Pyrococcus furiosus Prolyl Oligopeptidase: A Dynamic Supramolecular Host for Peptidase and Dirhodium Catalysis

Ken Ellis-Guardiola,[3,‡] Huan Rui,[2,‡] Ryan L. Beckner,[3] Poonam Srivastava,[3] Narayanasami Sukumar,[4*] Benoît Roux,[2,3*] Jared C. Lewis[1*]

[1]Department of Chemistry, Indiana University, Bloomington, IN 47405.

[2]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637.

[3]Department of Chemistry, University of Chicago, Chicago, IL 60637.

[4]NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Building 436E, Argonne National Laboratory, Argonne, IL 60439.

ABSTRACT: Supramolecular catalysis involves the design and characterization of synthetic macromolecules that catalyze chemical reactions. While enzymes are often cited as the inspiration for such catalysts, enzymes can also serve as hosts for non-native catalytic components. Protein-based hosts can be readily produced in *E. coli* and rapidly evolved for particular applications. Moreover, inherent properties of these systems, including their conformational dynamics, can be exploited for non-native transformations that occur within their interior. Studies on the peptidase activity of a prolyl oligopeptidase from Pyrococcus furiosus (*Pfu* POP) suggest that its unique two-domain architecture regulates substrate access and specificity. We have established that *Pfu* POP also serves as an efficient host for asymmetric cyclopropanation upon active-site modification with a dirhodium cofactor. To understand how *Pfu* POP controls both peptidase and dirhodium catalysis, we determined the crystal structures of this enzyme and its S477C mutant and used these

structures as starting points for MD simulations of both the apo structures and systems containing a covalently linked peptidase inhibitor or a dirhodium catalyst. *Pfu* POP was crystalized in an open conformation, and MD simulations reveal spontaneous transitions between open and closed states, in addition to a number of smaller scale conformational changes, suggesting facile inter-domain movement. Importantly, key aspects of previously reported peptidase kinetics and cyclopropanation selectivity can be rationalized in the context of this inter-domain opening and closing. This finding constitutes a remarkable example in which the conformational dynamics of a supramolecular host affect two different catalytic activities and suggests that *Pfu* POP could serve as a host for a wide range of non-native catalysts.

Introduction

Supramolecular catalysts afford chemists the opportunity to control the reactivity of guest molecules (e.g. substrates, catalysts, etc.) using various weak non-covalent interactions.[1,2] In this regard, these systems mimic enzymes that evolved in nature to catalyze a wide array of chemical reactions.[3] Enzymes can exhibit high catalytic proficiency[4] and the ability to override substrate controlled reactivity by regulating access to their active sites via conformational dynamics spanning a range of length and time scales[5]. Despite the mechanistic complexity that underlies these features, enzymes often catalyze non-native reactions[6] and can be evolved to possess improved activity, altered selectivity, or expanded substrate scope[7]. Even more remarkably, synthetic modifications can augment or replace the native catalytic capabilities of enzymes and thus exploit the supramolecular properties of these hosts for non-native catalysis.[8,9] Understanding the origins of this versatility--how enzymes and proteins in general can serve as hosts for both native and synthetic catalytic components and how conformational dynamics influence host

function and catalysis--would greatly improve our ability to harness proteins as supramolecular hosts for non-native catalysis.[10]

Several features of prolyl oligopeptidases (POPs)[11-13] led us to consider how these enzymes control reactions that occur within their active sites. POPs hydrolyze small peptide substrates following proline residues. With molecular weights of ~70-80 kDa, POPs are significantly larger than many prototypical proteases like trypsin or subtilisin (25-30 kDa), suggesting that the additional bulk of the former may contribute to their unique substrate specificity.[13,14] Indeed, structural studies have established that POPs are comprised of a ~30 kDa peptidase domain containing a Ser-His-Asp catalytic triad topped by a ~40 kDa 7-bladed β-propeller domain that restricts substrate access to a large active site cavity.[13,15-19] Experimental and computational studies have suggested that this domain undergoes large scale conformational changes that regulate substrate access to the active site. In effect, POP acts as a dynamic supramolecular host for peptide substrates that also harbors catalytic residues needed to hydrolyze those substrates.

Despite extensive structural characterization of bacterial and mammalian POPs, no structures for archaeal POPs have been reported. Particularly notable in this regard, given the extensive biophysical characterization dedicated to it,[20-24] is the POP from the hyperthermophilic archaeon, Pyrococcus furiosus (*Pfu*). *Pfu* POP is at most 28% identical to previously crystallized POPs, which are up to 44% identical, suggesting that *Pfu* POP could possess properties not present in these enzymes. Indeed, functional differences have been noted and most extensively discussed relative to porcine (*Sus scrofa*) POP. While porcine POP has melting points of 44.6 °C and 52.8 °C (pH 8), for example, *Pfu* POP melts at 91.7 °C and 109.5 °C (pH 8.4)[24] and exhibits maximum activity at 85 °C[20]. Both enzymes are activated by NaX (X = Cl, Br, etc.), but the nature of their response to [NaX] differs. Finally, while peptidase catalysis by both porcine and *Pfu* POP proceeds

via the general mechanism shown in Fig. 1A, substrate entry into the active sites of these enzymes and inter-domain conformational changes associated with this process appear to involve important differences (Fig. 1B).[25]

