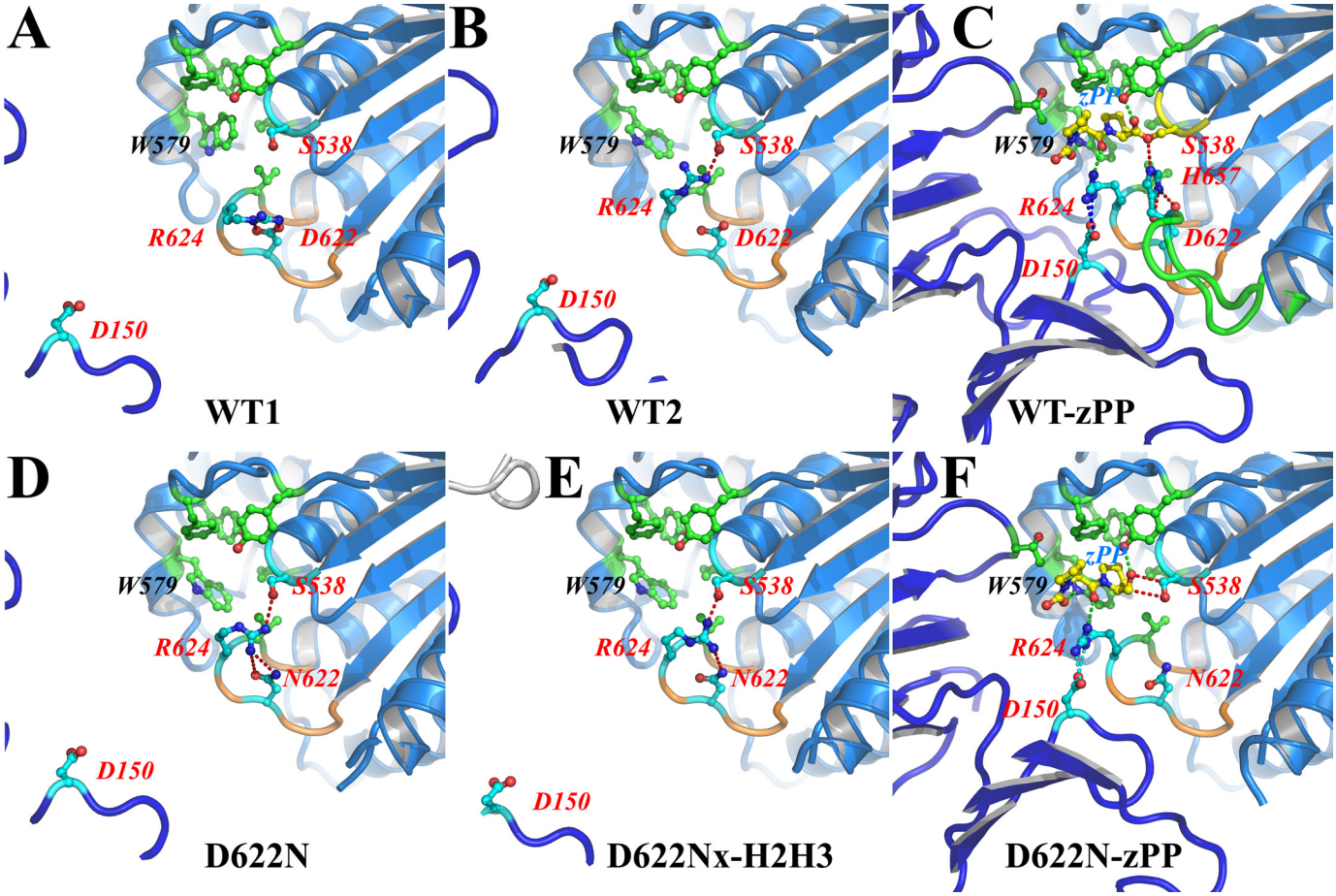


FIGURE 2. **Induced-fit catalytic mechanism in *ap*PEP.** Close-up views of the active site; α/β-hydrolase (*top right*) and β-propeller (*bottom left*) domains are colored *light blue* and *dark blue*, respectively. The active site of WTx is similar to WT2 and not shown. *Dashed lines* represent hydrogen bonds or van der Waals contact between Ser^{539} and the aldehyde group in D622N-zPP. In the native enzyme structures (WT1, WT2, D622N, and D622Nx-H2H3) the hydrophobic pocket (*green*) is preformed for recognizing a proline, the catalytic residues Ser^{538} and D622/N622 are kept in an inactive conformation by H-bonding with Arg^{624}, and the catalytic residue His^{657} is disordered. Addition of the inhibitor zPP (*yellow*) to preformed native crystals induces interdomain closure and brings residue Asp^{150} of the β-propeller domain into the active site to capture Arg^{624} in an orientation that allows it to H-bond the main chain carbonyl oxygen of the inhibitor (D622N-zPP). Upon further interdomain closure (WT-zPP), the catalytic histidine on the flexible loop (*green*) becomes fully ordered and forms the catalytic triad that is required for catalysis. For brevity, the inhibitor/enzyme interaction at the secondary site in WT-zPP is not shown or discussed.

across the inner and outer parts of the interface, respectively; Fig. 1*B*). Two loops (residues 189–206 and 652–662; Fig. 1, *A* and *C*), parts of which are unresolved in the electron density maps (residues 195–201 and 654–660), are inherently flexible and play an important role in substrate binding and catalysis. They are also disordered in the opened native enzyme structure *sc*PEP (residues 231–238 and 689–697 of Protein Data Bank entry 1YR2) (50) but are well ordered in the *p*PEP native enzyme structure in the closed state. The bacterial and porcine enzymes have nearly identical structures in the individual domains (root mean square deviation ~1.5 Å for β-propeller and ~0.9 Å for α/β-hydrolase domains) despite sharing only ~43% sequence identity (~40 and ~58% for the respective domains).

