

Use of Umbrella Sampling to Calculate the Entrance/Exit Pathway for Z-Pro-Prolinal Inhibitor in Prolyl Oligopeptidase

Jean-François St-Pierre,[†] Mikko Karttunen,[‡] Normand Mousseau,^{*,†} Tomasz Róg,[§] and Alex Bunker^{||,⊥}

[†]Département de Physique et Regroupement Québécois sur les Matériaux de Pointe, Université de Montréal, C.P. 6128, succursale centre-ville, Montréal (Québec), Canada H3C 3J7

[‡]Department of Applied Mathematics, The University of Western Ontario, 1151 Richmond Street North, London (Ontario), Canada N6A 5B7

[§]Department of Physics, Tampere University of Technology, P.O. Box 692, FI-33101 Tampere, Finland

^{||}Centre for Drug Research, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, FI-00014, University of Helsinki, Finland

[⊥]Department of Chemistry, Aalto University, PO Box 6100, FI-02015, Aalto, Finland

ABSTRACT: Prolyl oligopeptidase (POP), a member of the prolyl endopeptidase family, is known to play a role in several neurological disorders. Its primary function is to cleave a wide range of small oligopeptides, including neuroactive peptides. We have used force biased molecular dynamics simulation to study the binding mechanism of POP. We examined three possible binding pathways using Steered Molecular Dynamics (SMD) and Umbrella Sampling (US) on a crystal structure of porcine POP with bound Z-pro-prolinal (ZPP). Using SMD, an exit pathway between the first and seventh blade of the β -propeller domain of POP was found to be a nonviable route. US on binding pathways through the β -propeller tunnel and the TYR190-GLN208 flexible loop at the interface between both POP domains allowed us to isolate the flexible loop pathway as the most probable. Further analysis of that pathway suggests a long-range covariation of the interdomain H-bond network, which indicates the possibility of large-scale domain reorientation observed in bacterial homologues and hypothesized to also occur in human POP.

1. INTRODUCTION

Endopeptidases are a class of proteases that hydrolyze internal, i.e., nonterminal, peptide bonds. Prolyl oligopeptidase (POP) is a proline-specific endopeptidase that cleaves oligopeptides (<30-mer) at the C-side of an internal proline. *In vitro* analysis found that a wide variety of neuroactive peptides substrates can be cleaved by POP,^{1–5} and *in vivo* analysis indicates that these are its actual metabolic substrates.² It has been found, although somewhat inconsistently, that certain POP inhibitors can reverse memory loss caused by amnesic agents, neurological disorders, and aging, making POP an important drug target.^{6,7} In addition, an alteration in POP enzyme activity has been measured in serum samples taken from patients suffering from Parkinson's and Alzheimer's disease⁸ and multiple sclerosis.⁹ Experimental evidence exists that POP might have a role beyond its peptidase function. Examples of this include protein–protein interactions,^{2,10} intracellular transport,¹¹ inflammation,^{12,13} angiogenesis,¹⁴ and cancer development.^{14,15} The biochemical role of POP and its inhibitors has been reviewed in refs 1–5, 7, and 16.

Several fundamental questions regarding access to the active site, gating, selection, and the detailed inhibition mechanisms remain unanswered. Recent studies have combined a number of experimental and simulation techniques to address the details of inhibition mechanisms¹⁷ as well as binding and gating mechanisms,¹⁸ and progress has been made in developing combinatorial libraries for POP.¹⁹

In this paper, our focus is on the active site access (exit) pathways. This is a fundamental question in understanding the peptidase function. What makes this question particularly

challenging is its dynamic nature: entry and exit involve dynamic response. Hence, docking studies and short time-scale density functional calculations are not able to address this question. We study porcine POP, crystallized with a bound ZPP inhibitor described by Fülop et al.²⁰ (PDB database ID 1QFS). The structure is composed of two domains: the protease catalytic domain with an α – β hydrolase fold composed of amino acids 1–71 and 436–710 and the seven-bladed β -propeller, comprised of amino acids 72–435. The active site is on the surface of an internal cavity between the two protein domains. The ZPP inhibitor has a hydrophobic head that sterically blocks the active site and an aldehyde tail that forms a reversible covalent hemiacetal bond with the SER554 residue of the catalytic domain. The β -propeller domain has an unusual, mostly hydrophobic interaction between the first and seventh blades. This is called the “velcro rip” and has been proposed to act as a filter to the active site.^{20–23} Later studies have demonstrated that the domain by itself is more stable than in conjunction with the catalytic domain.²⁴ This finding is in agreement with a small scale computational study carried out by Fuxreiter et al.²⁵ and our previous computational work on POP with an unbound inhibitor in the binding cavity:²⁶ both indicate that the β -propeller is a highly stable structure. It was suggested that the entry point is most likely through the H-bonded network of loosely structured loops that connect the two protein domains, in particular the location of the TYR190-GLN208 flexible loop.

Received: December 8, 2010

Published: April 29, 2011

Recent electron microscopy experiments suggest that the interdomain region offers a wider entry point than the β -propeller tunnel.¹⁸ For example, Shan and collaborators find that the crystal structure of a distant POP relative, *Sphingomonas capsulata* (SC; PDB database ID 1YRZ), displays an open conformation in absence of the inhibitor, where the two domains are separated, exposing the catalytic triad to the solvent.²⁷ Experimental evidence of an open form of a homologous protein has also been captured by X-ray diffraction crystallography of *Aeromonas punctata* prolyl endopeptidase.²⁸ In that study, Li et al. captured the closed form of the protein after soaking the open form of the crystallized protein in a bath of the inhibitor. By maintaining the protein in this open form using glutaraldehyde driven lysine cross-linking, they obtained a complete absence of activity. This interdomain large-scale motion has not yet been observed in experiments performed on mammalian POP. Cysteine cross-linking experiments binding the two domains together have, however, been shown to lead to strongly reduced protein activity.²⁹

Elucidating the mechanism through which substrates gain access to the active site, i.e., identifying the ligand entry pathway, would be beneficial for the development of new classes of inhibitors for mammalian POP. It is important to perform full MD simulations with explicit solvent since water can access the active site and in some cases even have a decisive role.³⁰ To study the entry pathways, it is necessary to test the various possible trajectories directly. In this paper, we use SMD and US to measure the free energies of the three postulated exit/entry pathways of the ZPP inhibitor. Once we have determined the correct entry/exit pathway, we then study the interaction between inhibitor and the elements of the protein structure that comes in contact with it. Although SMD is first used to generate rough pathways, the bulk of our results are obtained using US. Using these methods, we determine that the most probable exit pathway is through the loosely structured loops between the two domains opposite the interdomain hinge.

2. METHODS

2.1. Software, Model, and Simulation Parameters. All simulations were performed using the GROMACS 4.0 simulation package³¹ at constant pressure (1 bar) and temperature (310 K; NPT). Temperature was maintained using the Nosé–Hoover thermostat^{32,33} and pressure using the Parrinello–Rahman barostat.³⁴ The coupling time constants were set to 0.1 and 1.0 ps for thermostat and barostat, respectively, and the protein and solvent were thermalized separately. Electrostatic interactions were computed using the Particle-Mesh-Ewald method (PME).^{35,36} The Lennard-Jones interactions were cut off at 1.0 nm. The same cutoff was used for the real-space part of PME. Charge groups were chosen to be small to avoid artifacts that may arise if the charge groups are spatially too large.³⁷

To parametrize the POP and ZPP molecules and the solution ions, we used the OPLS-AA (Optimized Parameters for Liquid Simulations, AA stands for all-atom) potential set.³⁸ Partial charges on ZPP were taken from our previous work.²⁶ For water, the TIP3P model was used.³⁹

The initial structure was taken from our previous 100 ns MD study of POP with the ZPP inhibitor unbound in the active site.²⁶ POP was solvated in a box of water of size $10 \times 10 \times 10$ nm. Potassium and chlorine ions were added to neutralize the system

to model physiological conditions (140 mM salt concentration). The solvated simulation box contained a total of 100 468 atoms.

Analysis and visualization were performed using the VMD (Visual Molecular Dynamics) package⁴⁰ and GROMACS^{31,41} analysis tools. Pathways were generated and evaluated using SMD and US, as described in the following sections.

2.2. SMD and US. In SMD, an external force, called force bias, F_{fb} , is applied to a single atom or a group of chosen atoms, through their center of mass. SMD is an irreversible approach, and its use is based on the Jarzynski equality $\langle \exp[-W/k_B T] \rangle = \exp[-\Delta G/k_B T]$, where k_B is the Boltzmann constant, T is temperature, and W is the total nonreversible work done on the system by F_{fb} during a nonequilibrium transition between two states connected by a reaction coordinate λ .^{42–44} The free energy difference between these two states is given by $\Delta G \equiv \Delta G(\lambda)$. The angular brackets stand for an ensemble average taken by repeating the simulation many times along the path connecting the initial and final states, λ_1 and λ_2 . The essence of the Jarzynski equality is that it links rigorously the work done in a nonequilibrium process to the change in the equilibrium free energy difference. This method has been used in small systems, e.g., single molecule conformational changes,^{45–47} and, more rarely, in larger systems involving, e.g., protein–protein interactions.⁴⁸

We define the direction of F_{fb} as \hat{r}_{fb} . All other degrees of freedom are allowed to react freely to this force. To drive the system, we apply a harmonic force

$$\vec{F}_{fb}(\lambda) = k_{fb}(\lambda \vec{r}_{fb} - \vec{r}_{cm}) \quad (1)$$

where k_{fb} is the force constant and \vec{r}_{cm} is the center of mass. A force bias with a fixed value of k_{fb} is introduced with λ increasing from zero to one at a continuous rate as the simulation proceeds. This rate is known as the *pulling rate*.