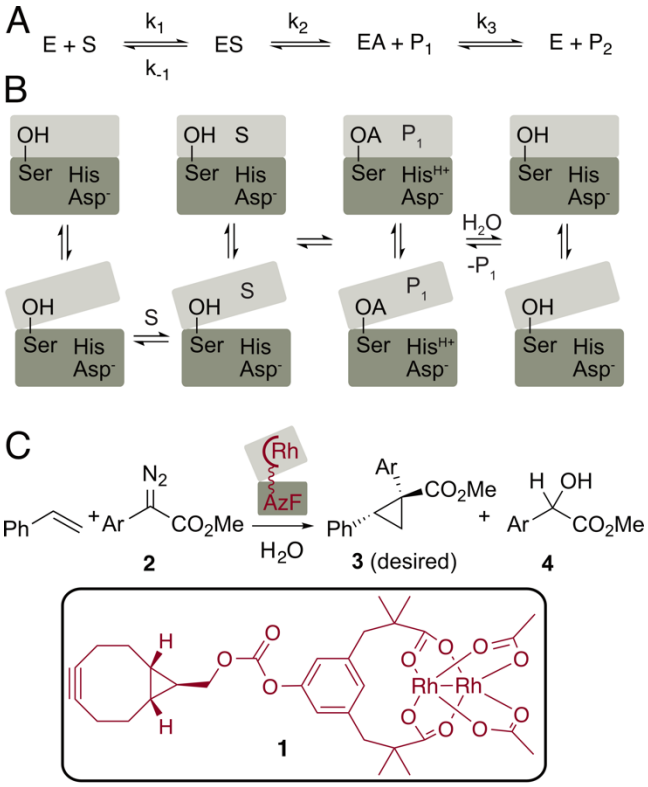

Figure 1. A) General scheme for peptidase catalysis involving an enzyme (E) and a peptide substrate (S) to generate two product peptides (P1 and P2) via enzyme-substrate (ES) and enzyme-acyl (EA) intermediates. B) Potential domain opening and closing during *Pfu* POP peptidase catalysis. C) Styrene cyclopropanation catalyzed by an artificial metalloenzyme generated by covalently linking cofactor **1** within the *Pfu* POP active site.

Our interest in *Pfu* POP was piqued by our finding that this enzyme serves as a uniquely effective host for dirhodium artificial metalloenzyme (ArM)[9,26] formation (Fig. 1C). Dirhodium

ArMs were generated by replacing the catalytic serine (S477) in *Pfu* POP with a genetically encoded azidophenylalanine residue and reacting the resulting protein with an alkyne-substituted dirhodium cofactor (**1**).[27] The resulting ArMs possess high stability and catalyze olefin cyclopropanation with high yield, enantioselectivity, and specificity for carbene insertion into olefins to give **3** over insertion into the O-H bond of water (the primary reaction catalyzed by the cofactor alone) to give **4** (Fig. 1C). Similar to the native enzyme, ArM activity was strongly dependent on [NaX], but the proposed binding of a histidine residue in the propeller domain to the dirhodium cofactor anchored in the peptidase domain would lead to an inter-domain crosslink, presumably altering native conformational dynamics. Despite the potential for crystallographic characterization to clarify different aspects of peptidase[13,22,24] and dirhodium[9,26] catalysis within the *Pfu* POP active site, and to thereby assist efforts to use this enzyme as a supramolecular host for catalysis in general, the crystal structure of this enzyme has remained elusive.

Herein, we report the structures of both *Pfu* POP and *Pfu* POP S477C. These structures, along with molecular dynamics simulations based on them, comparative analysis of previously reported POP structures, and existing biochemical data, have resolved several previously debated aspects of *Pfu* POP structure and function. These studies also suggest that the same large scale conformational dynamics control the selectivity and specificity of both the peptidase and cyclopropanase activity of POP-based systems. Together, these findings illustrate the utility of *Pfu* POP as a robust supramolecular host for catalysis.

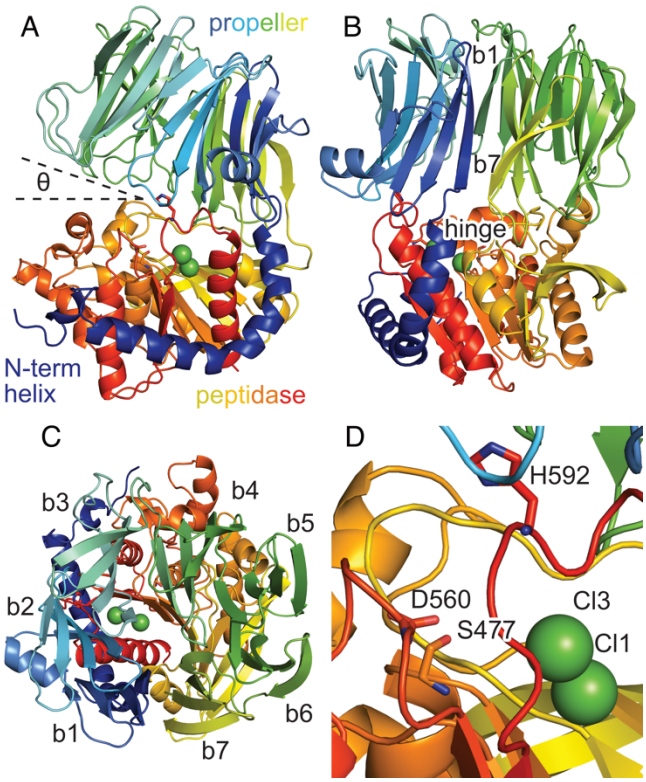

Figure 2. A-C) Annotated ribbon diagram of *Pfu* POP showing inter-domain angle (θ), blade (b1-7) and hinge locations, and key active site residues and Cl^- ions. D) expanded view of POP active site showing residues in the catalytic triad and bound chloride ions.