The hydrophobic pocket responsible for the proline recognition is preformed, with most residues except Trp^{579} and Arg^{624} packing against their neighbors in fixed conformations (Fig. 2 and supplemental Table IIS). Trp^{579} is the critical residue that packs with the proline ring of the substrate and in the absence of substrate adopts conformations that vary from structure to structure. Arg^{624} is conformationally flexible and forms hydrogen bonds (H-bonds) with catalytic residues Asp^{622} (WT1), Ser^{538} (WT2 and WTx), or both (D622N and D622Nx-H2H3) to keep them in the inactive conformation. The D622N substitution inactivates the enzyme without affecting the conformation of residue 622. The third member of the catalytic triad (His^{657}) resides on one of the flexible loops and is disordered. Arg^{624} and the three catalytic residues are absolutely conserved in sequence and have the same inactive conformation in the opened native enzyme structure *sc*PEP.

Induced-fit Mechanism for Substrate Entry and Enzyme Catalysis—To show the substrate-induced conformational changes directly, crystals of the native enzyme in the opened state were subsequently soaked with substrates or inhibitors. As shown in Fig. 1, *B* and *C*, supplemental Table IS, supplemental Fig. 1S, and supplemental Movie 1S, these loosely packed crystals (the Matthew's number of 2.9 gives a solvent content of 60%) have a 10–15 Å variation in the *b* and *c* unit cell lengths that reflect the large interdomain distance variation in the native enzyme. This loose packing also tolerates large scale conformational change upon the addition of substrate/inhibitor without altering the crystal packing environ-

