Article

Unveiling Prolyl Oligopeptidase Ligand Migration by Comprehensive Computational Techniques

Martin Kotev,[1] Daniel Lecina,[1] Teresa Tarragó,[2] Ernest Giralt,[2,3,*] and Víctor Guallar[1,4,*]
[1]Joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona Supercomputing Center, Barcelona, Spain; [2]Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain; [3]Department of Organic Chemistry, University of Barcelona (UB), Barcelona, Spain; and [4]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

ABSTRACT Prolyl oligopeptidase (POP) is a large 80 kDa protease, which cleaves oligopeptides at the C-terminal side of proline residues and constitutes an important pharmaceutical target. Despite the existence of several crystallographic structures, there is an open debate about migration (entrance and exit) pathways for ligands, and their coupling with protein dynamics. Recent studies have shown the capabilities of molecular dynamics and classical force fields in describing spontaneous binding events and nonbiased ligand migration pathways. Due to POP's size and to the buried nature of its active site, an exhaustive sampling by means of conventional long enough molecular dynamics trajectories is still a nearly impossible task. Such a level of sampling, however, is possible with the breakthrough protein energy landscape exploration technique. Here, we present an exhaustive sampling of POP with a known inhibitor, Z-pro-prolinal. In >3000 trajectories Z-pro-prolinal explores all the accessible surface area, showing multiple entrance events into the large internal cavity through the pore in the β-propeller domain. Moreover, we modeled a natural substrate binding and product release by predicting the entrance of an undecapeptide substrate, followed by manual active site cleavage and nonbiased exit of one of the products (a dipeptide). The product exit shows preference from a flexible 18-amino acid residues loop, pointing to an overall mechanism where entrance and exit occur in different sites.

INTRODUCTION

Prolyl oligopeptidase (POP; EC 3.4.21.26) (also known as prolyl endopeptidase, PREP, or postproline cleaving enzyme) is a serine protease that cleaves postproline bonds in short peptides (1). POP inhibitors might be valuable compounds in a variety of clinical conditions of the brain, such as the cognitive disturbances present in schizophrenia and bipolar affective disorder, as indicated by their neuroprotective and cognition-enhancing effects in experiments with animals (2). For these reasons, a plethora of POP inhibitors have been developed during the last 10 years for treatment of several central nervous system disorders (3,4). Two basic groups of inhibitors have been proposed: forming a covalent bond with the catalytic serine and noncovalent ones. Both of them dock at the same specific proline pocket, the main difference being the presence or lack of chemical groups capable of covalently binding to Ser-554. The development of POP inhibitors, however, has been based almost exclusively on modification of the canonical peptidomimetic compound Z-prolyl-prolinal (ZPP) that fits into the POP active site. This strategy does not take into account other possible POP binding surfaces such as surfaces involved in the entry of substrates and/or exit of products, which may trigger the discovery of innovative peptide scaffolds with biological activity. In addition to its enzymatic role, POP interacts with several proteins, α-synuclein being one of the most relevant. POP accelerates aggregation of α-synuclein in vitro, a process that can be reversed by specific inhibitors (5,6). Moreover, nuclear magnetic resonance spectroscopy studies have revealed that POP is a highly dynamic protein and that active site inhibition shifts this conformational equilibrium toward a less dynamic form (7). POP structural fluctuations and its importance for substrate/inhibitor delivery, however, is centering particular attention (8,9).

The crystal structures of POP indicate two domains, a catalytic one bearing the Ser-His-Asp triad and the so-called β-propeller domain, which covers a huge cavity around the catalytic center (1). Ligand access to this catalytic center, however, is under debate. Two main entry/exit areas have been investigated since release of the first crystal structures of porcine POP 16 years ago. The first one is a pore in the β-propeller domain, whereas the other is a ~18-residue flexible loop (some authors call it loop A (10)) standing close to the active site. The diameter of the pore (distance between two approximately opposite α-carbons) is around 11–13 Å (Fig. 1). This means that appropriate conformational orientations of some side chains in the area could open a passage for some inhibitors or small peptides. Two lysine side chains and two glutamic acid ones form salt bridges, which, together with a few hydrogen bonds reduce the propeller

Submitted June 23, 2014, and accepted for publication November 17, 2014.
*Correspondence: victor.guallar@bsc.es or ernest.giralt@irbbarcelona.org
Editor: Bert de Groot.

0006-3495/15/01/0116/10 $2.00
http://dx.doi.org/10.1016/j.bpj.2014.11.3453

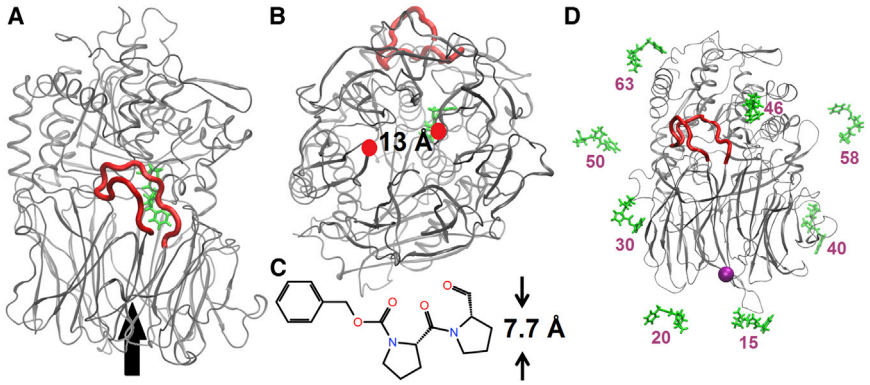

FIGURE 1 (*A*) Side view of the flexible loop (*underlined in red*) with a bound ZPP inhibitor (*green*). The arrow points to the β-propeller pore. (*B*) Bottom view showing also the pore distance between two α-carbons (*red circles*). (*C*) Structure and maximum width of ZPP. (*D*) All eight starting positions of ZPP (in *green*) for the entrance simulations and their approximate distance (*purple*) in Å to the center of the pore (*purple bead*). To see this figure in color, go online.

pore (Protein Data Base (PDB) entry 1QFS). Some experimental data shows forming of a disulfide bridge to block this pore but in this case the bridge is created aside and does not cover the central part of the pore (11). The mobility of the flexible loop, on the other side, has been suggested by trypsin cleavage assays (10). Recent experimental studies, however, question its involvement in ligand delivery. In the work of Szeltner and co-workers (10) a heptadecapeptide is better cleaved from a mutated POP containing a loop covalently locked by a disulfide bridge to the catalytic domain.

Of importance, there is an additional crystal structure from the bacterium *Aeromonas punctata* (12) where the two domains present a large opening, pointing to a clear entrance into the active site. In fact the domains are almost separated and only held by two covalent bonds—the hinge between the domains. However, there is no mammalian crystal structure showing such conformation, neither is there clear experimental proof of this opening (10). Furthermore, a porcine POP crystal structure, 99% similar to the human one, shows clear differences to the bacterial one in the noncovalent forces keeping together the two domains (12). Other studies also suggest that local conformational changes related to some flexible loops but not the entire domains could be responsible for the access to the active site of POP (13).