Results

Pfu POP Structural Overview

The structures of wild type *Pfu* POP and the corresponding S477C mutant were solved at 1.9 Å and 2.2 Å resolution (Table 1). As observed in previously reported POP structures,[12] both *Pfu* POP and the S447C mutant possess a two-domain architecture involving a peptidase domain with an α/β-hydrolase fold capped by a 7-bladed β-propeller domain (Fig. 2A-C). The N-terminus of the enzyme (residues 1-47) consists of an α-helical segment that wraps around the C-terminal peptidase domain (residues 367-616). The intermediate segment (residues 48-366) comprises the

β-barrel domain. The first and seventh blade of the propeller domain (residues 48-83 and 321-366, respectively) are joined by a number of hydrogen bonding and ion-pairing interactions in addition to hydrophobic interactions between the blade surfaces. The peptidase and β-barrel domains are joined covalently via a hinge region comprised of residues 47-50 and 361-367. An inter-domain angle of ~30° allows access to a large cavity at the domain interface. Residues comprising the catalytic triad reside at the peptidase/barrel interface, with S477 located centrally within the peptidase domain and D560 and H592 located in loops at the peptidase periphery (Fig. 2D).

Table 1. X-ray Diffraction Data Collection and Refinement Statistics[a]

	Pfu POP (5T88)	S477C (6CAN)
Data collection		
wavelength (Å)	0.98	0.98
space group	P21	P21
unit cell dimensions		
a, b, c (Å)	55.5, 176.8, 57.9	56.3, 178.8, 59.3
α, β, ϒ (°)	90, 106.0, 90	90, 104.4, 90
res. limit (Å)	50-1.9 (1.97-1.9)	100-2.2 (2.28-2.2)
I/Sigma(I)	12.3 (1.5)	9.5(1.2)
R_{pim} (%)	5.9(48.3)	7.1(44.0)
$CC^{1/2}$(%)	57.5	53.0
completeness (%)	99.9(100.0)	93.7(74.6)
redundancy	9.7 (9.3)	4.5(3.5)
Refinement		
resolution range (Å)	47.1 - 1.9	19.5 - 2.2
R_{work} (%)	19.3	19.7
R_{free} (%)	24.7	25.1
R (working + test) (%)	19.4	19.7
no. of reflections	83441	53366
Model		
no. of amino acids	1232	1232
no. of H_2O mol.	946	842
no. Cl^- ions	4	6
no. residues:		
in generously allowed region	12	13
in disallowed regions	0	0
Stereochemical ideality		

bonds (Å)	0.01	0.01
angles (°)	1.44	1.24
dihedral angles (°)	11.39	16.92
planarity (Å)	0.01	0.00

[a]Value in parentheses are for the outer shell.

Inter-domain Loop Composition and Structure

The composition and conformation of loops at the domain interface of different POPs possess many unique features,[28] and notable differences between the loops in these enzymes and *Pfu* POP are likewise apparent (Fig. S1). Extensive ion pairing is observed between residues located in these loop regions in closed POP structures,[17-19,22,29] but only a single inter-domain ion pair is present in the *Pfu* POP crystal structures. The loop containing the catalytic histidine (H592, residues 588-597) possesses a relatively high average B-factor in both structures (46.8/67.9 $Å^2$ versus 32.4/45.6 $Å^2$ for the full chains of *Pfu* and S477C, respectively, Table S1), the density of the main chain and the residues of this loop were apparent in the observed electron density (Fig. 3A, S11). This finding is notable because the histidine loop is disordered in all previously reported structures of POPs that crystallized in their open form[18,19]. Finally, the loop connecting β-strands 2 and 3 of blade 3 of the POP β-propeller domain (residues 158-169) is significantly shorter than in previously characterized homologues, and it is folded back onto the surface of the β-propeller domain (Fig. 3B). This conformation is enforced by two sequential proline residues (P168 and P169) and a salt bridge between the end of the loop (D164) and the β-propeller domain (R172). This orientation contrasts with other POP structures in which residues near the end of this loop often exhibit hydrogen-bonding and/or salt bridging interactions with the peptidase domain that "latch" the enzyme shut.[18]

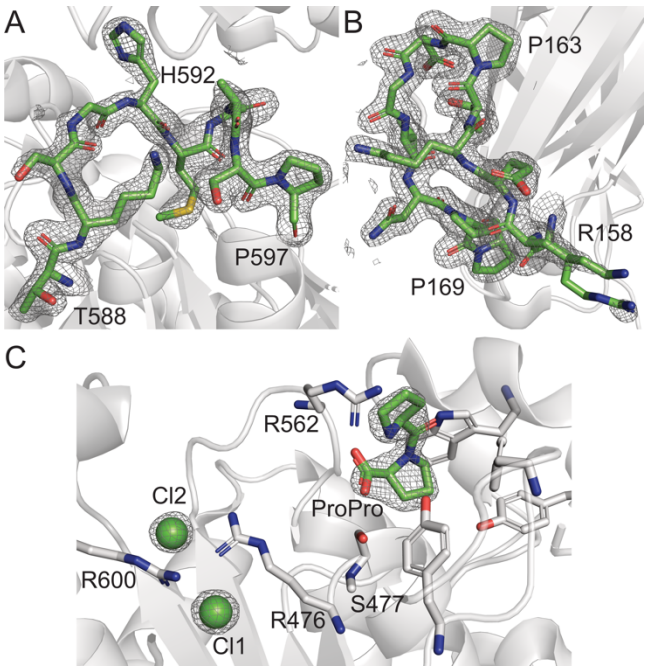

Figure 3. Electron density (contoured at 1.0 σ) for key features in the *Pfu* POP crystal structure: A) the loop containing His592, B) the "latch loop", C) bound prolylproline and chloride ligands.