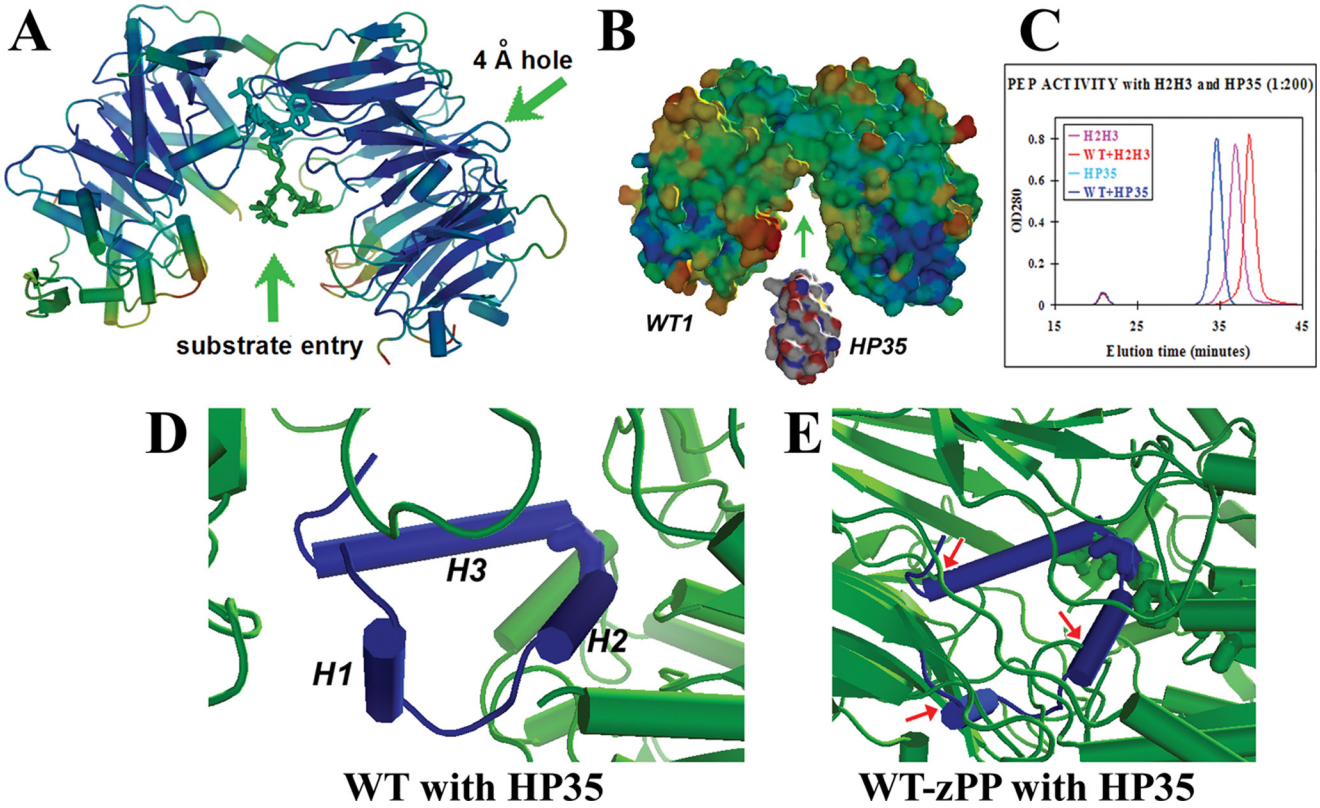


FIGURE 3. **Substrate entry and substrate size selection.** *A*, crystals of D622N that were cross-linked with glutaraldehyde prior to soaking with the substrate H2H3 (D622Nx-H2H3) showed substrate molecules binding nonspecifically at the entry point of the side opening (supplemental Fig. 2S) but not at the tip or interior of the β-propeller or active site. *B*, this model shows that the villin headpiece subdomain HP35 is small enough to get inside PEP while the enzyme is opened. *C*, gel filtration profile of a reaction mixture of PEP (small peak at ~22 min) and HP35 or H2H3 shows that HP35 is resistant to digestion, but H2H3 is completely digested. HP35 is highly stable (T_m = 69 °C) and has an internal proline residue at position 22, whereas H2H3 is unstructured in solution and has an internal proline residue in the middle of the peptide. *D*, this model shows that although HP35 can enter the active site, it is not cleaved because the catalytic residues are in an inactive conformation while the enzyme is opened. *E*, This model of PEP in the closed state (the prolyl rings of H2H3 and zPP are shown) shows that when the catalytic residues are in the active conformation, the internal cavity is too small for HP35 (some of the clashes are shown by *red arrows*). However, it is large enough to accommodate H2H3, which is unstructured in solution and can be reshaped to fit the active site.

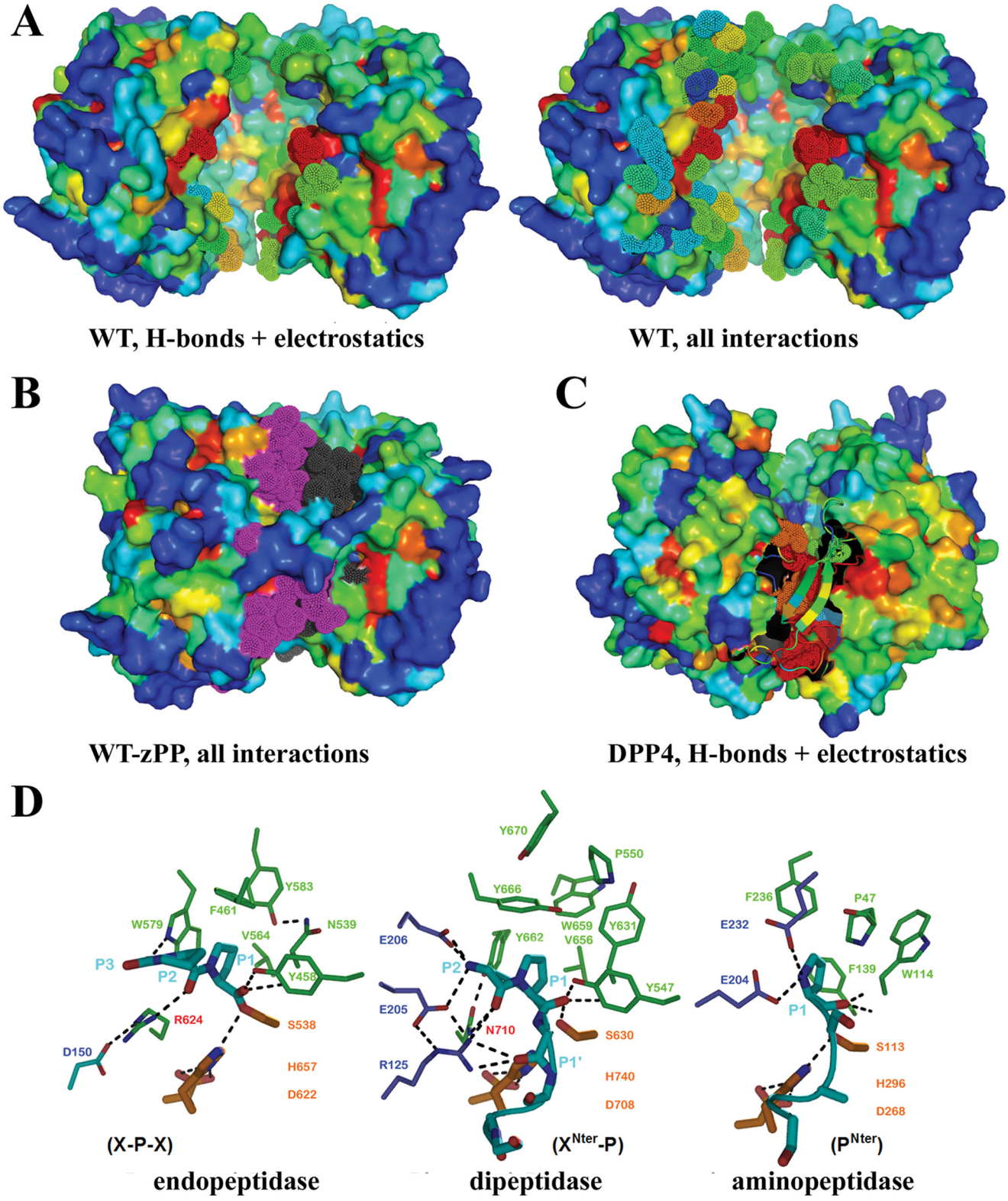
ment (supplemental Fig. 1S and supplemental Movie 1S). This eliminates the possibility that the observed changes were influenced by crystal packing, a possibility had co-crystallization been used to produce co-crystals in a different space group. In addition, the fact that the conformational changes were possible in these crystals means that the crystal packing forces here are energetically very weak, and thus the structures are not biased by crystal packing considerations. It is important to note that although using preformed crystals in soaking experiments may lead to false negatives in that tightly packed molecules may prevent conformational changes or substrate binding, this was not the case in this study as the crystals used were loosely packed.