Computational studies have also addressed POP's dynamics and its possible ligand migration pathways. Molecular dynamics (MD) simulations showed significant loop opening and exposure to the bulk solvent (13). Some authors in their previous study have used steered MD and umbrella sampling simulations to force the inhibitor ZPP exiting from the active site (8). In this work, the inhibitor was pulled in two possible directions: the loop one and to the β-propeller pore. Results show that the exit of ZPP is energetically more favorable through the loop region (8). Docking results and subsequent MD simulations from a docked pose of an inhibitor in the β-propeller pore have shown that the ligand can reduce some distance traveling toward the active site, which depicts potentiality of β-propeller in ingesting a ligand (9). None of the published simulations have shown indications for the interdomain opening. All of them reveal stable closed POP structures during the simulations except for some loop motions (13).

Using special purpose machines or graphical processors units, a nonbiased search accessing microsecond timescale simulations has recently been performed on small or medium systems (14,15). These computational approaches represent a significant computational cost, being still prohibitive when dealing with complex systems (buried active sites) such as POP. To address this issue we have used protein energy landscape exploration (PELE), a novel computational technique capable of exploring the nonbiased ligand diffusion and proteins dynamics (16). PELE combines a Monte Carlo stochastic approach with protein structure prediction algorithms, and it is capable of accurately reproducing long-timescale processes in a 1–2 order of magnitude faster manner than MD (17–20). Such a technological development, together with the use of the supercomputer Mare Nostrum, has allowed us to run 3000 trajectories, for an extensive exploration of ZPP interaction with both mammalian and bacterial POP. Our results indicate that entrance happens mainly through the bottom pore, with only smaller molecules being able to enter through the bacterial opened loop. Furthermore, we simulated the catalytic process of entering an 11-amino acid residue peptide as a substrate and the exiting of one of the two products. This full catalytic event indicates entrance through the β-propeller pore and exit of the cleaved small peptide through the loop area.

MATERIALS AND METHODS

System preparation

Initial coordinates for the closed POP structures were taken from the PDB entries 1QFS (mammalian porcine) (1) and 3IVM (bacterial *A. punctata*) (12). Semiopen bacterial POP coordinates were obtained from the PDB structure 3IUQ (12). Hydrogen atoms and titratable side chains were optimized with the Protein Preparation Wizard tool from Schrödinger (21) at physiological pH. The covalent bond with the ligand was broken (with the corresponding hydrogen additions) to assure the free exploration. The second ligand in PDB 3IVM was removed. Two missing flexible fragments (residues 194–201 and 654–660) of the 3IUQ PDB entry were recovered and filled with the Prime software (21).

PELE

The PELE algorithm is based on a consecutive iteration of three main steps: a ligand and protein (backbone) perturbation, a side-chain sampling, and a minimization (22,23). Thus, the procedure begins by a ligand perturbation involving a random translation and rotation of the ligand. In the case of the protein, the perturbation is based on the α-carbons anisotropic network model (ANM) (24); all atoms are displaced by a minimization where the α-carbons are forced to follow a randomly picked low eigenvector (within the lowest six modes) obtained in the ANM approach. In particular, three consecutive perturbations of 1.5 Å in the same mode (and direction) were used before randomly picking a new mode. The ANM network model used identical springs connecting all α-carbons within a 15 Å cutoff (additional details on the ANM setup can be found in (20). The algorithm defines the most excited side chains with the largest changes in energy after the ANM move and these are included in the next step, the side-chain prediction. Here, PELE proceeds by optimizing all side chains local to the ligand in a defined distance (6 Å) together with the hot side chains determined in the ANM step (22,23). The last procedure involves the minimization of the entire system, keeping the α-carbon with a weak constraint after the ANM move. These steps compose a move that is accepted (a new local minimum) or rejected based on a Metropolis criterion, forming a stochastic trajectory. PELE runs were carried out at a temperature of 1000° K. As emphasized in our original work (16,20), this high Metropolis temperature does not correspond to a real thermal bath, the effective temperature being significantly lower. PELE uses an OPLS (optimized potentials for liquid simulations) all-atom force field (OPLS-AA) (25) with an implicit surface generalized Born (SGB) continuum solvent model (26).

PELE's combination of random perturbations and protein structure prediction algorithms results in an effective exploration of the protein energy landscape, capable of reproducing large conformational changes associated with ligand migration (16–20). The method provides MD quality results (20) at a significantly faster rate. When compared to docking techniques, it provides a good induced fit description, allowing the docking in difficult cases (apo, cross-docking, etc.) (19). Moreover, when combined with Markov state models, PELE provides absolute binding free energies in a similar fashion to extensive (and more expensive) MD techniques (17).

PELE entrance/exit protocols

For ZPP, rotations and translations alternate between two different values: small ones using 30° rotation and 0.75 Å translation, and big ones with 60° and 1.50 Å, which were independently and randomly switched (with 50% overall probability). Two different ANM options were used for sampling the protein backbone. First type includes a random switch among the first (lowest) six calculated modes. The second type, aiming to bias the protein opening, used a dominant ANM mode describing the movement of opening and closing of the two domains.

For the entrance simulations, the ligand was placed at eight different random positions in the protein surface (see Fig. 1 *D* for the exact initial ligand positions). When studying the exit pathways, the ligand was chosen always to start from the active site in an equivalent position to the crystallographic structures.

Entropy corrections

Entropy loss estimates for the bound complex (respect to the solution value) were divided in the following contributions: translational, rotational, conformational, and vibrational. Translational, rotational, and vibrational entropies were obtained using the standard ideal gas approximation (for example as described in the Gaussian thermochemistry site, http://www.gaussian.com/g_whitepap/thermo.htm). Conformational entropy was obtained by screening all available dihedral conformations for the ligand and the neighboring protein side chains (in direct contact, <2 Å): $\Delta S_{conf} = k_B \ln (\Omega_C / \Omega_P\Omega_L)$, where Ω_C is the available dihedrals for the complex, Ω_L for the isolated ligand and Ω_P for the isolated protein.

Loop prediction

Loop prediction calculations were executed with the Prime package from Schrödinger software (21). The protocol includes a default sampling algorithm and ultra extended loop refinement method, specifically designed to overcome sampling problems with long loops (>10 residues). Side-chain refinement was limited to residues with a side-chain heavy atom within 7.5 Å of any β-carbon from the loop. Energy cutoff for the final minimization refinement was varied to 20 kcal/mol (default is 10 kcal/mol). The loop prediction included residues 190–208 for the porcine POP and residues 190–205 for the bacterial one.