Prolylproline and Chloride Binding within Pfu POP.

The active site of *Pfu* POP exhibits many features common to other POPs crystallized to date.[13] The hydrophobic proline-binding site (S1), comprised of F404, W518, and Y522, is conserved among all previously crystalized POPs,[13] and the active-site serine (S477) is predictably positioned within a GXSXGG motif (Fig. 3C)[30]. Notably, however, both chains of the *Pfu* POP structure contain electron density consistent with a bound prolylproline ligand, albeit with a significantly higher B-factor (58.0 $Å^2$) than the main chains (32.4 $Å^2$). The orientation of this ligand is similar to proline-based inhibitors in previously reported crystal structures of porcine POP in a closed conformation (Fig. S10).[31] In these structures, the proline residue binds in a hydrophobic S1 pocket while the carbonyl of the P1-P2 amide bond forms a hydrogen bond with R643.

Similarly, in the *Pfu* POP structure, the C-terminal proline carboxylate of the prolylproline ligand forms an ion pair with the homologous R562 residue.

The *Pfu* POP crystal structures also possess multiple Cl^- ions bound within the active site (Fig. 3C). WT *Pfu* POP contains two Cl^- ions bound via opposing arginine residues (R476 and R600) proximal to the active site serine. Each Cl^- is coordinated in a bidentate, end-on fashion to one arginine and in a monodentate, side-on fashion to the second with B-factors of 45.4 and 19.8 $Å^2$, respectively. In the S477C mutant, both of these sites are occupied by Cl^- ions with B-factors of 58.3 and 39.0 $Å^2$, and an additional chloride is also bound in the P1 pocket with a B-factor of 53.1 $Å^2$.

MD Simulations of Pfu POP Domain Dynamics

Microsecond MD simulations were carried out to investigate the domain dynamics of *Pfu* POP at 300 K and its optimal temperature for peptidase catalysis (358.15 K). Simulations were conducted in triplicate on wildtype *Pfu* POP (WT), WT containing a covalently linked ZPR inhibitor (IHBT), and the inactivated S477C mutant of WT (S477C), all hereafter referred to as "apo" states because they lack a covalently attached cofactor. Two simulations at 300K and 358.15 K were carried out on a model of a POP-based dirhodium ArM containing cofactor **1** (Fig. 1C).

The simulations show that the apo systems open and close spontaneously as the peptidase and the propeller domains pivot about a hinge region (Fig. 1B). An inter-domain angle (θ) was defined (Fig. 1A, S2) to quantify domain opening, and the time evolution of θ was evaluated at both 300K and 358.15K (Figs. 4 and S3). The open/closed transition occurs at a θ of ~17-23°, and θ ranges from 8-55°. At 358.15 K, the average θ is 20.0° ± 1.8°, 24.4° ± 4.2°, and 23.3° ± 0.9° in the WT, IHBT, and S477C simulations, respectively (Table S2). In general, the high temperature

simulations showed increased fluctuation of θ and more open/close transitions. The dirhodium ArM exhibits qualitatively similar dynamics to the apo systems (Fig. 4), indicating that both open and closed conformations can be sampled despite the presence of the bulky cofactor.

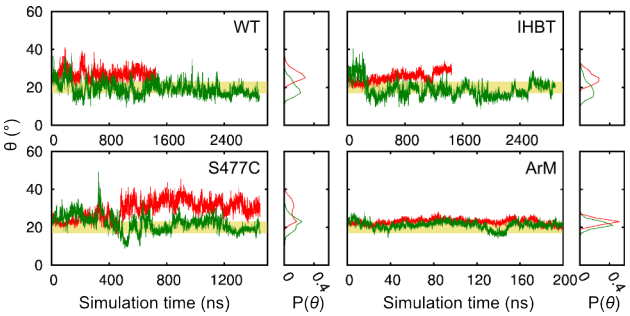

Figure 4. The open angle and its distribution in the MD simulations at 300 K (red) and 358.15 K (green) in ArM and replica 2 of the WT, IHBT, and S477C systems. The yellow banner indicates the intermediate open angle values between the open and the closed states.

Inter-domain opening renders the active site solvent-accessible, which can be visualized in trajectories from the IHBT simulations (Fig. 5A and B). This opening can also be seen in the ArM simulation (Fig. 5C). To quantitate this opening, the solvent-accessible area between each propeller blade and the peptidase domain was calculated for the apo systems (Fig. S4). This analysis shows that the largest opening occurs between blade 3 of the propeller domain and the peptidase domain (Tables S3 and S4). In the WT simulations, the average solvent accessible area is 219 ± 74 Å^2, even at the most constricted region along the blade 3 opening axis. Large fluctuations in the loop spanning D505-Y522 in the peptidase domain, opposite to blade 3, were also observed (Fig. S5). In contrast, the narrowest opening in the central pore of the propeller domain, which has been proposed as an alternate substrate entry site[10], has a solvent-accessible area of only 93 ± 24 Å^2 in the WT simulations. No significant opening between blades 1 and 7 of the propeller domain was observed in any of the simulations.