A surface representation colored by temperature factor (Fig. 1*B*) reveals that the interface between the two domains is highly labile and becomes more ordered when the two domains are brought together by inhibitor binding (Fig. 1, *B* and *C*). With the exception of the ordering of the two loops at this interface, and induced-fit adjustments of the active site (Fig. 2), the conformational changes between native opened and inhibitor-bound closed states involve mostly rigid-body rotation of the two domains with little changes in secondary structure elements (Fig. 1*C* and supplemental Movie 1S). This closure buries an additional 1240 Å^2 (D622N-zPP) to 2667Å^2 (WT-zPP) of solvent-accessible surface area at the interface. Upon closing, the β-propeller domain brings its residue Asp^{150} into the active site of the α/β-hydrolase domain. Its carboxylate group locks the side chain of Arg^{624} in an orientation that allows it to donate an H-bond to the main chain carbonyl oxygen (O9) of the inhibitor (Fig. 2*F*). Mutating this absolutely conserved Asp^{150} residue to an alanine reduces k_{cat}/K_m 2500-fold, presumably because the alanine side chain lacks the H-bond acceptors that lock Arg^{624} in the proper orientation. Mutating it to a leucine results in no measurable activity for presumably the same reason.

The remaining main chain carbonyl oxygens of the inhibitor (O2 and O16) are recognized by H-bonds from the side chains of Tyr^{458} and Trp^{579}, whereas the proline ring at position P1 is recognized by van der Waals stacking with the aromatic ring of Trp^{579}, and to a less extent, by van der Waals contacts with the remaining side chains of the hydrophobic pocket (Fig. 2, *C* and *F*). In the inactive enzyme substrate (*ES*) complex D622N-zPP, the catalytic Ser^{538} residue makes van der Waals contact with the aldehyde group of the inhibitor but is still oriented in the inactive conformation observed in all of the opened structures (WT1, WT2, D622N, and Protein Data Bank entry 1YR2), whereas His^{657} that resides on the flexible loop is unresolved in the electron density maps. In contrast, in the completely closed activated enzyme substrate (E*A*) transition state structure WT-

zPP, the full interaction network of previous closed inhibitor-bound (1H2Y and related entries, and *mx*PEP (2BKL)) and native enzyme (1H2W and related entries) structures (67) is observed. Ser^{538} is in a strained energy state, and His^{657} forms a catalytic triad with Ser^{538} and Asp^{622}. The distinct conformations of Ser^{538} and His^{657} in D622N-zPP and WT-zPP show that these representative *E*S and *E*A structures in the three-step mechanism ($E + S \leftrightarrow ES \leftrightarrow EA \leftrightarrow E + P$) proposed for PEP (59) are distinct.

Cross-linking Prevents Conformational Change, Substrate Binding, and Catalysis—To provide further support for this induced-fit mechanism, preformed crystals in the opened state



were first cross-linked with glutaraldehyde prior to the addition of substrate or inhibitor. This treatment locks molecules within the crystal in a fixed and rigid conformation by covalently linking surface lysine residues that are close to each other (especially between symmetry-related molecules) but otherwise does not alter the enzyme structure as judged from the structures of native (WT1, WT2, D622N) and cross-linked crystals (WTx). D622N crystals that were first cross-linked and then soaked with the substrate H2H3 have a structure identical with that of the untreated native enzyme and no substrate bound at the active site (see D622Nx-H2H3 in next section). This shows that substrate binding both induces and requires a conformational change, which then allows the formation of an active site that permits substrate binding and catalysis.

Cross-linking Confirms a Substrate Entry Point at a Cleft between the Two Domains—The D622Nx-H2H3 structure has two H2H3 molecules located at the entry point of the cleft between the two domains (Fig. 3*A*). The elevated temperature factor of these molecules (61, 30, and 40 Å^2 for H2H3, PEP and solvent molecules, respectively) and electron density for only parts of H2H3 (supplemental Fig. 2S) reflect the nonspecific nature of the binding to this pocket that is facilitated by the packing of two symmetry-related PEP molecules. No substrate molecule was observed at the exterior tip of the β-propeller domain where the 4 Å hole is located, at the interior cavity of the β-propeller domain, or at the active site. These structures thus show that although substrate can gain access to the active site while the enzyme is opened, substrate binding there is not possible because the active site cannot fully form until interdomain closure brings residue Asp^{150} of the β-propeller domain in close proximity to lock residue Arg^{624} of the α/β-hydrolase domain in an orientation that permits substrate binding.