MD

MD simulations were performed using the Desmond MD program (27,28). The bound ZPP structure (PDB entry 1QFS) was solvated in an orthorhombic box of 19 012 water molecules, and 66 sodium and 49 chloride ions were added to neutralize and create a 0.14 M solution of NaCl. We used the OPLS-AA force field and the simple point charge water model. The default relaxation protocol in Desmond was used, followed by a 70 ns production run in the NPT ensemble using the Nose-Hoover thermostat and the Martyna-Tobias-Klein barostat (29,30). The smooth particle mesh Ewald method was used for the long-range interactions.

RESULTS

Loop prediction calculations and MD simulations

Loop prediction calculations obtained with the Prime software indicate preference for the closed state. For both porcine and bacterial, the first loop pose corresponds to a structure in close agreement with the crystallographic closed one with 0.5 Å and 0.9 Å α-carbon root mean-square deviation (RMSD), respectively. Higher energy poses, however, introduce some degree of opening. In porcine, we find the most open structure as the fifth pose, with an energy increase of 9.1 kcal/mol, shown in Fig. 2 *A* in black. Easier opening is observed in the bacterial POP simulations. In this case, the second result by score, shown in Fig. 2 *A* in red, represents the most opened loop geometry, with an energy increase of only 2.3 kcal/mol. In both porcine and bacterial POP the loop is involved in interactions with another small flexible part of β-propeller domain. This is the loop constructed by residues 215–222 for the porcine and the same one for the bacterial analog (residues 212–219). The two most open structures were chosen as our initial models for the open state simulations in PELE (called porcine open and bacterial open, see Table 1).

The analysis of the MD simulations for the porcine structure also indicates some degree of loop opening. The loop starts the opening at ~10 ns of the simulations and after passing through a semiopen conformation tends to partially close again. Together with loop motion toward opening, ZPP starts moving in a direction showing partial exiting through the loop with the phenyl ring as a leading residue.

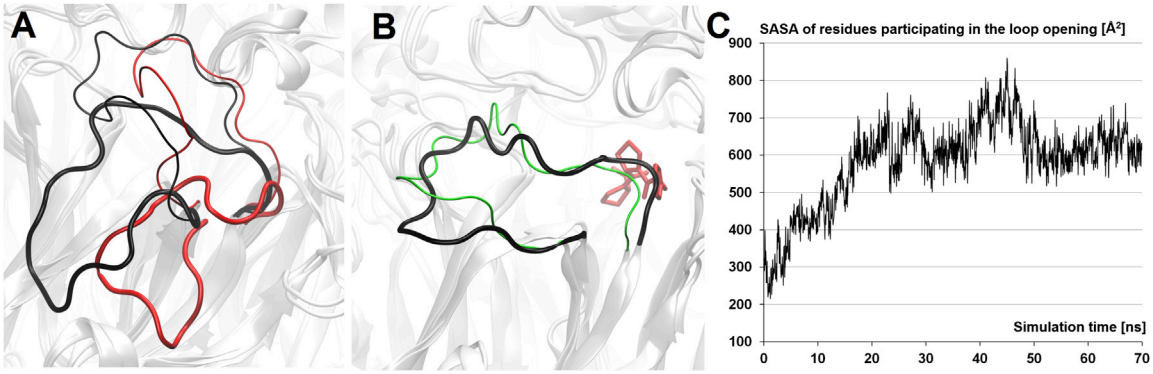

FIGURE 2 (*A*) Comparison of loop shapes from porcine (*thin black tube*) and bacterial (*thin red tube*) POP crystal structures, with PDB entries 1QFS and 3IVM, against most open structures from loop prediction calculations in the same colors but thicker tubes. (*B*) The 45 ns MD snapshot (*green*) is compared to the loop predicted structure (*black*). ZPP's position in the active site is underlined in *red licorice*. (*C*) Time evolution of SASA for residues 200–207, 590–594, and 641–644. Data were updated every 50 ps. To see this figure in color, go online.

The maximum open loop snapshot along the MD simulation, occurring at the 45 ns, is shown in Fig. 2 *B* in green. This opening is also clear when inspecting the evolution of the solvent accessible surface area for the main residues involved in the loop structure, as shown in Fig. 2 *C*. Interestingly, many of the structures from Prime's loop predictions have RMSD differences from MD snapshots lower than 1 Å (Fig. 2 *B*).

PELE explorations

The summary of PELE simulations exploring the entrance and exit pathways for the ZPP ligand in both porcine and bacterial POP is shown in Table 1. As a reminder, to model the open state, we used the most opened loop structures shown in Fig. 2 *A* (obtained with loop prediction techniques).

ZPP entrance pathway

As seen in Table 1, out of the 400 simulations for each system we obtain approximately the same number of entrances by the β-propeller pore in all of them (referred to as bottom pathway in Table 1). As indicated in the Materials and Methods, ZPP initial positions were randomly placed in the protein surface (Fig. 1 *D*). The remaining nonentering trajectories present structures where the ligand is associated with the surface (with some minor excursions into the bulk solvent). Furthermore, within each simulation the ligand explores a large fraction of the protein surface (see, for example, Fig. 3 or Movie S1 in the Supporting Material). In Fig. 3 we show a cross-section image where we display the ZPP ligand with blue beads (protein not shown) for the 50,000 snapshots along the porcine closed simulation. Clearly, we observe how the ligand covers the protein surface getting inside through the bottom entrance. In ~60% of these entrance events, ZPP enters the β-propeller pore by the hydrophobic phenyl moiety (Fig. 4 *A*). From the protein site, the most displaced blade of the β-propeller, after overlapping with the crystal structure, is the one bearing His-180 (Fig. 4 *B*). Two other blades, ones bearing Glu-134 and Lys-82, also show significant displacement. In the remainder 40% entrance events ZPP enters by the proline moiety, showing similar displacement of the blades. These three blade changes (*marked with stars* on Fig. 4 *B*), however, do not enlarge the pore significantly, and they seem to be induced by internal protein adjustments rather than by interaction with ZPP; analogous changes are seen with and without inhibitor.

TABLE 1 Entrance and exit pathways simulations

PELE experiment	ENTRANCE simulations	EXIT simulations
Protein structure	Pathway/**successful**/total number of trajectories	
Porcine closed	bottom[a]/**12**/400	bottom/**5**/400
Porcine open	bottom/**13**/400	bottom/**5**/400
Bacterial closed	bottom/**13**/400	bottom/**1**/400
Bacterial open	bottom/**14**/400[b]	loop/**10**/400[c]
	loop/**7**/400[b]	bottom/**1**/400[c]

[a]Bottom refers to the β-propeller pore.
[b,c]Same set of trajectories.