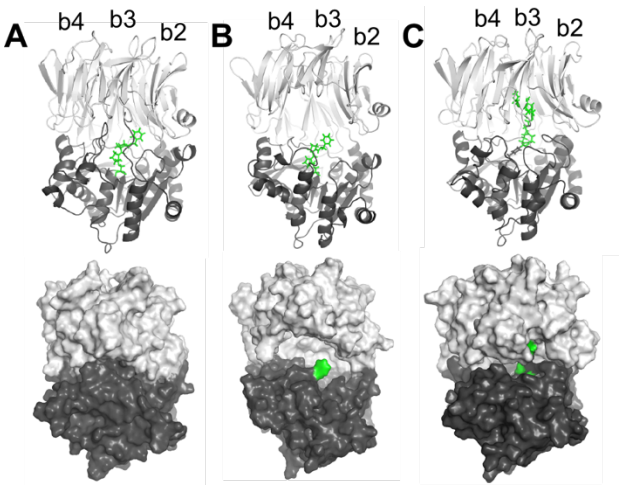

Figure 5. The cartoon and surface representation showing the largest opening between propeller (light grey) blade b3 and the peptidase domain (dark grey). The snapshots are taken from the covalently-linked inhibitor simulation IHBT_3 to illustrate (A) the closed state and (B) the open, substrate-accessible state enzyme. An open conformation of the *Pfu* POP from the ArM simulation is also shown in (C). The bottom panel shows a surface presentation of the same snapshot. To highlight the covalently bound inhibitor and the ArM, they are colored green.

Histidine Loop Conformational Dynamics

Large-scale domain movements are accompanied by smaller-scale conformational changes throughout the structures examined, but particularly notable are the dynamics of the histidine (H592) loop. In two simulations, the histidine loop achieves a conformation that places H592 within 4 Å from S477, its approximate orientation in crystal structures of POP enzymes in their closed forms (Fig. 6). The process of loop closing primarily involves changes in the ϕ - ψ backbone dihedral angles (Fig. S6) and sequential interaction of H592 with hydrogen bond accepting residues (Fig. 7). Histidine loop dynamics were also evaluated starting from a model of *Pfu* POP in the closed form (see supporting information). The similar ϕ /ψ dihedral angles

distribution from the trajectories of the engineered His loop and WT systems suggest that similar dynamics are observed from either the open or closed state of *Pfu* POP (Fig. S6).

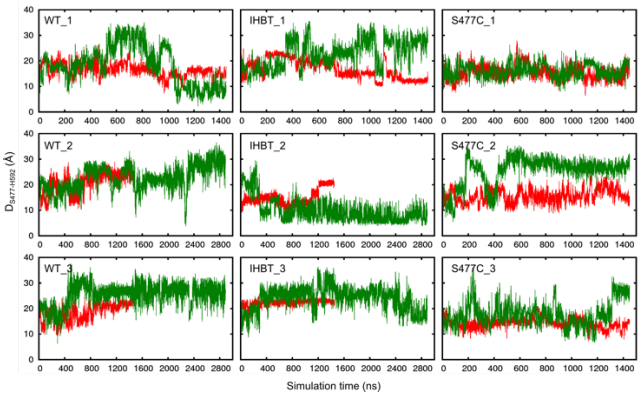

Figure 6. The time evolution of the S477 and H592 distance, DS477-H592. The hydroxyl oxygen (OG) in S477 and the imidazole amide hydrogen (HE2) are used to define the distance.

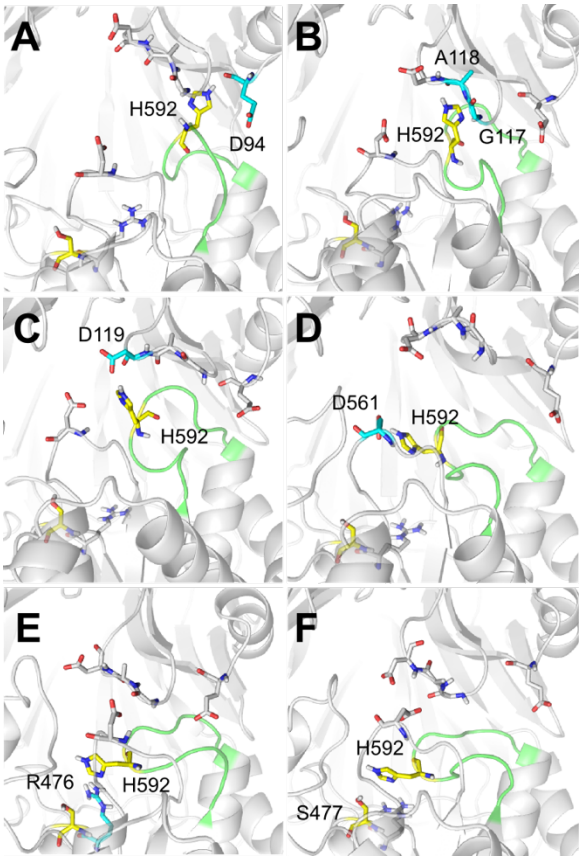

Figure 7. The H-bond network for H592 entry into the catalytic site. The H592 loop is colored green, H592 and S477 are colored yellow, and H-bonding residues are colored cyan.

Effects of Halide Binding on Histidine Loop Conformational Dynamics

As noted above, there are two Cl^- binding sites in the WT *Pfu* POP crystal structure and three in that of the S477C mutant. The sites involving bidentate Cl^- coordination by the guanidinium groups of R476 and R600, respectively, are common to both structures. The bidentate-R600 site has the lowest B-factors in the crystal structures, an intermittent binding to this site is observed in MD simulations. Starting from crystal structures with this site occupied by a Cl^-, both dissociation and association of Cl^- can be observed (Table S6); Cl^- ions bound at other sites dissociate early in the simulations and do not again associate. Interestingly, Cl^- binding to the R476/R600 site has a minor effect on the S477-H592 distance distribution in the WT and the inhibitor bound systems, but not for S477C (Fig. 8). The effect manifests itself as a shift in the distance distribution towards the lower end of the ranges observed (a shorter S477-H592 distance). The Cl^- binding also affects the inter-domain angle but to a lesser extent (Fig. S7).