It has been shown by Park and Schulten⁴⁹ that the Jarzynski equality is equivalent to calculating the free energy difference from the first and the second-order cumulants of the work done by the biasing force:

$$\Delta G = \langle W \rangle - \frac{\langle W^2 \rangle - \langle W \rangle^2}{2k_B T} \quad (2)$$

While both the Jarzynski equality and eq 2 are formally correct in the thermodynamic limit, finite sampling leads to potential problems since ΔG depends exponentially on W ; the result is easily dominated by the extremal values of the distribution.^{50,51} The impact of this sensitivity to rare pathways can be evaluated by comparing results obtained from these two methods.

US^{52,53} obtains the free-energy difference between two states from a set of equilibrated simulations. Like in SMD, a force bias is applied. However, now the configuration is equilibrated at each step, which we refer to as window. The value of the harmonic force constant used for each window is independent and can be set to a value that optimizes the efficiency with which the phase space is sampled. Parameters k_i and λ must be selected in such a way as to ensure that the phase space sampled by adjacent windows overlaps sufficiently, forming a continuous pathway between the initial and the final state. Results from all windows can then be combined using the weighted histogram analysis method (WHAM)⁵⁴ to provide the full thermodynamical evolution along the reaction coordinate.

Examples exist of systems where US has been successfully applied to include ion channels,^{55,56} unfolding of the I27 titin domain,⁵⁷ and the evaluation lipid transfer and peptide penetration

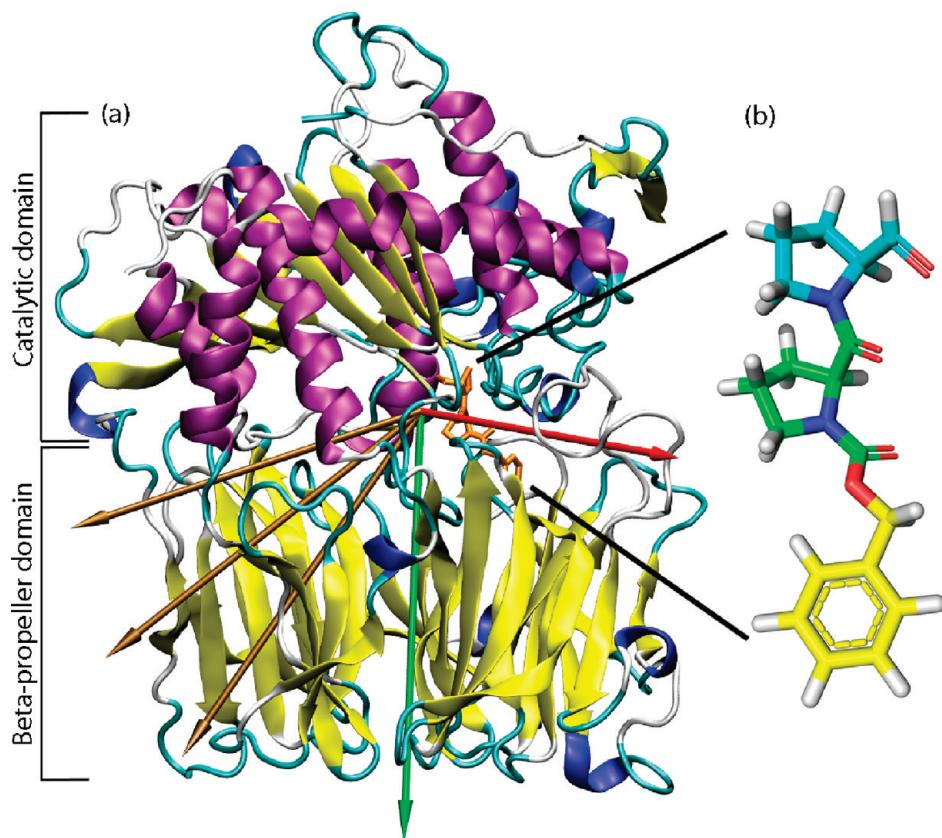


Figure 1. (a) SMD ZPP-pulling vectors for the flexible loop exit (red), the β -propeller tunnel exit (green), and three possible exits through the velcro-rip of the β -propeller (golden arrows). The ZPP inhibitor inside is in orange. (b) Zoom on the ZPP with carbons colored by our definition of its three regions: PHE (yellow), PRO1 (green), and PRO2 (cyan) groups. In its inhibition mode, the PRO2 aldehyde group is involved in a covalent bond to the SER554 of the protein. Oxygen and nitrogen atoms are left in red and dark blue, respectively.

in cellular membranes.^{58,59} SMD and US have been compared for ligand binding in the gramicidin A channel and Kv11.1 (also known as hERG) potassium channel.⁶⁰ Several comprehensive reviews are available.^{61,62}

2.3. Application of SMD and US to the Study of Our System. SMD employing the Jarzynski equality was performed on the system with the applied bias force dislocating the ZPP molecule from the binding pocket along the three proposed pathways shown in Figure 1a. We generated one pull vector each for the pathways through the β -propeller tunnel and the flexible loop. For the third pathway, the velcro-rip junction between the first and seventh blade of the β -propeller, three different pull vectors (see Figure 1a) were attempted to test this suggested exit pathway.^{20–23}

Since the center of mass and the orientation of the protein are not inherently conserved properties for the simulated system (protein + ZPP + solvent), it was necessary to add restraints to conserve the position and orientation of the protein. This was achieved by restraining the positions of a small number of α -carbon atoms positioned far from the sampled ZPP exit pathway with harmonic restoring forces with a force constant of $10\text{ kJ}/(\text{mol} \times \text{nm}^2)$. Two sets of weak restraints were selected, for the flexible loop exit and β -propeller tunnel pathways, and in both cases the harmonic force constant on these restraints was 2–3 orders of magnitude smaller than the restraining force bias applied to the ZPP ligand. For the β -propeller tunnel pathway, the restrained residues PRO7-ASP35, ASP431-GLY464,

TYR510-LYS546, and ILE610-GLN629 were all located on the catalytic domain and did not form part of the inner cavity. For the flexible loop exit, the atoms selected for restraint span both domains: on the catalytic domain, the α -carbon atoms of the residues VAL427-LYS458 were restrained, and on the beta-propeller, they were placed on amino acids TYR73-GLY108 and GLY288-LYS458. Although, formally, this position on both domains affects allosteric communication between the two domains, the restraints are weak enough and dispersed enough to minimize this possibility. With the orientation of the protein maintained, the force bias is relative to the center of mass of the protein, and the expression for the force bias vector becomes

$$\vec{F}_{fb} = k_{fb}(\lambda \vec{r}_{fb} - (\vec{r}_{cm-ZPP} - \vec{r}_{cm-protein})) \quad (3)$$

Following the work of Tskhovrebova et al.,⁶³ we first computed the free-energy barriers for the different exit pathways using SMD. This was performed with a harmonic force constant $k_{fb} = 5\text{ MJ}/(\text{mol} \times \text{nm}^2)$ and pulling rates of 0.5 nm/ns , as was used in previous work,⁶⁴ and also of 0.1 nm/ns . To make it possible for the ZPP to exit POP, the simulation box was extended by 2.0 nm in the direction of the force bias for the β -propeller tunnel pulling vector, which brings the total number of atoms in the periodic box to 130 004.