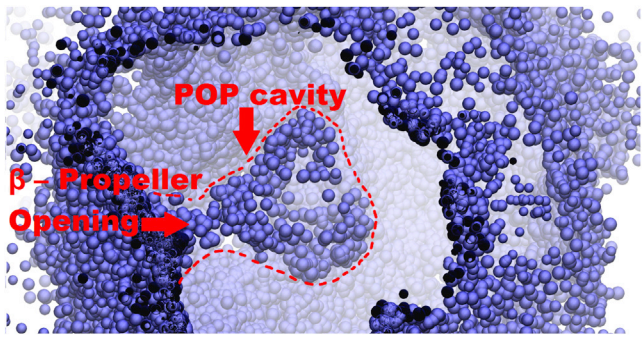

FIGURE 3 Cross section of POP from the side depicting the large cavity (*enclosed in red dashed curve*) and the ZPP surface exploration and bottom entrance. ZPP is presented with blue beads and the protein is omitted. To see this figure in color, go online.

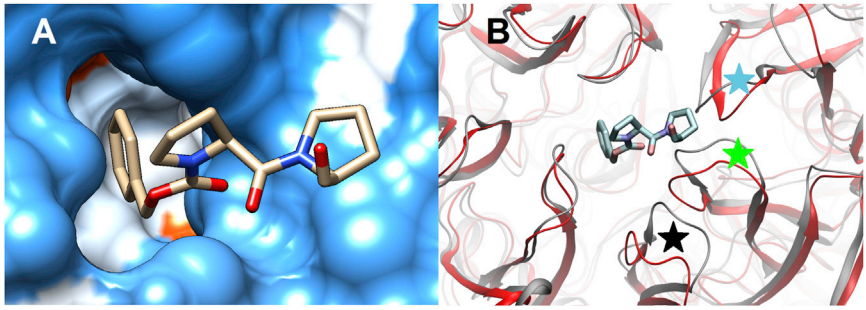

FIGURE 4 (*A*) Snapshot of entering ZPP through the β-propeller pore of porcine POP (surface presentation). (*B*) Bottom view of the entrance, gray cartoon, with an overlapped crystal structure (PDB entry 1QFS) shown in red cartoon. Black, green, and blue stars indicate bigger changes in the blades bearing His-180, Glu-134, and Lys-82. To see this figure in color, go online.

Our simulations indicate that the number of bottom entrances in both species is independent of the nature of the loop. Moreover, the mammalian POP does not show any ligand entrance by the loop even when starting by the open state. The bacterial one, however, shows seven entrances by the loop pathway when starting the simulation with the open state, the only instance where we observe entrance through the loop pathway.

ZPP exit pathway

Statistics on exit simulations, where the ligand starts in its active site position, show significantly different results from the entrance ones. In all cases the bottom exit was less probable than the entrance. Furthermore, contrary to the entrance, the exit through the bottom shows different results between porcine and bacterial, five events for porcine and only one for bacterial. Nevertheless, the exit through the bottom is still independent of the loop state.

Fig. 5 shows an entrance (*green*) and an exit (*red*) bottom trajectory for the closed porcine state. Fig. 5 A displays the ligand RMSD (*to the bound x-ray crystal*) along the PELE trajectory and the protein-ligand interaction energy. The entrance trajectory has initial high RMSD and interaction energies, decreasing accordingly along the entrance pathway. As expected, the opposite behavior is seen for the exit trajectory: an increase in RMSD and interaction energy. We should notice here that ZPP is a covalent inhibitor and that RMSD values are obtained in comparison to the bound crystal. The best binding ligand poses adopts an analogous crystal orientation but missing the last ~1–2 Å translation of the covalent bound formation, giving an overall RMSD ~5 Å. The ligand exits the bottom at approximately the 650 step, where we see an important barrier (~6 kcal/mol) in interaction energy. We want to note once more that the entrance trajectory starts significantly apart from the bottom pore.

Entropic contributions for ZPP at the bound state indicate a 31.2 kcal/mol correction to the binding free energy ($-T\Delta S$ term), obtained from 11.1, 15.0, 1.9, and 4.2 contributions from the translational, rotational, vibrational, and configurational entropy terms, respectively. This number, together with the PELE interaction energy, indicates an overall favorable binding event for ZPP in the noncovalent initial stage of the binding process. Moreover, we want to point to the nice correlation between the interaction energy and the RMSD in the last approach to the active site (*lower right green and left red corners in the bottom panel* of Fig. 5 *A*), indicating biological relevance for these pathways.

The most interesting aspect of the exit simulations, however, is the presence of 10 exit events through the loop pathway in the bacterial open state, see Fig. 6 *A*. Fig. 6 *B* shows a representative orientation of ZPP when crossing the loop, with the phenyl leading the pathway. In most of the cases exiting was observed with the participation of five residues, Trp-579, Phe-174, Tyr-233, Arg-232, and Tyr-190. Interestingly, this orientation adopted by the ligand is similar to the one observed in our MD simulation of porcine POP as a response of loop semiopening (Fig. 6 *B*).

Porcine POP interdomain opening

PELE can simulate protein motion according to the displacement of α-carbon-based ANM (24), an elastic model capable of describing large conformational changes. Inspection of the lowest six ANM modes showed that either the first or the second mode (depending on the initial structure) is associated with the interdomain opening direction. Thus, we forced PELE to sample this opening mode as the main ANM mode in porcine POP. Although open structures (similar to the open bacterial crystal) were produced when following (forcing) this mode, all of our attempts, including extreme temperatures, were unsuccessful in stabilizing them; the open structures spontaneously revert back to the closed one when not forcing the opening sampling mode.

Undecapeptide (substrate) entering and dipeptide (product) exiting simulations

Taking into account the results on ZPP, we modeled the entrance of a 11-residue peptide through the β-propeller pore in porcine POP. We also wanted to simulate the peptide cleavage in two products and the exit of the smaller (and more mobile) one in the presence of the other product in POP's cavity. For this purpose the Phe-Gly-Cys-Gly-Ala-Ser-Ala-Gly-Pro-Ala-Gly peptide, with two residues after the Pro cleavage point, was built. To facilitate the experiment the smaller product peptide (exiting part of