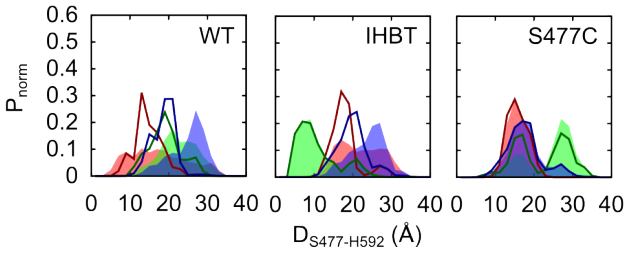

Figure 8. The impact on S477-H592 distance from the binding of Cl^- to the R600 site. The site is occupied when the Cl^- is within 1.5 Å to the crystal binding site of this Cl^-. The transparent, filled curves show the normalized population (P_{norm}) distributions of DS477-H592 computed from the entire trajectory at 358.15 K. The thick, colored lines are the P_{norm} distributions of DS477-H592 drawn from when Cl^- is bound to the R600 site. Three replicas are shown in red, green, and blue.

Discussion

Structural Origins of Pfu POP Stability

Stability improves the utility of enzymes for biocatalysis and for use as supramolecular hosts. The stability of *Pfu* POP makes this enzyme a promising candidate for both applications, and the crystal structures reported herein provide a number of insights into the origins of its stability. Comparing the molecular weights of *Pfu* POP (70.8 kDa) to those of the other POPs crystallized to date (76-83 kDa) highlights the relatively small size of the former. Aligning the *Pfu* POP sequence to those of other POPs crystalized to date shows that this smaller size results primarily from shortened loop regions between structurally conserved β-sheets in the propeller domain (Fig. S8). The structures of *Pfu* POP and the corresponding S477C mutant also have an average of 0.10-0.11 ion pairs/residue versus 0.061-0.090 for the other (mesophilic) POP enzymes crystalized to date (Table 2). Shortened loop regions and increased ion pair content have both been proposed to contribute to the stability of hyperthermophilic enzymes.[32,33]

Table 2. Structural data for previously crystalized POPs.

POP[a]	MW[b]	Additional Residues[c]		Ion Pairs[d]		Inter-domain Angle (°)[d]
		Peptidase domain	Propeller domain	Per residue	Inter-domain	
Pfu	70824	-	-	0.11	1	29
S477C	70840	-	-	0.11	0	32
*Pfu*_mod	70824	-	-	0.084	5	19
M. xanthus	76848	14	28	0.086	5	21
R. typhi	83074	27	54	0.061	3	23
S. scrofa	80770	27	40	0.076	4	22
H. sapiens	80700	27	40	0.061	3	22
A. caviae_c	76467	19	24	0.078	2	25
A. caviae_o				0.070	0	47
S. capsulata	78435	11	34	0.064	0	58

[a]*Pfu* chain B from 5T88; *Pfu* S477C chain B from 6CAN; *Pfu*_mod from a previously reported homology model; *M. xanthus* chain A from 2BKL; *R. typhi* from 4HVT; *S. Scrofa* from 1QFS; *H.*

sapiens from 3DDU; *A. caviae_c* from 3IVM; *A. caviae_o* from 3IUJ; *S. capsulata* from 1YR2. [b]Determined using the ExPASy server. [c]relative to *Pfu*. [d]Determined using VMD or UCSF Chimera as described in the SI.

Conformational Dynamics of Pfu POP

Despite the stability of *Pfu* POP, earlier kinetic analyses of this enzyme suggested that it might undergo significant conformational changes during catalysis.[22-24] Both open[18,19] and closed[17-19,22,29] conformations of previous POP crystal structures or homology models can be distinguished by aligning their peptidase domains to that of the *Pfu* POP crystal structures (RMSD = 0.18-1.06 $Å^2$, Fig. S8). Using the inter-domain angle (θ) definition described above, closed conformations have θ < 25°, while θ values of 47° and 58° were reported for open structures. At 29 and 32°, θ values for *Pfu* POP fall between these extremes, suggesting that it crystalized in an intermediate conformation. The single inter-domain ion pair in *Pfu* POP is also intermediate between these enzymes (Table 2), suggesting that ions pairs that could contribute to the stability of the closed form of the enzyme have only partially formed.

While crystal structures show that different domain angles can be achieved in different structures (Fig. S1), and both NMR spectroscopy[34,35] and MD simulations have been used to study conformational dynamics in individual POPs, none of the previous MD studies reported spontaneous domain opening or closing during simulations[25,28,36-39]. The microsecond MD simulations conducted herein show, for the first time, spontaneous conversion between the open and closed conformations of POP molecules (Fig. 3 and S3). The peptidase and propeller domains move largely as rigid bodies with important exceptions outlined below. Average θ values are similar to those observed in the crystal structures (Tables 2 and S2), but θ values ranging from 8-

55° are sampled, consistent with the notion that the 29 or 32° angles observed in the crystal structures are intermediate in nature. The inter-domain ion pairs observed in the crystal structures are observed in all MD simulations, and a negative correlation exists between the number of ion pairs and the inter-domain angle (Fig. S9). The covalently bound ZPR ligand in the IHBT system resulted in slightly larger θ values at 358.15 K, with an average θ of 24.4° ± 4.2° compared to the 20.0° ± 1.8°. This result contrasts with experimental findings for porcine POP, which suggested that inhibitor binding favors the closed form of the enzyme.[40] This difference could reflect either fundamental differences in how the dynamics of *Pfu* and porcine POP respond to ZPR binding or insufficient simulation time to sample the lowest energy substrate-bound closed state of *Pfu* POP. Converting the wildtype enzyme to the corresponding ArM, which contains a covalently linked dirhodium cofactor, does not significantly impact POP conformational dynamics. Open/closed transitions and a similar average θ value are observed during the simulation (Fig. 4).