For our US, we used states obtained from the SMD as starting configurations. For each window, we selected a state where the value of the component of the displacement vector (the vector

Table 1. Umbrella Sampling Windows Parameters for the Flexible Loop Exit^a

window	<i>z</i> (nm)	<i>T</i> (ns)	<i>k</i> _{fb} (MJ/(mol × nm ²))	source	window	<i>z</i> (nm)	<i>T</i> (ns)	<i>k</i> _{fb} (MJ/(mol × nm ²))	source
1	0.3	10	2.5	SMD	18b	1.95	7.2	10	18
2	0.4	10.1	2.5	SMD	19	2.1	10.1	2.5	SMD
3	0.5	10	2.5	SMD	20	2.2	10.1	2.5	SMD
4	0.6	10.1	2.5	SMD	20b	21.4	10.1	10	20
5	0.7	10	2.5	SMD	21	2.3	10.1	2.5	SMD
6	0.8	10	2.5	SMD	22	2.4	5	2.5	SMD
7	0.9	10	2.5	SMD	23	2.5	10	2.5	SMD
8	1	10	2.5	SMD	24	2.6	10	2.5	SMD
8b	1	8.4	10	8	25	2.7	10.1	2.5	SMD
9	1.1	10	2.5	SMD	26	2.8	10.1	2.5	SMD
9b	1.05	8.4	10	9	27	2.9	7	2.5	SMD
9c	1.1	7.1	10	9	28	3	10	2.5	SMD
10	1.2	10.1	2.5	SMD	29	3.1	10	2.5	SMD
10b	1.15	8.4	10	10	30	3.2	10	2.5	SMD
10c	1.2	8.4	10	10	31	3.3	10	2.5	SMD
11	1.3	10	2.5	SMD	32	3.4	10	2.5	SMD
12	1.4	10	2.5	SMD	33	3.5	10	2.5	SMD
13	1.5	10	2.5	SMD	34	3.6	10	2.5	SMD
13b	1.47	10	10	13	35	3.7	10	2.5	SMD
14	1.6	10.1	2.5	SMD	36	3.8	10	2.5	SMD
15	1.7	10.1	2.5	SMD	37	3.9	10	2.5	SMD
15b	1.65	8.4	10	15	38	4	10	2.5	SMD
15c	1.7	8.5	10	15	39	4.1	10	2.5	SMD
16	1.8	10	2.5	SMD	40	4.2	10	2.5	SMD
16b	1.75	7.1	10	16	41	4.3	11	2.5	SMD
16c	1.8	8.5	10	16	42	4.4	10	2.5	SMD
17	1.9	10	2.5	SMD	43	4.5	10	2.5	SMD
17b	1.85	7.2	10	17	44	4.6	10	2.5	SMD
17c	1.9	7.1	10	17	45	4.7	10	2.5	SMD
18	2	10	2.5	SMD	46	4.8	10	2.5	SMD

^a *z* is the reaction coordinate, the equilibrium distance between the ZPP and protein's center of mass for each window. *k*_{fb} is the force constant of the spring restraining the ZPP at distance *z*. *T* is the length of time of the window's MD simulation. The "source" column indicates what was the source of the initial conformation of the window, where SMD means it was extracted from the close position in the steered molecular dynamics and where a number points to the US window for which the last conformation was extracted.

connecting the centers of mass of the POP and ZPP along the direction of the force bias) matched the value of λ for the window.

Only the flexible loop and the β -propeller exit pathways were studied with US. The pathways were divided into 46 and 48 windows, respectively, with the reaction coordinate values separated by 0.1 nm. In each window, $k_{fb} = 2.5 \text{ MJ}/(\text{mol} \times \text{nm}^2)$, and force biased MD was performed for 10 ns and 8–9 ns, respectively. These initial simulations were started using conformations obtained from the SMD, as shown in Tables 1 and 2. To ensure sufficient overlap between distributions sampled in adjacent windows, undersampled regions were identified, and simulations were launched in these windows.

For the loop opening, 14 extension windows were added in regions where the resistance to sampling was particularly high (energy barrier regions); the sampling was extended by 5–8 ns, performed with a higher force constant of 5–10 MJ/(mol × nm²) to obtain a minimum sampling of 50 000 conformations per bin of size 0.01 nm along the reaction coordinate. For the β -propeller tunnel exit, 20 extension simulations with a force constant of 5 MJ/(mol × nm²) of length 5 ns each were

added to reach a minimum sampling of 100 000 conformations per bin of size 0.0095 nm. From this data, the free energy difference profile of ZPP-POP unbinding was computed using WHAM,⁵⁴ as implemented in GROMACS 4.5.3.⁶⁵ These extension windows were initiated with the last conformation of a nearby window, as indicated by the "source" columns of Tables 1 and 2.

3. RESULTS

In order to describe the ZPP molecule and its interactions with its environment, we use the same formalism as in our previous publication²⁶ to define the structure of the ZPP molecule in terms of three atomic groups. As shown in Figure 1b, PHE represents the aromatic phenyl head, PRO1 the middle proline, and PRO2 the terminal proline containing the aldehyde group (involved in the hemiacetal bond with SER554). In the following section, we describe our SMD results followed by the US results, and an analysis of the exit pathway.

3.1. Exploration of the Exit Pathways Using SMD. As the first attempt to estimate the free energy barriers associated with the exit pathways, we conducted a set of three SMD simulations

Table 2. Umbrella Sampling Windows Parameters for the β -Propeller Exit^a

window	<i>z</i> (nm)	<i>T</i> (ns)	<i>k</i> _{fb} (MJ/(mol × nm ²))	source	window	<i>z</i> (nm)	<i>T</i> (ns)	<i>k</i> _{fb} (MJ/(mol × nm ²))	source
1	-0.15	9.1	2.5	SMD	27b	2.19	9.5	5	27
2	-0.05	9.7	2.5	SMD	28	2.35	5	2.5	SMD
3	0.05	8.6	2.5	SMD	29	2.45	7.3	2.5	SMD
4	0.15	9.4	2.5	SMD	29b	2.36	8.3	5	29
5	0.25	9.3	2.5	SMD	30	2.55	8.3	2.5	SMD
5b	0.2	5	5	S	30b	2.48	5	5	30
6	0.35	9.3	2.5	SMD	31	2.65	7.9	2.5	SMD
7	0.45	9.3	2.5	SMD	31b	2.59	5	5	31
8	0.55	9	2.5	SMD	32	2.75	6.3	2.5	SMD
9	0.65	9.4	2.5	SMD	33	2.85	7.8	2.5	SMD
10	0.59	5	5	SMD	34	2.95	8.7	2.5	SMD
11	0.75	8.5	2.5	SMD	34b	2.88	5	5	34
12	0.85	8.1	2.5	SMD	35	3.05	8.6	2.5	SMD
12b	0.8	5	5	12	35b	3	5	5	35
13	0.95	8.2	2.5	SMD	36	3.15	8.6	2.5	SMD
14	1.05	7.8	2.5	SMD	37	3.25	8.3	2.5	SMD
15b	1	5	5	15	38	3.35	8.5	2.5	SMD
16	1.15	9.2	2.5	SMD	38b	3.31	5	5	38
17	1.25	7.9	2.5	SMD	39	3.45	8.2	2.5	SMD
18	1.35	8.1	2.5	SMD	39b	3.5	5	5	39
18b	1.3	5	5	18	40	3.55	8.6	2.5	SMD
19	1.45	8.2	2.5	SMD	40b	3.6	5	5	40
20	1.55	7.7	2.5	SMD	41	3.65	8.4	2.5	SMD
21	1.65	5	2.5	SMD	41b	3.7	5	5	41
21b	1.6	8.2	5	21	42	3.75	8.5	2.5	SMD
22	1.75	5	2.5	SMD	42b	3.8	5	5	42
22b	1.72	9.2	5	22	43	3.85	8.7	2.5	SMD
23	1.85	8.9	2.5	SMD	44	3.95	8.8	2.5	SMD
24	1.95	5	2.5	SMD	45	4.05	8.5	2.5	SMD
24b	1.95	9.3	5	24	45b	4.05	5	5	45
25	2.05	5	2.5	SMD	46	4.15	8.5	2.5	SMD
25b	2.05	8.4	5	25	47	4.25	8.8	2.5	SMD
26	2.15	9.3	2.5	SMD	48	4.35	8.6	2.5	SMD
27	2.25	5	2.5	SMD					

^a *z* is the reaction coordinate, the equilibrium distance between the ZPP and protein's center of mass for each window. *k*_{fb} is the force constant of the spring restraining the ZPP at distance *z*. *T* is the length of time of the window's MD simulation. The "source" column indicates what was the source of the initial conformation of the window, where SMD means it was extracted from the close position in the steered molecular dynamics and where a number points to the US window for which the last conformation was extracted.

with a pulling rate of 0.1 nm/ns and 35 SMD simulations at a rate of 0.5 nm/ns for both the flexible loop and the β -propeller exit pathways with a total simulation time near 250 ns per pulling direction. The work (*W*) required to move the ZPP over the entire path was calculated for each run. We computed the free energy difference from the distribution of *W* values using the two separate methods discussed previously and found a significant discrepancy between their results. The pull rate was found to strongly influence the sampled pathways: in all 35 SMD trajectories with the higher pulling rate along the loop exit pathway, the ligand exited either around or through the TYR190–GLN208 flexible loop, while two of the three slower SMD pathways favored a concerted opening of the same loop. Moreover, the orientation of the ZPP at the exit of these two trajectories was reversed. The trajectory requiring the lowest work had ZPP oriented such that the phenyl group PHE was directed toward the catalytic domain and the PRO2 proline group toward the β -propeller.

The second trajectory showing a concerted opening of the flexible loop had a value of *W* that was greater by 40 kJ/mol than the previous pathway. In this case, ZPP exited the protein in a reversed orientation with the PHE and PRO2 groups directed toward the β -propeller and catalytic domain, respectively. The lack of convergence in the sampled work obtained at the slow pulling rate of 0.1 nm/ns indicates that a slower pulling rate would be required for the ZPP to adopt the preferred conformation in most pulling trials. Given the large number of trials necessary for sufficient statistics, and the high computational cost of each pulling trial, this option was not retained.