Substrate Size Selection—PEP cleaves internal proline residues in peptides up to ~30 amino acids but not in proteins greater than this size (60, 61). To investigate the structural basis for substrate size selection, we tested the ability of *ap*PEP to cleave HP35 and H2H3. HP35 is a peptide fragment of the villin headpiece subdomain that folds quickly into a highly stable compact α-helical domain without the need for ligands or ions, with an engineered mutant being the fastest folding protein to date (62). HP35 has a proline residue at position 22. The smaller fragment H2H3 spans the second and third helices of HP35 (the C-terminal 24 residues), has a proline residue at position 10, and is unstructured in solution as determined by CD. A 200-fold excess of each peptide was incubated with PEP for 1 h, and then the mixture was fractionated on an analytical Superdex 75 gel filtration column. The elution profile shows that H2H3 but not HP35 is completely cleaved (Fig. 3*C*). A structural explanation for this is that although HP35 can theoretically enter the internal cavity and gain access to the active site while the enzyme is opened (Fig. 3*B*), cleavage does not occur because the catalytic residues are in an inactive conformation (Fig. 3*D*). In contrast, when the enzyme is in the closed state the internal cavity is too small for HP35 because HP35 is highly stable (T_m = 69 °C) and cannot be unfolded to fit inside (Fig. 3*E*). This is not the case for H2H3, which is two-thirds the size and unstructured in solution.

DISCUSSION

Conserved Induced-fit Mechanism for All PEPs—The initial insight that a conformational change is necessary in the catalytic mechanism of PEP was obtained in the early 1990s when the Polgár group showed that changing the identity the substrate's leaving group has little effect on enzyme activity. They interpreted this to mean that a physical rather than a chemical step is rate-limiting (44–47). However, the *p*PEP enzyme which they used in all subsequent crystallographic studies could only be crystallized in the closed state and showed little structural differences between substrate-free and substrate-bound forms (34). A subsequent structure of an inactive mutant, where residue T597C forms a disulfide bond with residue Cys^{255} across the interface to keep the enzyme permanently closed, provides indirect evidence that a conformational change (*e.g.* interdomain opening and closing) is required for activity (49). The Khosla group was the first to crystallize a PEP in an opened conformation, that from the bacteria *S. capsulata* (*sc*PEP) (50). However, this opened conformation may have been the result of an intrusion of a 10-residue His$_6$ tag C-terminal helix from one symmetry-related molecule into the interdomain cavity of another (supplemental Fig. 1S). They inferred that a conformational change is required for catalysis by comparing that structure with that of the enzyme·zPP complex from *M. xanthus* (*mx*PEP-zPP) in the closed state. Although both crystals are $P2_1$, the unit cell dimensions are very different, resulting in two molecules in the asymmetric unit in *mx*PEP-zPP and one in *sc*PEP and completely different crystal packing (50). Given that we have crystallized an orthologous enzyme in yet another different crystal packing environment, how do we

FIGURE 4. **Conserved induced-fit mechanism in all PEP.** *A–C*, sequence conservation among 28 aligned PEPs is mapped onto the surface of *ap*PEP, with *red*, *orange*, *yellow*, *green*, *cyan*, and *blue* representing absolute to low sequence identity, respectively. Note that the sequence is most conserved at the interdomain interface and least conserved at the extremities. *A*, residues at the interface that are involved in interdomain H-bonds and electrostatic interactions (*left*) plus van der Waals contacts (*right*) when the enzyme is closed are depicted as *spiked balls*. Note that residues that are required for proper formation of the active site (including Asp^{150} and Arg^{624}) are spread across two domains and absolutely conserved in sequence, indicating that the induced-fit mechanism of *ap*PEP is conserved in all PEPs. *B*, interactions of *A* are shown for WT-zPP, with residues of different domains colored in *black* or *magenta*. Note that the *floppy loop* that covers the entrance of the two domains in the closed state (*green* in Fig. 1*C*) is not conserved. *C*, same analysis shows strong sequence conservation in DPP4, suggesting that it also uses an induced-fit mechanism. To reveal the underlying interface, the loop overlaying the interface is rendered in *ribbon*. *D*, the three subfamily members have a common recognition mechanism where residues from two separate domains must be close together for the active site to form, suggesting that the induced-fit mechanism is conserved across the family. All members have a hydrophobic pocket (*green*) for recognizing the prolyl ring, noncatalytic residues for recognizing with the substrate main chain atoms (*dashed lines* depict H-bonds), and a Ser/His/Asp catalytic triad. Discrimination of the position of the proline within the substrate is determined by Arg^{624}, recognizing the main chain carbonyl oxygen in PEP (WT-zPP) and by glutamates recognizing the positively charged N terminus in DPP4 (1R9N) and APP (model based on 1QTR). Discrimination of the first *versus* second proline from the N terminus is by the distance of the β-propeller domain relative to the α/β-hydrolase domain, which is farther away in DPP4 than in APP. TPP has a structure almost identical to DPP4, but cleaves the third instead of second proline from the N terminus because its two domains are separated further than DPP4. This shows that inhibitors can be designed against different classes of prolyl peptidases. The structures shown are our 3IUL (*A*) and 3IVM (endopeptidase in *B* and *D*), 1R9N (dipeptidase in *C* and *D*) (65), and 1QTR (aminopeptidase in *D*) (66).

know that the induced-fit mechanism we have shown for *ap*PEP in conserved in all PEPs?