Large-scale domain dynamics of POPs have been proposed to be modulated by a so-called "latching loop" (i.e., R158-P169 in *Pfu* POP). Previously reported POP structures show this loop bound to the peptidase domain in closed structures or disordered in open structures.[19,29] The "latching" action of the loop observed in these structures is primarily mediated through polar and ionic interactions. A homology model of *Pfu* POP built from the porcine structure predicted that this loop would extend to the peptidase domain,[22] contrary to the conformation observed in the *Pfu* POP crystal structures. The conformation of the "latching loop" in *Pfu* POP remains relatively constant throughout all MD simulations and makes minimal contact with the hydrolase domain in the closed structures from the simulations. There is no persisting pattern of interactions between the "latching loop" and the hydrolase domain, which could contribute to the apparent favorability of the open form of *Pfu* POP.

The conformation of the histidine (H592) loop also plays a pivotal role in POP peptidase catalysis. In the other POP crystal structures reported to date, the histidine loop was modeled only in closed structures[18,29] and omitted from open structures due to poor electron density.[18] This trend holds for the structures of most other POP family enzymes, although the histidine loop of oligopeptidase B from *T. Brucei*[41] was resolved in the structures of both the open and closed conformations. While the closed form of *Pfu* POP was not observed in any of the crystal structures, overlaying the structures obtained with those for other closed POPs shows that H592 would have to travel 7-12 Å to achieve an orientation similar to the closed structures (Fig. S1). The structure of *Pfu* POP thus offers a rare view[41] of the position of the catalytic histidine in a POP prior to formation of a competent catalytic triad.

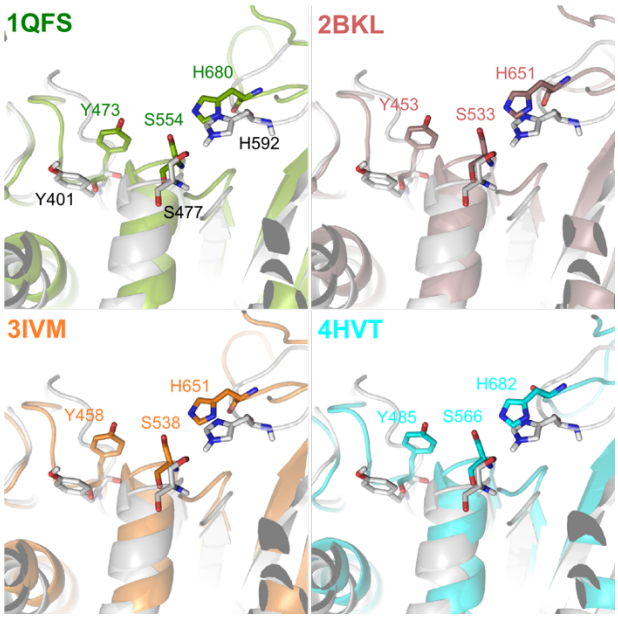

Figure 9. Overlay of closed crystal structures with a simulation snapshot with a closed H592 loop. Important residues are shown with sticks.

Two approaches were used to examine the conformational dynamics involved in forming active site configurations analogous to those observed in closed POP structures. In the first,

triplicate MD simulations were performed on the WT, IHBT, and S477C systems as discussed in the His Loop Conformational Dynamics section above. During these simulations, a number of open-to-closed transitions were observed (Figs. 3 and S3), but only two of these events led to the formation of active site configurations analogous to those observed in the structures of closed, inhibitor-bound POPs (Figs. 9 and S1). The slight structural deviation from the closed structures of these covalently inhibited structures is in part due to the fact that their formation requires bond breakage and formation, which is outside of the MD simulation regime. Nonetheless, competent-like active site configurations were only seen in snapshots of closed POP structures with small open angles, suggesting that domain closing shifts the conformational distribution of the H592 loop to an orientation suitable for catalysis. In the second approach, model structures of *Pfu* POP in its closed form were generated using the H592 loop φ/ψ angle information from the crystal structures of POP in closed state. The MD simulations of these model structures without any inhibitors bound show that the H592 loop rapidly moves away from the closed conformation to adopt an orientation similar to that seen in the open conformation. The observed dynamics of these model systems after the H592 loop moves away from the active site are similar to those observed in the WT simulations according to the φ/ψ angle distributions for both systems (Fig. S6), providing a clear picture of the large-scale inter-domain opening/closing that occurs in *Pfu* POP.

Effects of Pfu POP Dynamics on Peptidase and Dirhodium Catalysis

Crystal structures and MD simulations for different POPs suggest that these enzymes undergo large conformational changes, but the extent to which these changes influence POP catalysis (e.g. substrate binding, formation of a competent active site, etc.) remains an open question. The idea that conformational dynamics could impact native and non-native activities

within the same scaffold is particularly interesting and to our knowledge, without precedent. Understanding how this could occur in *Pfu* POP requires analysis of previous studies on the influence and timing of conformational changes on amide hydrolysis (Fig. 1A)--that is, the extent to which conformational dynamics impart utility to *Pfu* POP as a supramolecular host for peptidase catalysis (Fig. 1B).