We also used SMD to investigate a third possible exit pathway proposed by a number of groups, which involves ZPP moving through the velcro-rip between the first and seventh blades of the β -propeller domain (see Figure 1).^{20–23} We generated three SMD simulations using three different pulling vectors (gold vectors in Figure 1) and a pulling rate of 0.5 nm/ns. In all three

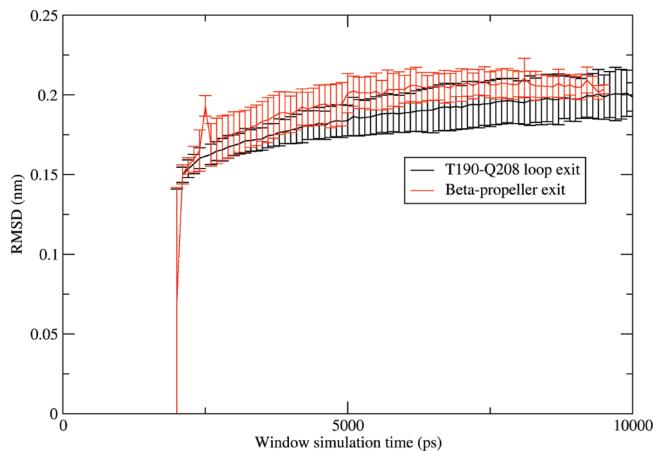


Figure 2. Average root-mean-square deviation evolution in the loop exit (black) and the β -propeller (red) as a function of displacement from the binding site. Error bars express the standard deviation.

cases, the ZPP molecule failed to pass through the velcro rip, either exiting by the β -propeller exit or through the hinge region linking the catalytic and the β -propeller domains. Clearly, the velcro-rip is extremely stable. This result is consistent with previous simulation results that indicated that the β -propeller domain is highly stable.^{26,25} Thus, this exit pathway was ruled out and the rest of our analysis focused only on the two remaining candidate pathways.

While SMD can serve to rule out a proposed exit pathway, this method requires too many simulations to converge. We did, however, use the configurations from generated pathways as starting points for US trajectories.

3.2. Free Energy Difference Calculations with US. The initial configurations for launching MD in all of the US windows were obtained from the SMD pathways. US for both the flexible loop and the β -propeller exit pathways was performed as described previously. The sum of all simulation times (i.e., including all windows) is equivalent to 539 ns of MD for the loop exit and 499 ns for the β -propeller exit. In all individual simulation windows, statistics were accumulated after an initial 2 ns equilibration with positions, simulation times, and force constants as listed in Tables 1 and 2. To evaluate convergence toward equilibrium, we computed the root-mean-square deviation (rmsd) of the protein for each window. After 4 ns, the rmsd values converged, on average, to $0.20 \text{ nm} \pm 0.02 \text{ nm}$ for the loop exit and $0.20 \text{ nm} \pm 0.01 \text{ nm}$ for the β -propeller pathway Figure 2.

Figure 3 shows the number of configurations sampled along the reaction coordinate for the loop and the β -propeller exit pathways, respectively, with bins of sizes 0.01 and 0.0095 nm. The smallest count for the loop exit is 50 585 conformations at $z = 1.45 \text{ nm}$ and 122 193 conformations at $z = 1.27 \text{ nm}$ for the β -propeller. They correspond to visiting times of 101 and 245 ps, respectively.

The resulting potential of mean force obtained from WHAM for the loop exit and β -propeller exit is presented in Figure 4. The error bars give the standard deviation as calculated using 100 iterations of statistical bootstrapping using the histogram Bayesian bootstrap available in Gromacs version 4.5.3.⁶⁵ For the loop exit, we find the lowest free energy at position 0.8 nm in the cavity, the transition peak at 1.8 nm, and the solvated free energy at 4.15 nm in the reaction coordinate. This reveals a free energy difference between the bound and free ZPP conformations of

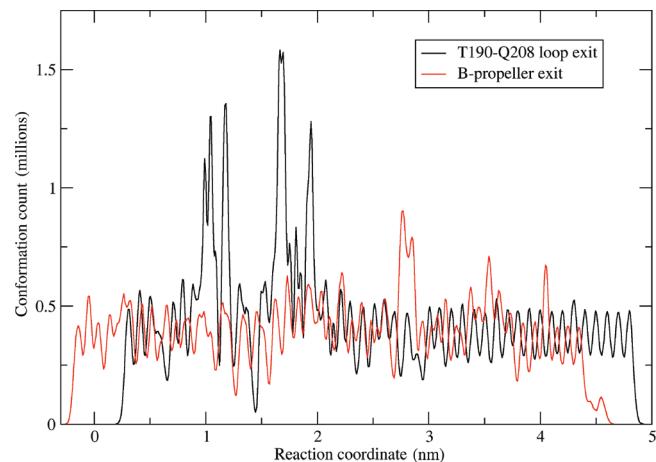


Figure 3. Histogram of the reaction coordinate along the loop exit as a function of displacement from the binding site using a bin size of 0.01 nm (black) and along the β -propeller tunnel exit using a bin size of 0.0095 nm (red).

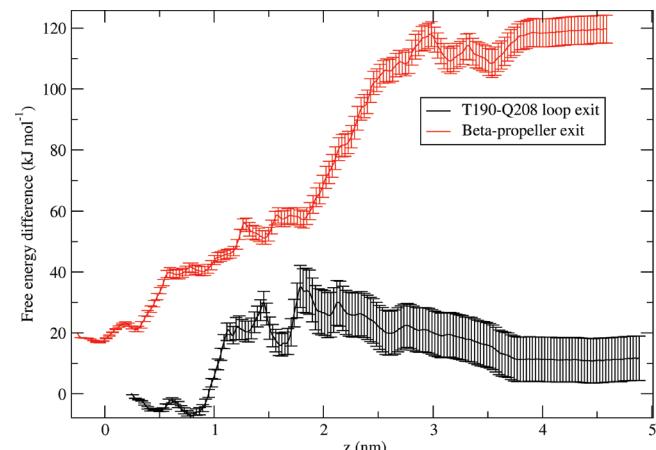


Figure 4. Potential of mean force for the loop exit (black) and β -propeller tunnel exit (red) as a function of displacement from the binding site. The red curve was shifted vertically for better legibility. Error bars express the standard deviation.

$-18.5 \pm 8.2 \text{ kJ/mol}$ with a transition energy barrier of $25.1 \pm 8.1 \text{ kJ/mol}$ in the entry direction. For this interaction, we can calculate a constant of inhibition $K_i = 0.8 \text{ mM}$ using the formula:

$$K_i = [1M] e^{-\Delta G/RT} \quad (4)$$

where R is the perfect gas constant and T is the absolute temperature. The accuracy of ΔG was calculated to be $\pm 8.2 \text{ kJ/mol}$, on the basis of one standard deviation on the binding free energy difference; thus our the confidence interval for K_i is $[32 \mu\text{M}, 18 \text{ mM}]$. This inhibition constant is much weaker than the empirical value of $K_i = 0.35 \text{ nM}$ ⁶⁶ since it does not include the formation of the favorable hemiacetal bond, an event not simulated in our study. However, the K_i found in our study is on the same order as other inhibitors who do not form a hemiacetal bond like suc-Gly-Pro-Nan with a $K_i = 0.278 \pm 0.35 \text{ mM}$ at pH 5.6 and Z-Gly-Pro-OH with a $K_i = 0.253 \pm 0.18 \text{ mM}$ at pH 8 or $K_i = 21.2 \pm 0.5 \mu\text{M}$ at pH 7.35.⁶⁷

For the β -propeller tunnel exit, a WHAM analysis performed on the US data (Figure 4 in red) yielded a free energy minimum at -0.05 nm and a transition free energy of -101.6 ± 4.7 kJ/mol, corresponding to a plateau starting at $z = 4.0$ nm, at which point the ZPP is free in the solvent. Contrary to the loop exit, this pathway shows no global minimum, rising systematically as the

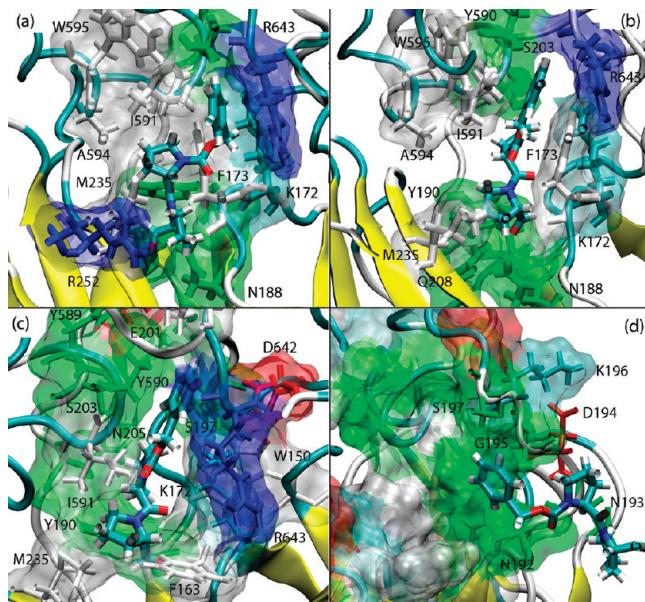


Figure 5. Main amino acids (colored by type) making contact with ZPP from the windows. $z = 1.0$ nm (a), $z = 1.3$ nm (b), $z = 1.6$ nm (c), and $z = 3.0$ nm (d).