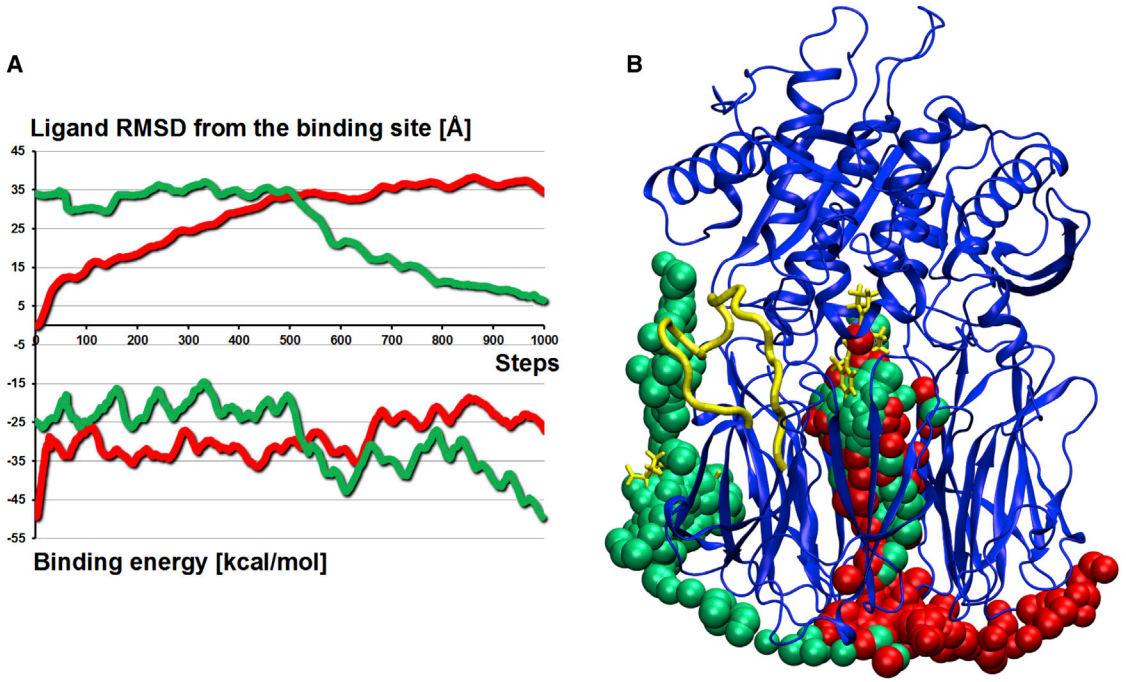

FIGURE 5 (*A*) Ligand interaction energy profile for a representative ZPP's entrance (*green*) and exit (*red*) trajectories. (*B*) Superposition of same trajectories (with same colors scheme) on one POP structure. ZPP positions along the trajectories are shown with beads, the flexible loop in *yellow tube*, and the initial inhibitor position in *yellow licorice*. To see this figure in color, go online.

the simulations) was chosen to be a dipeptide (Ala-Gly). The undecapeptide was placed around the bottom pore and guided to the α-carbon of the catalytic Ser-554 using the spawning algorithm in PELE. This algorithm aims to reduce the distance between two atoms (the α-carbons of Ser-554 and the substrate Pro) by random perturbation of the ligand and by using a tolerance distance window, 3 Å in our simulation. Every time the trajectory has a distance value larger (by the tolerance value) than the best registered distance, it will abandon the search and start with the best coordinates. Obviously, the best registered distance is updated when a shorter distance is found. Applying such protocol, the substrate cannot move further away, and explores freely, in a reduced window, possible structures that will reduce the desired distance. In this way, we can model difficult cases like the entrance of a large substrate by the bottom pore.

Fig. 7 shows the interaction energy profile and the Ser-554-Pro distance along the guided entrance process. As mentioned previously, we should keep in mind that this value reflects only internal energies which, due to the peptide size (forming numerous hydrogen bonds), are significantly larger. Rotational, vibrational, and translational entropic corrections amount for ~34 kcal/mol. Conformational entropy is out of our reach due to the presence of 30 rotatable bonds. Nevertheless, it was recently estimated to be on the order of ~60 kcal/mol for a nine-residue flexible peptide, giving rise to a total corrections on the order of ~94 kcal/mol (31). As seen in Fig. 7, no significant energy barrier is observed along the initial entrance process, in

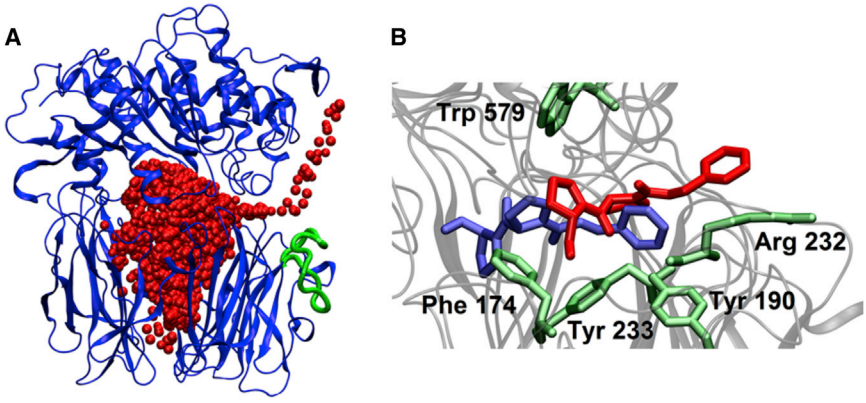

FIGURE 6 (*A*) A representative ZPP exit simulation for the open-loop (shown in *green*) bacterial POP. ZPP is shown in *red beads*. (*B*) A detailed view of the exit through the open loop and participating residues (in *green*). ZPP in blue corresponds to a superimposed structure from a MD snapshot in porcine POP. To see this figure in color, go online.

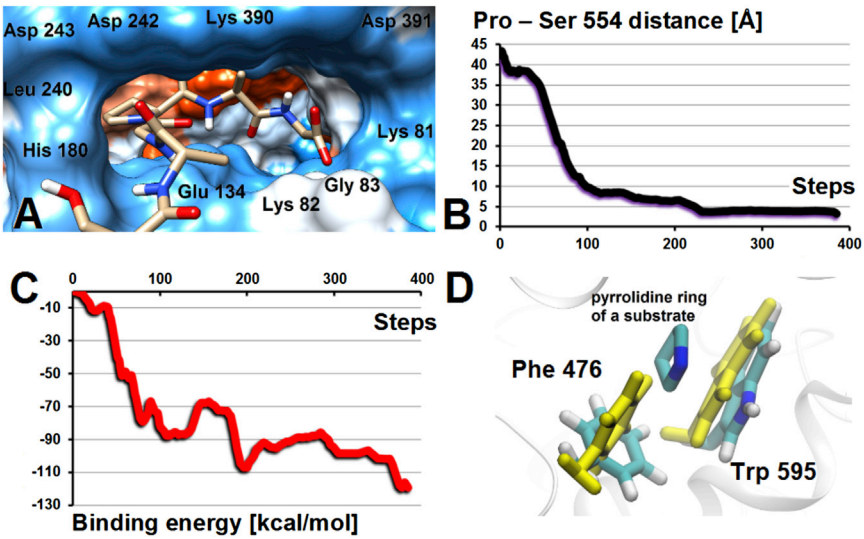

FIGURE 7 (*A*) Entering snapshot of the undecapeptide (C-terminus shown) and the residues forming the entrance. (*B*) Average distance (*angstroms*) between the carbonyl carbon of the undecapeptide proline and the oxygen atom of POP's Ser-554 for all entering trajectories. (*C*) Average binding energy profile for the entering trajectories (in kcal/mol). (*D*) Details of Phe-476 rearrangement to accommodate the pyrrolidine proline ring. To see this figure in color, go online.

agreement with a smooth reduction in the guiding distance. Thus, it seems like the big internal cavity can accommodate and easily allow the passage of large peptides. Around step 150 (Fig. 7) and after reaching a low Ser-554-Pro distance, we observe a significant side-chain rearrangement, giving rise to a better fitting (lower interaction energies) in the active site pocket. In particular, it involves mainly residues Phe-476 and Trp-595, two main actors in the active site proline pocket (as seen in the crystals 1QFS and 1E8N), where we observe changes from a closed state (Fig. 7 *D*, *yellow*) to an open one (Fig. 7 *D*, *atom type color*) to better accommodate the pyrrolidine proline ring.