There are several reasons why this induced-fit mechanism is real and universally conserved in all PEPs. It is real in our study because the conformational change from opened to closed states is directly caused by the addition of substrate/inhibitor to preformed crystals, and when this change is physically prevented by glutaraldehyde pretreatment of the native crystals, no substrate binding at the active site is observed. It is both consistent with and explains previous kinetic studies which proposed that the enzyme exists in two different conformational states (44–47). It is also consistent with molecular dynamics simulations showing domain movement upon substrate binding (48). It is universally conserved in all PEPs because the active site residues are absolutely conserved: among 28 PEP orthologs, the identity of residues at positions 458, 461, 564, 579, and 583 are 28Y, 28F, 28V, 28W, and 25Y/3F, respectively. In addition, the residues involved in the interdomain interactions, which include Asp^{150} from the β-propeller domain and Arg^{624} from the α/β-hydrolase domain, are highly conserved whereas those at periphery that have no role in catalysis are not (Fig. 4*A*). We thus can conclude that this induced-fit mechanism applies to all PEPs and that the crystallization of the opened or closed forms of this conformationally flexible enzyme is the result of inherent differences in amino acid sequence at the protein surface (Fig. 4*A*) that favor the fortuitous formation of different crystal contacts (supplemental Fig. 1S and supplemental Movie 1S) and not because the mechanism is different among PEPs. Indeed, the crystallization of orthologous enzymes as a strategy to capture their different conformational states has been used for the enzyme lipase. Lipase has a very similar α/β-hydrolase domain to PEP and an internal substrate binding pocket that is either solvent-exposed or occluded by a small secondary domain. The conformation of the second domain depends on the crystallization conditions and species identity (63).

Mechanistic Implications for Other PrP Members—All of the structures of DPP4 solved to date, with or without substrate or inhibitor, are identical to each other and in the closed state but with a side opening between the two domains. This suggests that unlike PEP, the catalytic mechanism in DPP4 does not involve substrate-induced interdomain closure from a yet to be identified opened state. However, as was the case for *p*PEP before our current work, lack of evidence does not necessarily mean that a conformationally based mechanism is not possible for DPP4. Although the participation of residues from both domains in the active site has been shown in all previous closed structures of PrPs, we have now shown for PEP that these residues are distantly positioned from each other in the native enzyme in an opened state and must be brought together by substrate-induced interdomain closure for the active site to form properly for substrate binding and catalysis (Fig. 2). A similar situation, the participation of residues Arg^{125}, Glu^{205}, and Glu^{206} from the β-propeller domain in the active site and conservation of the interdomain interface, exists in DPP4 (Fig. 4, *C* and *D*). However, whether or not they are also positioned far away from the active site in an opened state and then are brought into the active site of the α/β-hydrolase domain by substrate-induced interdomain closure remains to be determined. Similarly, the case for other DPPs, TPP, and APP remains to be determined. Note in APP that the noncatalytic domain being all α-helical instead of a β-propeller does not necessarily negate the possibility that its catalytic mechanism is similar to PEP. The architectural arrangement of its noncatalytic domain is similar to the enzyme lipase, which adopts different conformations depending on the crystalline environment (63).

Limitations of Current Inhibitors—Current drug design efforts against PrPs have focused on the limited substrate specificity where selection does not extend beyond a dipeptide motif (*i.e.* interactions with the proline ring, one residue to the N- or C-terminal side, plus a positive charge for DPP inhibitors). Although these efforts have yielded inhibitors that can discriminate among the different members of this enzyme family, there is currently no inhibitor that can target a specific substrate, a level of specificity that is ultimately required because of the existence of multiple substrates for each enzyme and multiple enzymes capable of cleaving the same substrate(s). For example, although Januvia, Galvus, and other DPP4 inhibitors in clinical trials have been shown to inhibit the processing of GLP1 by DPP4 (15–17), they are not GLP1-specific, and it is not known whether inhibition of DPP4 by these compounds has other physiological consequences. DPP4 also cleaves, with even higher efficiency than GLP1, stromal cell-derived factor 1α (which is involved in stem cell homing, engraftment, and tissue regeneration), monocyte-derived chemokine, gastrin-releasing peptide 3–27, interferon-inducible T cell α, neuropeptide Y (which is involved in neurogenic inflammation and blood pressure regulation), monocyte-derived chemokine 3–69, IP10, and Mig (64). Although no significant negative side effects of Januvia and Galvus have been reported (18, 19), these drugs act transiently, and follow-up studies will be needed to ensure the safety of their long term usage.

CONCLUSION

We have crystallized an opened form of a PEP native enzyme. By adding substrate or inhibitor to induce a conformational change in the enzyme in preformed crystals directly, we have for the first time revealed that the catalytic mechanism involves both large scale conformational change from opened to closed states and induced-fit formation of the active site for substrate binding and catalysis. Inhibiting the conformational change abolishes substrate binding at the active site. Substrate entry and exit occur via the interdomain opening. This mechanism has remained conserved throughout the evolution of this enzyme. Whether or not it is utilized by the entire enzyme family remains to be determined.

Acknowledgments—We thank Dr. Wei Yang for use of the Mosquito crystallization robot and Drs. David Worthylake, Arthur Haas, and Alison Hickman for constructive comments on the manuscript. All figures of protein structure were made with PyMOL (Schrödinger, LLC).