ZPP moves out of the protein. Furthermore, the free energy barrier corresponds to a K_i of 7.6×10^{-18} M with a confidence interval [1.2 aM, 50 aM] based on one standard deviation of the free energy difference, 9 orders of magnitude smaller than the experimental value, hence very strong binding.

These two observations suggest that the β -propeller exit is not sampled sufficiently, in spite of a total of 499 ns of MD simulation dedicated to this pathway. We hypothesize that the cause of this undersampling is the pathway's extreme constriction, which leads to very low mobility of the ZPP in the windows, with a ZPP position ranging from $z = 1.95$ nm to $z = 2.95$ nm, presented in Figure 6, making it almost impossible to fully sample the accessible conformations along the trajectory. Thus, even though the displacement may vary and overlap with neighboring windows, the phase space along the exit path is not properly sampled due to the strong dependence on initial conditions for the MD run in each window. To verify this hypothesis, we compared the average radius of gyration of ZPP for both the loop and the β -propeller pathways, the number of conformation clusters adopted by ZPP based on a rmsd clustering and the standard deviation of the angular distribution of ZPP as a function of the position along the reaction coordinate (Figure 7).

The above quantities are associated with the mobility and the conformational entropy for ZPP as each new window is explored. For all three properties, the window-to-window fluctuations are 20 to 30% larger for the β -propeller exit than the flexible-loop exit. Since there is sufficient overlap between neighboring windows, we would expect that these three properties evolve smoothly from window to window, along the reaction coordinate. The large fluctuations observed for the β -propeller exit indicate rather that there is an imperfect overlap in the

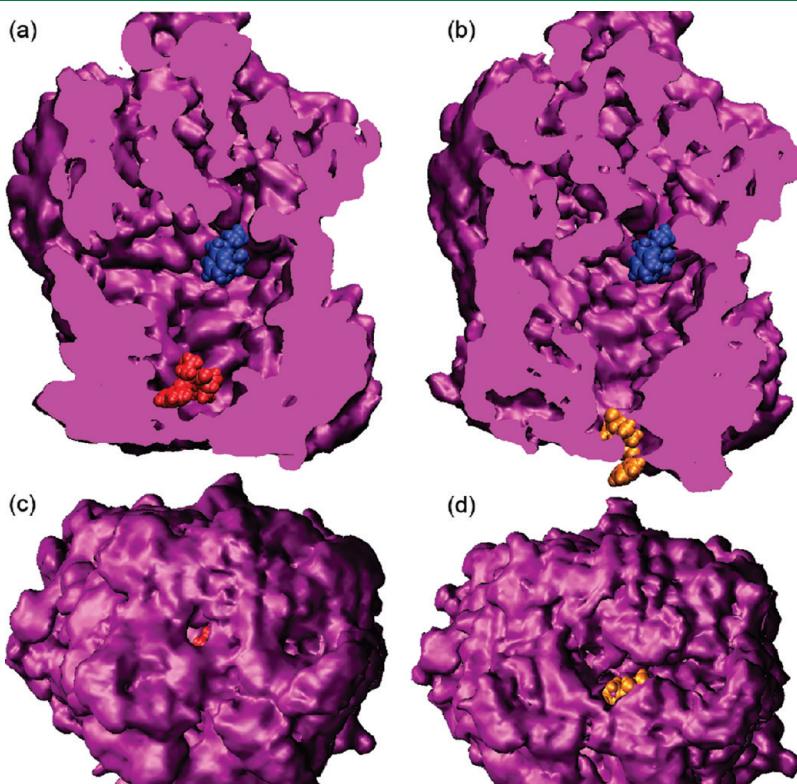


Figure 6. Constriction of the β -propeller exit pathway. Position of the ZPP is presented (a) in red for window $z = 1.95$ nm and (b) in orange for $z = 2.95$ nm, with the ZPP from window $z = -0.15$ nm in blue as a reference to the starting position. The bottom view of the β -propeller tunnel is presented in c and d, respectively. The protein's surface was computed from a volumetric density map averaged over the trajectory of the respective window.

configuration samples within adjacent windows leading to sharp conformational transitions between windows. The overlap of the position of the center of mass of ZPP between neighbor windows does not translate to a POP–ZPP complex conformation overlap of the same windows. This suggests that the visited states are strongly influenced by the initial configuration selected for each window. While we cannot, under these conditions, extract an accurate free-energy barrier for the β -propeller exit pathway, the sampling difficulties observed can be directly associated with the presence of a significant free-energy barrier. This suggests that the pathway is significantly less favorable than the loop exit. It is thus not unreasonable to conclude that this will not be the access pathway to the active site, leaving the loop exit as the only feasible pathway.

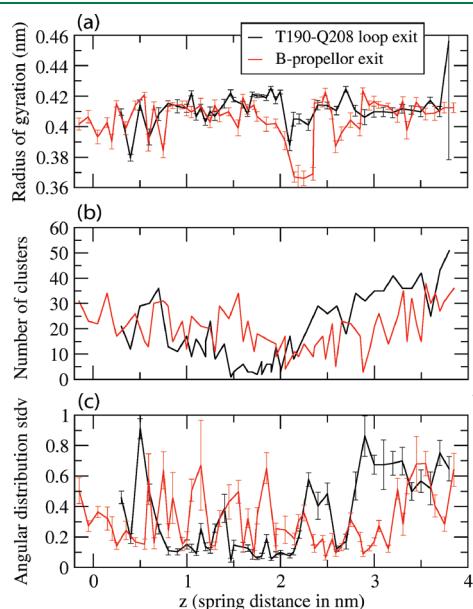


Figure 7. (a) Average radius of gyration of ZPP as a function of the displacement from the binding site. (b) Average number of conformation clusters using a rmsd clustering algorithm of cluster size 0.07 nm. (c) Standard deviation of the angular distribution of ZPP as a function of displacement. Error bars in a and c are obtained through a 5000 bootstrap evaluation of 10% of the available data and a confidence probability of 95%.

3.3. Interaction between ZPP Inhibitor and POP in Loop-Exit Pathway. Now that the loop exit, associated with three loosely structured loops between the two protein domains, has been identified as the most probable entrance/exit pathway, we investigate this pathway in further detail.

We investigated the relation between the free energy as a function of displacement z along the pathway with various structural properties to understand the nature of the energy barrier. As seen in Figure 4, two free energy peaks exist, at $z = 1.46$ and 1.8 nm, of approximately the same height, separated by a relatively deep local free-energy minimum at $z = 1.65$ nm. While it is difficult to determine the exact nature of these free-energy features, they are well correlated with specific alterations in the contacts and the H-bond network. Variation in the amino acids making contact with the different parts of ZPP can be linked to these transitions and are listed for the PRO2 (Table 3), PRO1 (Table 4), and PHE (Table 5) groups. As the ZPP moves out of the protein, and z approaches 1.3 nm, it comes into contact with the cavity wall, in particular, a group of hydrophobic side chains: MET235, ILE591, PHE173, and TRP595. ZPP makes contact predominantly with positively charged ARG252 and ARG643, positioned on each side of the ligand (see Figure 5a). As it moves through this constrained region, however, the ligand also forms contact with hydrophobic amino acids PHE173, ILE591, and MET235 Figure 5b, as well as with hydrophilic amino acids of the TYR190–GLN208 flexible loop (Figure 5c), strongly reducing ZPP's access to the solvent.

The ligand's outward displacement along the reaction coordinate is sterically constricted by a group of amino acids that block ZPP's direct access to the flexible loop. Looking more specifically at the configurations sampled around the free energy peak at $z = 1.46$ nm, we see that ZPP is sterically constrained between THR590 on one side of the PHE group and TRP150, LYS172, and ARG643 on the other side. This reduces significantly the PRO1 and PRO2 groups' conformation flexibility, thus lowering the entropy available to the ZPP and raising the free energy difference.