When the proline α-carbon reached ~4 Å from the hydroxyl oxygen of Ser-554 (part of the catalytic triad) we cleaved the substrate into two peptide fragments. At this point, we repeated the nonbiased exit simulations for the small dipeptide product as performed with ZPP. We used the open loop state and we ensured that the remaining nine-residue product peptide was not blocking the bottom pore, facilitating the possible exit of the two-residue product fragment along both pathways. Of importance, and contrary to the results with ZPP, from a total of 400 trajectories, we observe now 38 exits along the loop pathway, with only seven events through the bottom. Movie S2 from the full process is deposited in the Supporting Material. We should emphasize that in this simulation only the two-residue fragment is perturbed (asked to leave) in PELE's simulation, and that it does it in the presence of a bulkier nine-residue fragment, which remains the entire time in the POP's cavity. An additional (and last) simulation was performed after removing (by deleting it) the nonapeptide from the POP's cavity. Thus, here the small product was let free to explore all internal volume before leaving the protein. In this case, we observed 19 exits through the loop opening and 29 through the bottom from a total number of 400 trajectories.

The dipeptide exit through the loop happens mostly from two different areas associated with a larger opening around residues Leu-206 and Thr-204 and a smaller one around Thr-202 (Fig. 8 *B*). Both combined, could result into an opening similar to the one predicted for bacterial POP (Fig. 2, *red*). For this small dipeptide product, five exit trajectories cross another loop (some authors call it loop B (13)), involving catalytic domain residues 578–604; three of them involve close interactions to Tyr-589. In only one trajectory, ZPP exits around residue Pro-74 between the hinge keeping together the catalytic and β-propeller domains (exits 578–604 and through the hinge are not shown on Fig. 8). Exiting around loop 578–604 shows another potential flexible part of POP. Finally, analogous to ZPP, entropic contributions for the dipeptide product gave 25.5 kcal/mol, which added to an ~10 kcal/mol interaction energy in the solvent gave a total exothermic energy profile for product release.

DISCUSSION

The loop prediction results, both using Prime and MD, indicate that the loop in the mammalian POP has significant mobility, confirming previous experimental and computational results (10,13). The excellent agreement of the predicted structures with the experimental ones, 0.5 Å and 0.9 Å α-carbon RMSD with porcine and bacterial crystals, respectively, indicates the quality of Prime's algorithm in sampling long loops and gives credit to the open structure predictions. Moreover, we observe a large degree of overlapping between the loop prediction techniques and the MD results for the semiopen loops. Thus, one would expect that being able to run a longer MD we would observe a larger opening of the loop, similar to the one predicted by Prime. Moreover, a similar argument could be expected for domain opening. Our mammalian simulations do not

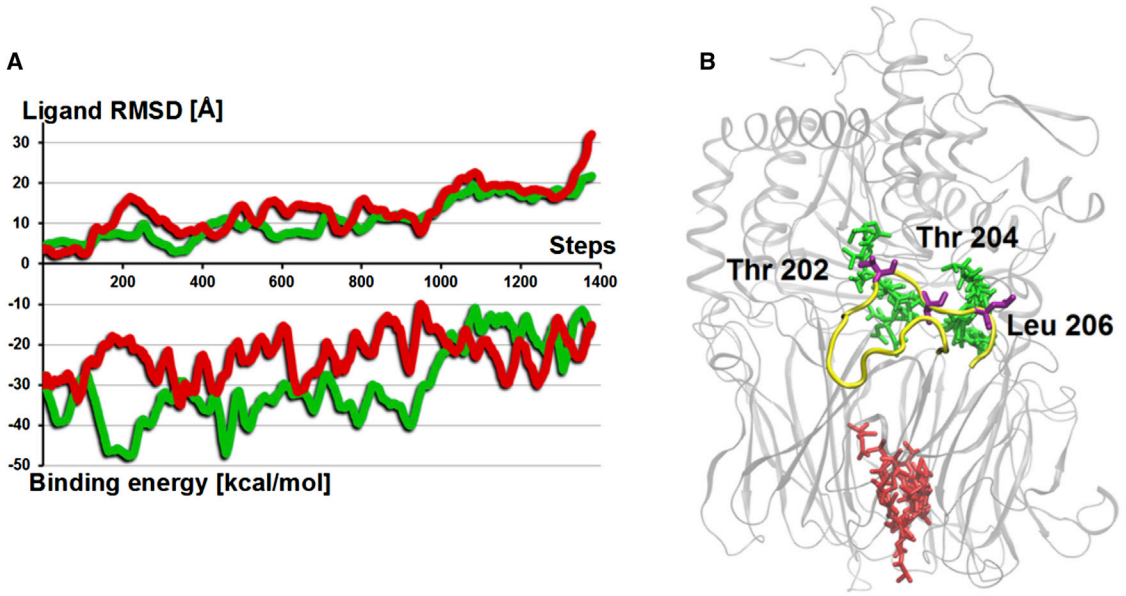

FIGURE 8 (*A*) Ligand interaction energy and RMSD profiles for the bottom (*red*, exiting at ~30 Å of ligand RMSD) and loop (*green*, exiting at ~15 Å of ligand RMSD) exit pathways for the Ala-Gly product. (*B*) Representative exit snapshots along the exit pathways (same color scheme). In *purple licorice*, we underline some residues from the loop (in *yellow*) where the dipeptide exits POP's cavity. To see this figure in color, go online.

present significant interdomain conformational change as the one present in the bacterial crystals. Inspection of the interdomain contacts seems to confirm larger difficulty in opening mammalian POP. Normal mode analysis, however, still indicates that the lowest modes describe domain-domain movement. Thus, one could expect that considerably larger MD simulations could introduce partially opened structures.

In previous experimental work, it showed not only the mobility of this loop but also the important contributions to the substrate enter/exit mechanisms, i.e., lower activity of a porcine POP mutant lacking this loop (10). The largest predicted opening in porcine POP is, however, not as significant as the one observed in bacterial POP. Nevertheless, the exit of the two-residue product during the simulation of the undecapeptide substrate seems to indicate that the loop opening is enough for some small product release (see below).

Our simulations indicate a clear preference for the bottom entrance. Only for the bacterial opened state do we observe partial entrance by the loop pathway, yet the statistics for this state show higher occurrence for the bottom entering (Table 1). We should keep in mind that simulations were performed with a relatively small molecule size (compared to average POP size substrates) like ZPP. Thus, for larger peptides one would expect even a larger contribution of bottom entrances.