REFERENCES

1. Irazusta, J., Larrinaga, G., González-Maeso, J., Gil, J., Meana, J. J., and Casis, L. (2002) *Neurochem. Int.* **40,** 337–345
2. Maes, M., Goossens, F., Scharpé, S., Calabrese, J., Desnyder, R., and Meltzer, H. Y. (1995) *Psychiatry Res.* **58,** 217–225
3. Maes, M., Lin, A. H., Bonaccorso, S., Goossens, F., Van Gastel, A., Pioli, R., Delmeire, L., and Scharpe, S. (1999) *J. Affect. Disord.* **53,** 27–34
4. Maes, M., Monteleone, P., Bencivenga, R., Goossens, F., Maj, M., van West, D., Bosmans, E., and Scharpe, S. (2001) *Psychoneuroendocrinology* **26,** 17–26
5. Atack, J. R., Suman-Chauhan, N., Dawson, G., and Kulagowski, J. J. (1991) *Eur. J. Pharmacol.* **205,** 157–163
6. Portevin, B., Benoist, A., Rémond, G., Hervé, Y., Vincent, M., Lepagnol, J., and De Nanteuil, G. (1996) *J. Med. Chem.* **39,** 2379–2391
7. Bellemère, G., Morain, P., Vaudry, H., and Jégou, S. (2003) *J. Neurochem.* **84,** 919–929
8. Bellemère, G., Vaudry, H., Morain, P., and Jégou, S. (2005) *J. Neuroendocrinol.* **17,** 306–313
9. Welches, W. R., Brosnihan, K. B., and Ferrario, C. M. (1993) *Life Sci.* **52,** 1461–1480
10. Ito, K., Inoue, T., Kabashima, T., Kanada, N., Huang, H. S., Ma, X., Azmi, N., Azab, E., and Yoshimoto, T. (2000) *J Biochem.* **128,** 673–678
11. de Meester, I., Lambeir, A. M., Proost, P., and Scharpé, S. (2003) *Adv. Exp. Med. Biol.* **524,** 3–17
12. Banbula, A., Mak, P., Bugno, M., Silberring, J., Dubin, A., Nelson, D., Travis, J., and Potempa, J. (1999) *J. Biol. Chem.* **274,** 9246–9252
13. Mentlein, R. (1999) *Regul. Pept.* **85,** 9–24
14. Ito, K., Nakajima, Y., Xu, Y., Yamada, N., Onohara, Y., Ito, T., Matsubara, F., Kabashima, T., Nakayama, K., and Yoshimoto, T. (2006) *J. Mol. Biol.* **362,** 228–240
15. Ahrén, B., Landin-Olsson, M., Jansson, P. A., Svensson, M., Holmes, D., and Schweizer, A. (2004) *J. Clin. Endocrinol. Metab.* **89,** 2078–2084
16. Drucker, D. J. (2003) *Expert Opin. Investig. Drugs* **12,** 87–100
17. Holst, J. J. (2003) *Adv. Exp. Med. Biol.* **524,** 263–279
18. Aschner, P., Kipnes, M. S., Lunceford, J. K., Sanchez, M., Mickel, C., and Williams-Herman, D. E. (2006) *Diabetes Care* **29,** 2632–2637
19. Ristic, S., Byiers, S., Foley, J., and Holmes, D. (2005) *Diabetes Obes. Metab.* **7,** 692–698
20. Ludwig, A., Schiemann, F., Mentlein, R., Lindner, B., and Brandt, E. (2002) *J. Leukocyte Biol.* **72,** 183–191
21. Richard, E., Alam, S. M., Arredondo-Vega, F. X., Patel, D. D., and Hershfield, M. S. (2002) *J. Biol. Chem.* **277,** 19720–19726
22. Chen, W. T. (2003) *Adv. Exp. Med. Biol.* **524,** 197–203
23. Park, J. E., Lenter, M. C., Zimmermann, R. N., Garin-Chesa, P., Old, L. J., and Rettig, W. J. (1999) *J. Biol. Chem.* **274,** 36505–36512
24. Jones, B., Adams, S., Miller, G. T., Jesson, M. I., Watanabe, T., and Wallner, B. P. (2003) *Blood* **102,** 1641–1648
25. Abbott, C. A., Yu, D. M., Woollatt, E., Sutherland, G. R., McCaughan, G. W., and Gorrell, M. D. (2000) *Eur. J. Biochem.* **267,** 6140–6150
26. Olsen, C., and Wagtmann, N. (2002) *Gene* **299,** 185–193
27. Pitman, M. R., Sulda, M. L., Kuss, B., and Abbott, C. A. (2009) *Front. Biosci.* **14,** 3619–3633
28. Reinhold, D., Bank, U., Täger, M., Ansorge, S., Wrenger, S., Thielitz, A., Lendeckel, U., Faust, J., Neubert, K., and Brocke, S. (2008) *Front. Biosci.* **13,** 2356–2363
29. Ansorge, S., Bank, U., Heimburg, A., Helmuth, M., Koch, G., Tadje, J., Lendeckel, U., Wolke, C., Neubert, K., Faust, J., Fuchs, P., Reinhold, D., Thielitz, A., and Täger, M. (2009) *Clin. Chem. Lab. Med.* **47,** 253–261
30. Geiss-Friedlander, R., Parmentier, N., Möller, U., Urlaub, H., Van den Eynde, B. J., and Melchior, F. (2009) *J. Biol. Chem.* **284,** 27211–27219
31. Lee, H. J., Chen, Y. S., Chou, C. Y., Chien, C. H., Lin, C. H., Chang, G. G., and Chen, X. (2006) *J. Biol. Chem.* **281,** 38653–38662