The drop in free energy, as ZPP moves past $z = 1.46$, is associated with a decrease in the average number of H bonds of 1.6 ± 0.4 (relative to the average number of H bonds for the region $z = 0.3$ to $z = 1.25$ nm) between the whole protein and its TYR190–GLN208 loop, easing ZPP's diffusion pathway

Table 3. Average Probability of Existence of the Most Persistent Contacts in Three Regions of the Reaction Coordinate z for the PRO2-Body Contacts and PRO2-Loop Contacts

PRO2-body						PRO2-loop					
region	0.3–1.3 nm	region	1.4–2.0 nm	region	2.1–3.7 nm	region	0.3–1.3 nm	region	1.4–2.0 nm	region	2.1–3.7 nm
MET235	0.62	PHE173	0.85	PHE173	0.42	TYR190	0.41	TYR190	0.84	GLN192	0.42
PHE173	0.44	MET235	0.79	MET235	0.34	GLN208	0.32	ALA189	0.45	SER197	0.34
ILE591	0.35	ILE591	0.61	LYSH172	0.27	ASN205	0.15	ASN205	0.43	GLY195	0.30
ASN188	0.32	LYSH172	0.45	ILE591	0.26	SER203	0.06	GLN192	0.33	LYSH196	0.23
ARG252	0.31	ASN188	0.28	TRP150	0.19	LEU206	0.03	GLN208	0.31	ASP194	0.23
GLY236	0.26	SER174	0.20	ARG170	0.16	ALA189	0.01	SER203	0.18	GLN193	0.23
CYS175	0.20	TRP234	0.19	ASN188	0.12			GLY195	0.04	ASP198	0.17
SER174	0.19	TRP150	0.15	SER174	0.09						
TRP234	0.19	ALA594	0.14	TRP234	0.06						
ALA594	0.18	TRP595	0.05	VAL171	0.04						
GLY237	0.17	LYSH233	0.02	ALA594	0.04						
LYSH172	0.15			GLU169	0.03						

Table 4. Average Window Probability of Existence of the Most Persistent Contacts in Three Regions of the Reaction Coordinate z for the PRO1-Body Contacts and PRO1-Loop Contacts

PRO1-body						PRO1-loop					
region	0.3–1.3 nm	region	1.4–2.0 nm	region	2.1–3.7 nm	region	0.3–1.3 nm	region	1.4–2.0 nm	region	2.1–3.7 nm
ILE591	0.67	ASN205	0.77	TRP150	0.25	TYR190	0.092	SER203	0.75	GLN192	0.44
MET235	0.56	TRP150	0.70	ASN205	0.14	LEU206	0.048	TYR190	0.57	LYSH196	0.40
TRP595	0.48	LYSH172	0.66	PHE173	0.10	SER203	0.031	GLN192	0.51	GLY195	0.36
PHE173	0.46	THR590	0.51	LYSH172	0.07			SER197	0.50	SER197	0.34
ALA594	0.36	ILE591	0.45	ILE591	0.05			GLU201	0.32	ASP198	0.18
ILE478	0.26	PHE173	0.29	VAL171	0.04			THR204	0.13	ARG170	0.18
ARG252	0.21	MET235	0.11	MET235	0.04			ASP194	0.13	ASP194	0.18
PHE476	0.16	TYR589	0.03	THR590	0.01			LYSH196	0.10	TYR190	0.16

Table 5. Average Window Probability of Existence of the Most Persistent Contacts in Three Regions of the Reaction Coordinate z for the PHE-Body Contacts and PHE-Loop Contacts

PHE-body						PHE-loop			
region	0.3–1.3 nm	region	1.4–2.0 nm	region	2.1–3.7 nm	region	1.4–2.0 nm	region	2.1–3.7 nm
ILE591	0.84	THR590	0.92	TRP150	0.17	SER197	0.68	LYSH196	0.38
ARG643	0.87	LYSH172	0.81	ARG170	0.12	GLY199	0.61	SER197	0.38
PHE173	0.57	VAL645	0.56	PHE173	0.08	SER203	0.30	GLN192	0.32
TRP595	0.46	PHE173	0.53	THR590	0.06	ASP198	0.28	ASP198	0.31
THR590	0.36	ILE591	0.47	LYSH172	0.05	GLU201	0.23	GLY195	0.29
ASN555	0.24	ASP642	0.40	TYR589	0.03	THR202	0.12	SER203	0.24
HISB680	0.24	ARG643	0.36	LYSH23	0.03	THR200	0.08	GLU201	0.22
ASP149	0.23	TYR589	0.33			LYSH196	0.06	ASP194	0.18
LYSH172	0.22	VAL644	0.29			GLN192	0.05	ASN205	0.17
PHE476	0.21	VAL580	0.21						
ILE478	0.11	TRP150	0.19						
SER554	0.11	TRP595	0.06						

(Figure 8). Table 6 gives a list of H bonds between the above loop and the rest of the protein that are modulated by the passage of ZPP through position $z \geq 1.46$ nm. For $z < 1.46$, the TYR190–GLN208 loop maintains an average of 11.75 ± 0.12 H bonds with the rest of the protein. There are only 17 pairs of atoms forming H bonds between the above loop and the protein that exist for more than 25% of the simulation time and another 28 pairs of atoms present between 5% and 25% of that simulation time.

The second free energy barrier ($z = 1.8$ nm) can be linked with the breaking of H bonds that connect the β -propeller and catalytic domains. The average number of H bonds between these two domains increases from 15.5 ± 0.22 at $z = 1.6$ nm to 19.4 ± 0.15 at $z = 1.8$ nm and falling back to 15.1 ± 0.26 at $z = 1.95$ nm (Figure 8), which correlates well with the observed energy peak at $z = 1.8$ nm. The bulk of the variation in the H-bond network between the two domains is not directly linked to the opening of the flexible loop into the solvent since we can see that the number of H bonds between the TYR190–GLN208 loop and the catalytic domain varies only slightly, going from 1.5 ± 0.1 H bonds at $z = 1.6$ nm to 1.8 ± 0.1 at $z = 1.8$ nm and decreases to 1.3 ± 0.1 at $z = 1.95$ nm (data not presented).

To identify the H bonds modulated by the position of ZPP, we calculated the correlation coefficient between the average

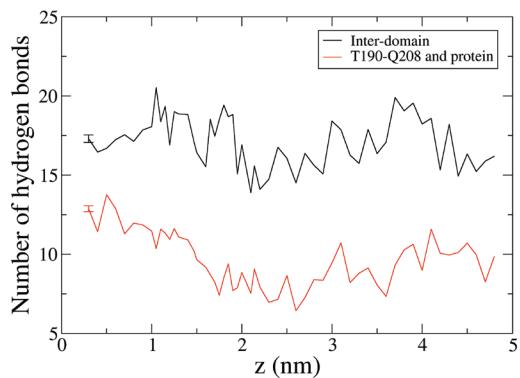


Figure 8. Number of H bonds formed between the two domains of POP (black) and between the TYR190–GLN208 flexible loop and the protein body (red) as a function of the spring equilibrium length. Maximum errors evaluated to ± 0.26 and ± 0.18 , respectively, are obtained through a 5000 bootstrap evaluation of 10% of the available data and a confidence probability of 95%.

number of H bonds per US window and the probability of existence of each individual H bond for that window. We define this probability of existence as the percentage of time over the length of a window for which the H bond exists. Table 7 presents

those H bonds with (absolute value) minimum correlation coefficients of 0.4 for windows around the energy peaks. When combined, the fluctuations of this group of H bonds (including those with negative correlation coefficients) explains 75% of all of the interdomain H-bond fluctuations. Although a few H bonds are formed and broken in the vicinity of the ZPP (SER148–ASP642, ARG128–ASP641, LYSH172–ASP642), many are located opposite the loop opening in the interface region near the velcro-rip (THR686–ASN96, LYSH75–TYR71, LEU94–ASP72, LYSH428–GLU69 in Figure 9).

Once the ZPP has moved past the second free energy maximum, it gradually loses contact with the body of the protein, only keeping contact with the TYR190–GLN208 loop (Figure 5d). Due to the higher mobility of the solvent-exposed ZPP, no strong contact dominates in this region. As the ZPP is pulled outward, the flexible loop adopts an extended conformation to maintain contacts with it. The amino acids having the most frequent contacts with the ligand are those situated on the 192–198 segment of the loop, which can conformationally extend the furthest into the solvent.

When examining the evolution of the H-bond network between the two domains of POP, we can see a drop in the average number of H bonds as the ZPP moves outward from the second free energy peak ($z = 1.8$ nm), starting from 19.4 ± 0.15

and decreasing to 13.9 ± 0.17 at $z = 2.1$ nm. This number then increases as the ZPP moves further along the trajectory up to a value of 18.4 ± 0.22 at $z = 3.0$ nm. This increase of 4.5 ± 0.4 H bonds from $z = 2.1$ nm to $z = 3.0$ nm is mainly due to an increase of 1.9 ± 0.3 H bonds between the flexible loop and the catalytic domain, indicating that the TYR190–GLN208 loop is folding back onto the protein. A correlation analysis between the presence of each individual H bond and the average number of H bonds (Table 8) shows fewer H bonds forming atom pairs with high correlations: Specifically, all of the H-bond forming pairs with an absolute value minimum correlation of 0.3 had to be selected to explain 75% of the variation of the average number of H bonds.

Interestingly, we observed a prolonged interaction between the TYR190–GLN208 loop and ZPP in the z window, where the inhibitor is completely solvated. Starting from a ZPP–POP distance of $z = 2.7$ nm, and moving further outward from the protein, the large majority of the ligand's contacts are with amino acids on the flexible loop, even though ZPP's motion is only constrained along the reaction coordinates and the ligand can move freely in the perpendicular hyperplane. For example, in the $3.2 < z < 3.7$ nm region, ZPP makes contacts with 5.4 amino acids compared to 12.6 contacts on average for the $0.3 < z < 2.0$ nm region of the trajectory (data not presented). Small in number, these interactions are nevertheless sufficient to stabilize ZPP's position and keep it in contact with the loop for at least 80% of the simulation time. This suggests that the role of the flexible loop is not simply to open up and leave a pathway open for the ZPP entrance. The TYR190–GLN208 loop could play an active role in recruiting the ZPP ligand by binding to it in the solvent and directing it to the entry pathway, helping the ZPP to go through a first free energy barrier at $z = 1.8$ nm.