The entrance by the bottom pore is in agreement with recent umbrella sampling simulations (8) where the authors monitor the energy profile when forcing exit pathways. In another study, using MD from a docked inhibitor in the bottom of the large internal cavity, the authors show spontaneous migration of the ligand toward the active site region (9). Along the different entrances through the bottom pore, we find nine residues (Glu-134, His-180, Leu-240, Ser-241, Asp-242, Asp-243, Gln-388, Lys-389, and Lys-390) having contacts closer than 4 Å with ZPP atoms. One could expect that mutations that introduce bulkier side chains in some of these positions would result in a weaker inhibitory activity (with possibly a large alteration of binding kinetics) of ZPP. Within our statistical limitations, our results indicate no orientation preference along the β-propeller pore entrance for small inhibitors. Furthermore, such size molecules do easily rotate and translate in the POP's internal huge cavity. Bigger peptide substrates, however, might need some guiding to pass preferably with its C-terminus. The amount of Lys side chains around the bottom pore (Lys-81, 82, 84, 157, 162, 183, 389, 390) could be this guiding tool (some of them shown in Fig. 7 *A*).

Although the bottom entrance in bacterial POP has similar probability to the mammalian one, exiting by the same pathway is severely more restricted (Table 1). A close look at both structures reveals clear differences in the pore residues. Mammalian Lys-81 and Lys-389, are replaced by the bulkier Arg-83 and His-377. Similarly, Asp-242 in porcine is replaced by a longer Glu-240. More importantly, Arg-135 replaces Glu-134. We observe, in ~50% of cases where ZPP is in close proximity to exit by the pore, how Arg-135 blocks ZPP passage by interacting with Glu-240 and Asp-237 (Fig. 9). In the reverse (entrance) cases, however, the inhibitor molecule has more mobility and interacts closely with the pore, leading to bigger changes (including Arg-135) and to more successful entry trials—total number

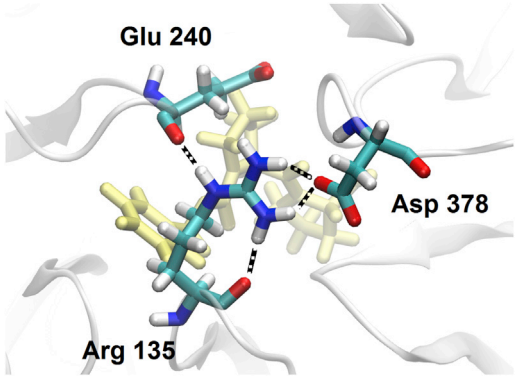

FIGURE 9 Bottom view of the β-propeller (*white ribbons*) of bacterial POP showing Arg-135 blocking the center of the pore by interacting with Glu-240 and Asp-378. ZPP is shown with yellow licorice. To see this figure in color, go online.

of 27 pore entrances to only two exits for the bottom pathway of bacterial POP (Table 1).

Our results correlate the larger degree of loop opening in bacterial POP with the appearance of exit events through it. Along the exit pathway, the hydrophobic interactions of ZPP's proline with Phe-174 and Tyr-233 closely resemble the interactions of the inhibitor in the active site (for example with residues Phe-476 and Trp-595 in porcine POP). In addition, the orientation of ZPP when exiting, Fig. 6 *B*, agrees with the one adopted when entering: the proline moiety finding first this hydrophobic cavity. Orientations of the side chains of Phe-174 and Tyr-233 change dynamically through both processes, adopting conformations that consecutively interact with the phenyl or proline rings in ZPP. Additional residues showing large motion in some exiting trajectories are Tyr-190 (also located in porcine POP) and Arg-232.

In an attempt to model the entire process, we diffused an 11-residue peptide from the bottom pore to the porcine active site. The substrate reached the active site with a smooth energy profile, agreeing with the preference observed in ZPP for the bottom entrance. Interestingly, after the cleavage and forming the two-residue product we observe a preference for the dipeptide exiting through the loop pathway. This difference could come not only because of the partial shielding of the bottom pore from the remainder nonapeptide, but also because of the size of the leaving ligand. We studied this hypothesis with an additional 400 simulations where we removed the nonapeptide. The 19/29 ratio for loop/bottom exits shows higher occurrence through the bottom pore, although the loop one still participates in 40% of the successful cases. This means that the option for exiting products through the bottom pore will be further impeded with the presence of a large molecule (i.e., our modeled nonapeptide product) in the POP cavity or even totally obstructed. Of significant importance in this case is the conformation of the N-terminus of the substrate and which part of the POP β-propeller domain has been occupied. Our 11-residue peptide simulations show that the tail of the peptide prefers extended conformations covering the area close to the bottom. Extensive preliminary active site search and docking calculation by us, showed good interactions in proximity of the β-propeller pore. These observations seem to agree with the only crystal structure with a bigger substrate—a bulky octapeptide in porcine POP, PDB entry 1E8N (32). In this crystal, the N-terminus is pointing toward the bottom, whereas the C-terminus is not well resolved. Thus, all together this indicates that longer peptide will obstruct the bottom passage and drive the products release through the loop opening.

Our binding energy and entropy estimates indicate large compensation effects. Although this topic has been under debate, numerous recent calorimetric studies seem to support its importance (33,34). We have studied three peptide-like substrates with large flexibility and entropy loss upon binding, in agreement with recent observations (31). Interestingly, the polar large cavity in POP seems to have evolved to compensate for this reduction in mobility by increasing the number of protein-ligand interactions. This is clear when inspecting the binding energy plots where we observe a sudden large increase (in absolute value) once the ligand enters the cavity (see, for example, Fig.7 *C* at steps ~50). Due to the large errors in entropy calculations, however, any attempt to obtain accurate binding free energies should use more sophisticated methods; our binding energies are only of qualitative nature.

In addition to the large polar cavity, our PELE and MD simulations show a hydrophobic pocket (similar to the one for proline in the active site of POP) buried and uncovered by the loop motions, which could be a trigger mechanism for peptide release. Thus, product exit seems to follow three steps: initial ligand binding to the predocking site followed by a larger opening of the loop, which pulls out the products in the same direction and then exit through it. This mechanism would agree with the experimental results showing that mutants lacking a flexible loop convert POP to an inefficient enzyme (10).

In summary, our extensive computational analysis reveals a clear preference for ligand entrance through the β-propeller pore. Exit conditions, however, seem to be more specific of the species, degree of loop opening, and nature of the substrate. Overall, cleavage of a small peptide at the active site seems to be correlated with its exit along the loop. This loop is shown to be very flexible in our simulations; modeling domain-domain opening in mammalian POP (if present) will require considerably longer simulations.

SUPPORTING MATERIAL

Two movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)04666-9.