32. Lankas, G. R., Leiting, B., Roy, R. S., Eiermann, G. J., Beconi, M. G., Biftu, T., Chan, C. C., Edmondson, S., Feeney, W. P., He, H., Ippolito, D. E., Kim, D., Lyons, K. A., Ok, H. O., Patel, R. A., Petrov, A. N., Pryor, K. A., Qian, X., Reigle, L., Woods, A., Wu, J. K., Zaller, D., Zhang, X., Zhu, L., Weber, A. E., and Thornberry, N. A. (2005) *Diabetes* **54,** 2988–2994
33. Chiravuri, M., Schmitz, T., Yardley, K., Underwood, R., Dayal, Y., and Huber, B. T. (1999) *J. Immunol.* **163,** 3092–3099
34. Fülöp, V., Böcskei, Z., and Polgár, L. (1998) *Cell* **94,** 161–170
35. Rasmussen, H. B., Branner, S., Wiberg, F. C., and Wagtmann, N. (2003) *Nat. Struct. Biol.* **10,** 19–25
36. Engel, M., Hoffmann, T., Wagner, L., Wermann, M., Heiser, U., Kiefersauer, R., Huber, R., Bode, W., Demuth, H. U., and Brandstetter, H. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100,** 5063–5068
37. Thoma, R., Löffler, B., Stihle, M., Huber, W., Ruf, A., and Hennig, M. (2003) *Structure* **11,** 947–959
38. Hiramatsu, H., Kyono, K., Higashiyama, Y., Fukushima, C., Shima, H., Sugiyama, S., Inaka, K., Yamamoto, A., and Shimizu, R. (2003) *Biochem. Biophys. Res. Commun.* **302,** 849–854
39. Aertgeerts, K., Levin, I., Shi, L., Snell, G. P., Jennings, A., Prasad, G. S., Zhang, Y., Kraus, M. L., Salakian, S., Sridhar, V., Wijnands, R., and Tennant, M. G. (2005) *J. Biol. Chem.* **280,** 19441–19444
40. Strop, P., Bankovich, A. J., Hansen, K. C., Garcia, K. C., and Brunger, A. T. (2004) *J. Mol. Biol.* **343,** 1055–1065
41. Rigolet, P., Mechin, I., Delage, M. M., and Chich, J. F. (2002) *Structure* **10,** 1383–1394
42. Medrano, F. J., Alonso, J., García, J. L., Romero, A., Bode, W., and Gomis-Rüth, F. X. (1998) *EMBO J.* **17,** 1–9
43. Huse, M., and Kuriyan, J. (2002) *Cell* **109,** 275–282
44. Polgár, L. (1992) *Biochem. J.* **283,** 647–648
45. Polgár, L. (1991) *Biomed. Biochim. Acta* **50,** 721–726
46. Polgár, L., Kollt, E., and Hollósi, M. (1993) *FEBS Lett.* **322,** 227–230
47. Polgár, L. (1991) *Eur. J. Biochem.* **197,** 441–447
48. Fuxreiter, M., Magyar, C., Juhász, T., Szeltner, Z., Polgár, L., and Simon, I. (2005) *Proteins* **60,** 504–512
49. Szeltner, Z., Rea, D., Juhász, T., Renner, V., Fülöp, V., and Polgár, L. (2004) *J. Mol. Biol.* **340,** 627–637
50. Shan, L., Mathews, II, and Khosla, C. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 3599–3604
51. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J., and Davies, D. R. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 7517–7522
52. Lusty, C. J. (1999) *J. Appl. Crystallogr.* **32,** 106–112
53. Otwinowski, Z., and Minor., W. (1997) in *Methods in Enzymology* (Carter, C. W. Jr., and Sweet, R. M., eds) Vol. 276, pp. 307–326, Academic Press, New York
54. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1772–1779
55. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
56. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
57. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 484–491
58. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
59. Walter, R., and Yoshimoto, T. (1978) *Biochemistry* **17,** 4139–4144
60. Koida, M., and Walter, R. (1976) *J. Biol. Chem.* **251,** 7593–7599
61. Taylor, W. L., and Dixon, J. E. (1980) *Biochem. Biophys. Res. Commun.* **94,** 9–15
62. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., and Hofrichter, J. (2006) *J. Mol. Biol.* **359,** 546–553
63. Schrag, J. D., Li, Y., Cygler, M., Lang, D., Burgdorf, T., Hecht, H. J., Schmid, R., Schomburg, D., Rydel, T. J., Oliver, J. D., Strickland, L. C., Dunaway, C. M., Larson, S. B., Day, J., and McPherson, A. (1997) *Structure* **5,** 187–202
64. Lambeir, A. M., Proost, P., Scharpé, S., and De Meester, I. (2002) *Biochem. Pharmacol.* **64,** 1753–1756
65. Aertgeerts, K., Ye, S., Tennant, M. G., Kraus, M. L., Rogers, J., Sang, B. C., Skene, R. J., Webb, D. R., and Prasad, G. S. (2004) *Protein Sci.* **13,** 412–421
66. Yoshimoto, T., Kabashima, T., Uchikawa, K., Inoue, T., Tanaka, N., Nakamura, K. T., Tsuru, M., and Ito, K. (1999) *J Biochem.* **126,** 559–565
67. Szeltner, Z., Rea, D., Renner, V., Fulop, V., and Polgár, L. (2002) *J. Biol. Chem.* **277,** 42613–42622