4. DISCUSSION

In this work, we identify the most probable pathway for binding ZPP to POP. Using SMD and US simulations, we eliminated two proposed pathways: first, through the velcro-rip and, second, through the β -propeller tunnel. The first pathway is ruled out because we were unable to even generate a pathway using SMD, demonstrating the extreme resistance to this path.

Table 6. Average Window Probability of Existence of the Most Persistent H Bonds between the TYR190–GLN208 Loop and Protein Body in the 0.3 nm to 1.8 nm Section of the Loop Exit Pathway

donor	acceptor	presence time ratio
SER203N	LYS588O	0.42
HIS593N	THR204OG	0.37
TYR190O	TRP234O	0.36
GLN208NE	TRP234O	0.33
LYS196NZ	GLU169O	0.31
ASN188ND2	GLN208OE1	0.30
SER203OG	THR590O	0.30
TRP234N	TYR190O	0.28
LYS196NZ	TRP150O	0.28

Table 7. Hydrogen Bonds Located at the Interdomain Interface with Activity Modulated by the Position of Zpp on the Reaction Coordinate As Identified by the Pearson's Correlation Coefficient against the Average Number of H Bonds for Two Regions, $z = [1.3, 2.0]$ and $z = [1.05, 2.1]$ ^a

Hydrogen bond	Region $z = [1.3, 2.0]$			Region $z = [1.05, 2.1]$		
	Correl.(p-val.)	Prob.	Std-dev	Correl.(p-val.)	Prob.	Std-dev
GLN439NE2-GLN439HE22-ASP356O	0.51(0.044)	0.30	0.08	0.39(0.060)	0.32	0.09
THR597OG1-THR597HG1-GLY254O	0.50(0.049)	0.16	0.34	0.33(0.115)	0.11	0.28
THR686OG1-THR686HG1-ASN96OD1	0.45(0.080)	0.34	0.42	0.52(0.009)	0.32	0.42
SER148OG-SER148HG-ASP642OD2	0.44(0.088)	0.29	0.31	0.31(0.140)	0.27	0.35
LYSH75NZ-LYSH75HZ3-TYR71OH	0.43(0.096)	0.28	0.42	0.39(0.060)	0.27	0.42
SER148OG-SER148HG-ASP642OD1	0.41(0.114)	0.26	0.35	0.20(0.349)	0.26	0.32
LYSH677NZ-LYSH677HZ3-ASP122OD1	0.40(0.125)	0.33	0.37	0.26(0.220)	0.36	0.38
ARG128NH1-ARG128HH12-ASP641OD1	-0.40(0.125)	0.16	0.33	-0.40(0.053)	0.16	0.30
LYSH172NZ-LYSH172HZ3-ASP642OD2	-0.43(0.096)	0.16	0.27	0.07(0.745)	0.22	0.32
LYSH428NZ-LYSH428HZ3-GLU69OE1	-0.54(0.031)	0.77	0.31	-0.31(0.140)	0.81	0.26
LEU94N-LEU94H-ASP72OD1	-0.54(0.031)	0.19	0.35	-0.06(0.780)	0.23	0.38

^a In both cases, the average probability of existence and standard deviation of each h-bond in their respective subset of windows is also given.

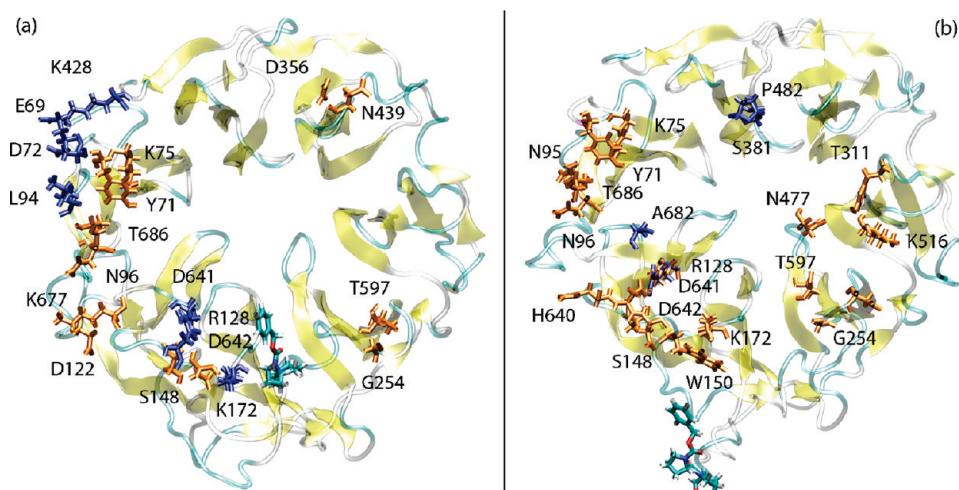


Figure 9. View of the amino acids forming interdomain H bonds modulated by the position of ZPP on the reaction coordinate as identified by Pearson's correlation coefficient of the involved H bond against the average number of H bonds. Amino acids in orange are positively correlated with the average number of H bonds, while those in blue are negatively correlated. The catalytic domain is hidden to facilitate the view. ZPP crossing $z = 1.3$ nm (a) and $z = 3.0$ nm (b).

Table 8. Hydrogen Bonds Located at the Inter-Domain Interface with Activity Modulated by the Position of ZPP on the Reaction Coordinate As Identified by the Pearson's Correlation Coefficient against the Average Number of H Bonds for the Hydrogen Bonds with the Highest Correlation for the Region $z = [1.75, 3.7]$ ^a

hydrogen bond	correl (p-val.)	prob.	std-dev
THRS97OG1–THR597HG1–GLY254O	0.50 (0.008)	0.14	0.29
TRP150N–TRP150H–ASP642OD2	0.48 (0.011)	0.15	0.29
THR686OG1–THR686HG1–ASN96OD1	0.48 (0.012)	0.34	0.40
LYSH516NZ–LYSH516HZ3–ASP256OD1	0.43 (0.025)	0.13	0.23
THR686OG1–THR686HG1–GLN95O	0.41 (0.031)	0.11	0.27
ASN477ND2–ASN477HD22–TYR311OH	0.38 (0.051)	0.21	0.18
THR686N–THR686H–ASN96OD1	0.38 (0.051)	0.26	0.35
LYSH7SNZ–LYSH7HZ3–TYR71O	0.37 (0.058)	0.31	0.40
SER148N–SER148H–HISA640O	0.36 (0.066)	0.19	0.22
ARG128NH2–ARG128HH22–ASP641OD1	0.32 (0.099)	0.19	0.31
LYSH172NZ–LYSH172HZ3–ASP642OD1	0.32 (0.106)	0.14	0.24
ARG128NH2–ARG128HH22–ALA682O	-0.31 (0.113)	0.08	0.19
SER381N–SER381H–PRO482O	-0.34 (0.086)	0.18	0.34

^aAlso included are the average probability of existence and standard deviation of each H bond in the selected subset of windows.

The resulting β -propeller exit trajectory free energy difference profile was unphysical. Detailed analysis suggests that this unphysical profile is associated with very constrained regions of the pathways where sampling is particularly difficult. On the basis of this evidence, we were also able to eliminate this pathway. The appropriate behavior of the ZPP when pulled through the flexible loop region and the physical nature of the free energy difference profile indicate that this is the correct access pathway.

Whether or not the access to the binding cavity involves a large domain reconfiguration could not be definitively resolved by our study. The natural substrates of POP, peptides of length <30 residues, are much larger than the ZPP inhibitor. Thus, the access mechanism of ZPP may not be the general pathway involved in its catalytic activity. The long-range destabilization of the H-bond network that was seen to occur as the ZPP left the binding cavity

could possibly be interpreted as evidence of the large scale interdomain motion²⁷ that has been hypothesized to play a role in the access pathway. In addition, the prolonged association of the TYR190–GLN208 loosely structured loop with the ZPP as it left the protein provides evidence that this region of the protein could possibly have a role in ligand recognition and recruitment.

AUTHOR INFORMATION

Corresponding Author

*E-mail: normand.mousseau@umontreal.ca.

ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council of Canada (M.K., N.M.), the Academy of Finland (T.R., M.K.), Canada Research Foundation (N.M.), and the GALENOS program (J.-F.S.-P.). We are grateful to the Réseau québécois de calcul de haute performance (RQCHP), SharcNet, and the Finnish IT Centre for Science for their generous allocation of computer time.