ACKNOWLEDGMENTS

This study was supported by The European Research Council (2009-Adg25027-PELE) to V.G., MICIN-FEDER (BIO2013-40716R) to E.G., the Generalitat de Catalunya (XRB and 2014SGR-521) to E.G., and SEV-2011-00067 of Severo Ochoa Program, awarded by the Spanish Government to D.L.

REFERENCES

1. Fülöp, V., Z. Böcskei, and L. Polgár. 1998. Prolyl oligopeptidase: an unusual β-propeller domain regulates proteolysis. *Cell.* 94:161–170.
2. Brandt, I., S. Scharpé, and A.-M. Lambeir. 2007. Suggested functions for prolyl oligopeptidase: a puzzling paradox. *Clin. Chim. Acta.* 377:50–61.
3. Lawandi, J., S. Gerber-Lemaire, …, N. Moitessier. 2010. Inhibitors of prolyl oligopeptidases for the therapy of human diseases: defining diseases and inhibitors. *J. Med. Chem.* 53:3423–3438.
4. López, A., T. Tarragó, and E. Giralt. 2011. Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010. *Expert Opin. Ther. Pat.* 21:1023–1044.
5. Brandt, I., M. Gérard, …, A.-M. Lambeir. 2008. Prolyl oligopeptidase stimulates the aggregation of α-synuclein. *Peptides.* 29:1472–1478.
6. Myöhänen, T. T., M. J. Hannula, …, A.-M. Lambeir. 2012. A prolyl oligopeptidase inhibitor, KYP-2047, reduces α-synuclein protein levels and aggregates in cellular and animal models of Parkinson's disease. *Br. J. Pharmacol.* 166:1097–1113.
7. Kichik, N., T. Tarragó, …, E. Giralt. 2011. ^{15}N relaxation NMR studies of prolyl oligopeptidase, an 80 kDa enzyme, reveal a pre-existing equilibrium between different conformational states. *ChemBioChem.* 12:2737–2739.
8. St-Pierre, J.-F., M. Karttunen, …, A. Bunker. 2011. Use of umbrella sampling to calculate the entrance/exit pathway for Z-Pro-prolinal inhibitor in prolyl oligopeptidase. *J. Chem. Theory Comput.* 7:1583–1594.
9. Kaushik, S., and R. Sowdhamini. 2011. Structural analysis of prolyl oligopeptidases using molecular docking and dynamics: insights into conformational changes and ligand binding. *PLoS One.* 6:e26251.
10. Szeltner, Z., T. Juhász, …, L. Polgár. 2013. The loops facing the active site of prolyl oligopeptidase are crucial components in substrate gating and specificity. *Biochim. Biophys. Acta.* 1834:98–111.
11. Szeltner, Z., D. Rea, …, L. Polgár. 2004. Concerted structural changes in the peptidase and the propeller domains of prolyl oligopeptidase are required for substrate binding. *J. Mol. Biol.* 340:627–637.
12. Li, M., C. Chen, …, T. K. Chiu. 2010. Induced-fit mechanism for prolyl endopeptidase. *J. Biol. Chem.* 285:21487–21495.
13. Kaszuba, K., T. Róg, …, A. Bunker. 2012. Molecular dynamics, crystallography and mutagenesis studies on the substrate gating mechanism of prolyl oligopeptidase. *Biochimie.* 94:1398–1411.
14. Buch, I., T. Giorgino, and G. De Fabritiis. 2011. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA.* 108:10184–10189.
15. Shan, Y., E. T. Kim, …, D. E. Shaw. 2011. How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* 133:9181–9183.
16. Borrelli, K. W., A. Vitalis, …, V. Guallar. 2005. PELE: protein energy landscape exploration. A novel Monte Carlo based technique. *J. Chem. Theory Comput.* 1:1304–1311.
17. Takahashi, R., V. A. Gil, and V. Guallar. 2014. Monte Carlo free ligand diffusion with Markov state model analysis and absolute binding free energy calculations. *J. Chem. Theory Comput.* 10:282–288.
18. Madadkar-Sobhani, A., and V. Guallar. 2013. PELE web server: atomistic study of biomolecular systems at your fingertips. *Nucleic Acids Res.* 41 (Web Server issue):W322–W328.
19. Borrelli, K. W., B. Cossins, and V. Guallar. 2010. Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility. *J. Comput. Chem.* 31:1224–1235.
20. Cossins, B. P., A. Hosseini, and V. Guallar. 2012. Exploration of protein conformational change with PELE and meta-dynamics. *J. Chem. Theory Comput.* 8:959–965.
21. Schrödinger Suite. 2013. Protein Preparation Wizard; Epik version 2.6, Schrödinger, LLC, New York, NY; Impact version 6.1, Schrödinger, LLC; Prime version 3.4, Schrödinger, LLC.
22. Jacobson, M. P., R. A. Friesner, …, B. Honig. 2002. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* 320:597–608.
23. Jacobson, M. P., G. A. Kaminski, …, C. S. Rapp. 2002. Force field validation using protein side chain prediction. *J. Phys. Chem. B.* 106:11673–11680.
24. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 2:173–181.
25. Kaminski, G. A., R. A. Friesner, …, W. L. Jorgensen. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptidesy. *J. Phys. Chem. B.* 105:6474–6487.
26. Ghosh, A., C. S. Rapp, and R. A. Friesner. 1998. Generalized Born model based on a surface integral formulation. *J. Phys. Chem. B.* 102:10983–10990.
27. Molecular, Dynamics System, 2012, version 3.1, D. E. Shaw Research, New York, NY. Maestro-Desmond Interoperability Tools, version 3.1, Schrödinger LLC, New York, NY.
28. Bowers, K. J., E. Chow, ..., D. E. Shaw. 2006. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *Proc. ACM/IEEE Conf. Supercomputing (SC06).* Tampa, Florida, November 11–17, 2006.
29. Hoover, W. G. 1986. Constant-pressure equations of motion. *Phys. Rev. A.* 34:2499–2500.
30. Martyna, G. J., D. J. Tobias, and M. L. Klein. 1994. Constant pressure molecular-dynamics algorithms. *J. Chem. Phys.* 101:4177–4189.
31. Killian, B. J., J. Y. Kravitz, …, M. K. Gilson. 2009. Configurational entropy in protein-peptide binding: computational study of Tsg101 ubiquitin E2 variant domain with an HIV-derived PTAP nonapeptide. *J. Mol. Biol.* 389:315–335.
32. Fülöp, V., Z. Szeltner, …, L. Polgár. 2001. Structures of prolyl oligopeptidase substrate/inhibitor complexes. Use of inhibitor binding for titration of the catalytic histidine residue. *J. Biol. Chem.* 276:1262–1266.
33. Chodera, J. D., and D. L. Mobley. 2013. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* 42:121–142.
34. Portman, K. L., J. Long, …, D. J. Scott. 2014. Enthalpy/entropy compensation effects from cavity desolvation underpin broad ligand binding selectivity for rat odorant binding protein 3. *Biochemistry.* 53:2371–2379.