REFERENCES

- Polgár, L. *Cell. Mol. Life Sci.* **2002**, 59, 349–362.
- Garcia-Horsman, J. A.; Männistö, P. T.; Venäläinen, J. I. *Neuropeptides* **2007**, 41, 1–24.
- Rea, D.; Fülop, V. *Cell Biochem. Biophys.* **2006**, 44, 349–365.
- Brandt, I.; Scharpé, S.; Lampier, A.-M. *Clin. Chim. Acta* **2007**, 377, 50–61.
- Szeltner, Z.; Polgár, L. *Curr. Prot. Pept. Sci.* **2008**, 9, 96–107.
- Yoshimoto, Y.; Kado, K.; Matsubara, F.; Koriyama, N.; Kaneto, H.; Tsura, D. *J. Pharmacobiodyn.* **1987**, 10, 730–735.
- Männistö, P. T.; Venäläinen, J. I.; Jalkanen, A.; Garcia-Horsman, J. A. *Drug News Perspect.* **2007**, 20, 293–305.
- Mantle, D.; Falkous, G.; Ishiura, S.; Blanchard, P. J.; Perry, E. K. *Clin. Chim. Acta* **1996**, 249, 129–139.
- Tenorio-Laranga, J.; Coret-Ferrer, F.; Casanova-Estruch, B.; Burgal, M.; García-Horsman, J. A. *J. Neuroinflammation* **2010**, 7, 23.

- (10) Di Daniel, E.; Glover, C. P.; Grot, E.; Chan, M. K.; Sanderson, T. H.; White, J. H.; Ellis, C. L.; Gallagher, K. T.; Uney, J.; Thomas, J.; Maycox, P. R.; Mudge, A. W. *Mol. Cell. Neurosci.* **2009**, *41*, 373–382.
- (11) Schulz, I.; Zeitschel, U.; Rudolph, T.; Ruiz-Carrillo, D.; Rahfeld, J.-U.; Gerhartz, B.; Bigl, V.; Demuth, H.-U.; Rossner, S. *J. Neurochem.* **2005**, *94*, 970–979.
- (12) O'Reilly, P. J.; Hardison, M. T.; Jackson, P. L.; Xu, X.; Snelgrove, R. J.; Gaggard, A.; Galin, F. S.; Blalock, J. E. *J. Neuroimmunol.* **2009**, *217*, 51–54.
- (13) Gaggard, A.; Jackson, P. L.; Noerager, B. D.; O'Reilly, P. J.; McQuaid, D. B.; Rowe, S. M.; Clancy, J. P.; Blalock, J. E. *J. Immunol.* **2008**, *180*, 5662–5669.
- (14) Liu, J.-M.; Kusinski, M.; Ilic, V.; Bignon, J.; Hajem, N.; Komorowski, J.; Kuzdaz, K.; Stepien, H.; Wdzieczak-Bakala, J. *Anticancer Res.* **2008**, *28*, 2813–2817.
- (15) van Gool, A. R.; Verkerk, R.; Fekkes, D.; Sleijfer, S.; Bannink, M.; Kruit, W. H.; van der Holt, B.; Scharpé, S.; Eggermont, A. M. M.; Stoter, G.; Hengeveld, M. W. *J. Interferon Cytokine* **2008**, *28*, 283–286.
- (16) Lawandi, J.; Gerber-Lemaire, S.; Juillerat-Jeannere, L.; Moitessier, N. *J. Med. Chem.* **2010**, *53*, 3423–3438.
- (17) Lawandi, J.; Toumieux, S.; Seyer, V.; Campbell, P.; Thielges, S.; Juillerat-Jeanneret, L.; Moitessier, N. *J. Med. Chem.* **2009**, *52*, 6672–6684.
- (18) Tarrago, T.; Martin-Benito, J.; Sabido, E.; Claasen, B.; Madurga, S.; Gairi, M.; Valpeusta, J. M.; Giralt, E. *FEBS Lett.* **2009**, *583*, 3344–3348.
- (19) Comellas, G.; Kaczmarśka, Z.; Tarrago, T.; Teixidó, M.; Giralt, E. *PLoS One* **2009**, *4*, e6222.
- (20) Fülpö, V.; Böcskei, Z.; Polgár, L. *Cell* **1998**, *94*, 161–170.
- (21) Fülpö, V.; Jones, D. T. *Curr. Opin. Struct. Biol.* **1999**, *9*, 715–721.
- (22) Fülpö, V.; Szeltner, Z.; Polgár, L. *EMBO Rep.* **2000**, *1*, 277–281.
- (23) Szeltner, Z.; Renner, V.; Polgár, L. *J. Biol. Chem.* **2000**, *275*, 15000–15005.
- (24) Juhász, T.; Szeltner, Z.; Fulop, V.; Polgár, L. *J. Mol. Biol.* **2005**, *346*, 907–917.
- (25) Fuxreiter, M.; Magyar, C.; Juhász, T.; Szeltner, Z.; Polgár, L.; Simon, I. *Proteins Struct. Funct. Bioinf.* **2005**, *60*, 504–512.
- (26) Kaszuba, K.; Rög, T.; St. Pierre, J.-F.; Männistö, P. T.; Karttunen, M.; Bunker, A. *SAR QSAR Environ. Res.* **2009**, *20*, 595–609.
- (27) Shan, L.; Mathews, I. L.; Khosla, C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3599–3604.
- (28) Li, M.; Chen, C.; Davies, D. R.; Chiu, T. K. *J. Biol. Chem.* **2010**, *285*, 21487–95.
- (29) Szeltner, Z.; Rea, D.; Juhász, T.; Renner, V.; Fülpö, V.; Polgár, L. *J. Mol. Biol.* **2004**, *340*, 627–637.
- (30) Kaszuba, K.; Rog, T.; Bryl, K.; Vattulainen, I.; Karttunen, M. *J. Phys. Chem. B* **2010**, *114*, 8374–8386.
- (31) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (32) Nosé, S. *Mol. Phys.* **1984**, *52*, 255–268.
- (33) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (34) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (35) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (36) Karttunen, M.; Rottler, J.; Vattulainen, I.; Sagui, C. *Curr. Top. Membr.* **2008**, *60*, 49–89.
- (37) Wong-ekkabut, J.; Miettinen, M. S.; Dias, C.; Karttunen, M. *Nat. Nanotechnol.* **2010**, *5*, 555–557.
- (38) Jorgensen, W.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (39) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (41) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
- (42) Jarzynski, C. *J. Stat. Phys.* **2000**, *98*, 77–102.
- (43) Jarzynski, C. *Phys. Rev. E* **1997**, *56*, 5018–5035.
- (44) Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120*, 5946–5961.
- (45) Sun, S. X. *J. Chem. Phys.* **2003**, *118*, 5769–5775.
- (46) Xiong, H.; Crespo, A.; Martí, M.; Estrin, D.; Roitberg, A. E. *Theor. Chem. Acc.* **2006**, *116*, 338–346.
- (47) Ytreberg, F. M.; Zuckerman, D. M. *J. Chem. Phys.* **2004**, *120*, 10876.
- (48) Cuendet, M. A.; Michelin, O. *Biophys. J.* **2008**, *95*, 3575–3590.
- (49) Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120*, 5946–5961.
- (50) Oberhofer, H.; Dellago, C.; Geissler, P. L. *J. Phys. Chem. B* **2005**, *109*, 6902–6915.
- (51) Baştug, T.; Kuyucak, S. *Chem. Phys. Lett.* **2007**, *436*, 383–387.
- (52) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (53) Roux, B. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
- (54) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (55) Chen, P.-C.; Kuyucak, S. *Biophys. J.* **2009**, *96*, 2577–88.
- (56) Mustafa, M.; Henderson, D. J.; Busath, D. D. *Proteins* **2009**, *76*, 794–807.
- (57) Mills, M.; Andricioaei, I. *J. Chem. Phys.* **2008**, *129*, 114101.
- (58) Xing, C.; Faller, R. *J. Chem. Phys.* **2009**, *131*, 175104.
- (59) Yesylevskyy, S.; Marrink, S.-J.; Mark, A. E. *Biophys. J.* **2009**, *97*, 40–9.
- (60) Baştug, T.; Chen, P.-C.; Patra, S. M.; Kuyucak, S. *J. Chem. Phys.* **2008**, *128*, 155104.
- (61) Rodriguez-Gomez, D.; Darve, E. *J. Chem. Phys.* **2004**, *120*, 3563–3578.
- (62) Ytreberg, F. M.; Swendsen, R. H.; Zuckerman, D. M. *J. Chem. Phys.* **2006**, *125*, 184114.
- (63) Tskhovrebova, L.; Trinick, J.; Sleep, J. A.; Simmons, R. M. *Nature* **1997**, *387*, 308–312.
- (64) Liu, M.; Sun, T.; Hu, J.; Chen, W.; Wang, C. *Biophys. Chem.* **2008**, *135*, 19–24.
- (65) Hub, J. S.; de Groot, B. L.; van der Spoel, D. *J. Chem. Theory Comput.* **2010**, *6*, 3713–3720.
- (66) Bakker, A. V.; Jung, S.; Spencer, R. W.; Vinick, F. J.; Faraci, W. S. *Biochem. J.* **1990**, *271*, 559–562.
- (67) Szeltner, Z.; Renner, V.; Polgár, L. *Protein Sci.* **2000**, *9*, 353–360.