Several findings from previous studies provide key insights into the elementary steps of *Pfu* POP peptidase catalysis (Fig. 1A): 1) The rate-limiting step in peptide hydrolysis differs at lower temperatures (25-40 °C) versus higher temperatures (40-95 °C), according to observed non-linearity in a plot of $\ln(k_{cat}/T)$ vs $(1/T)$.[24] 2) A 70-80-fold increase in k_1 was observed from 25-55 °C, suggesting a large kinetic barrier for formation of an enzyme-substrate complex at lower temperatures.[24] 3) A plot of k_{cat} versus temperature for hydrolysis of Z-Gly-Pro-pNA was linear between the ranges reported (60-90 °C), indicating that a single elementary step was rate limiting over this range.[24] 4) Peptide hydrolysis at high temperatures (56 and 85 °C) exhibited a solvent isotope effect, implicating a rate-limiting chemical step in this range.[22,24] 5) Lack of a leaving group effect at high temperatures (55-75 °C) was observed, which, in conjunction with points 3 and 4, suggests that k_3 (chemical hydrolysis of the acyl intermediate) is rate-limiting at high temperatures.[24] 6) A sigmoidal increase in k_{cat}/K_M as a function of [NaX] (X = Cl, Br, F) and no change in k_{cat} was observed for *Pfu* POP up to ~2M NaX, suggesting that halide binding activates *Pfu* POP hydrolase activity (~2-3 fold) via a decrease in K_M.[22] Thus, while a chemical step is rate limiting at high temperatures, substrate binding, and any conformational changes associated with it, are rate limiting below 40 °C, the temperature range at which all POP ArM catalysis was also examined.

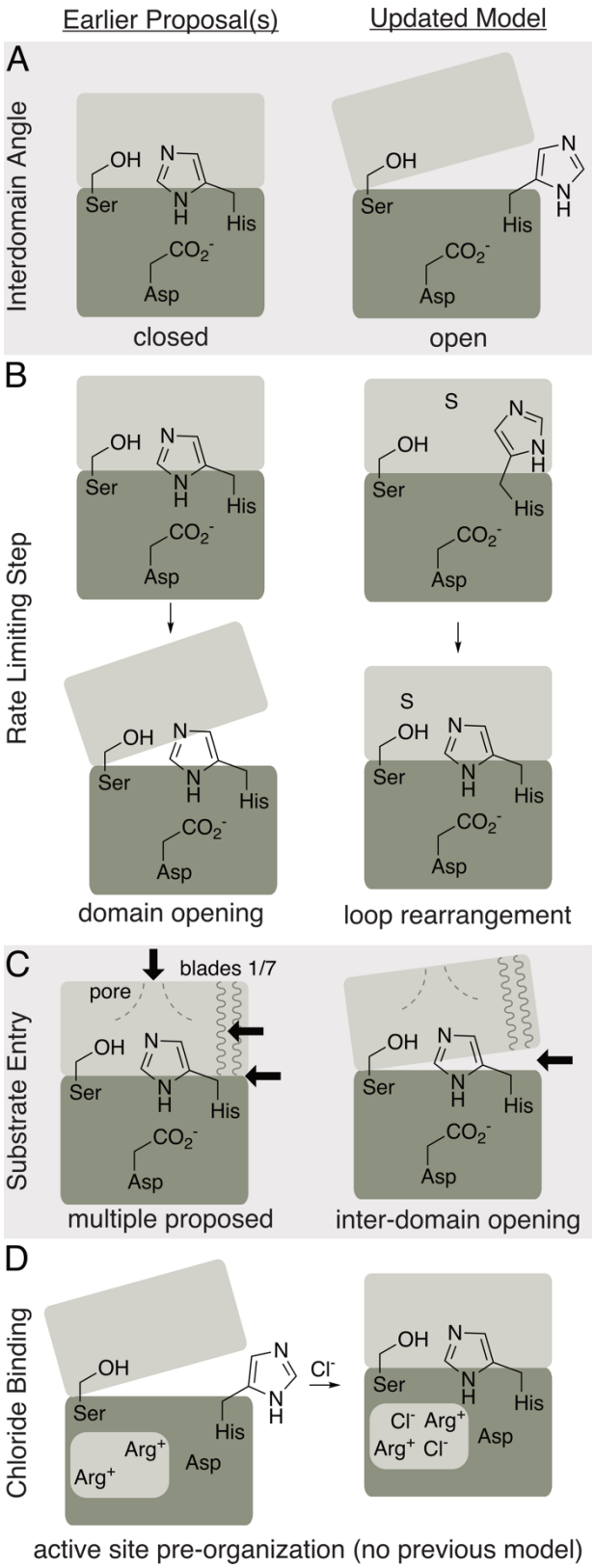
Earlier Proposal(s)
Updated Model
A
Interdomain Angle
OH
N
Ser
N
H
His
CO2-
Asp
closed
open
B
Rate Limiting Step
S
domain opening
loop rearrangement
C
Substrate Entry
blades 1/7
pore
multiple proposed
inter-domain opening
D
Chloride Binding
Cl-
Arg+
active site pre-organization (no previous model)

Figure 10. Conceptual summary of updated model for effects of POP structure and dynamics on peptidase activity (S = substrate).

It was previously hypothesized that *Pfu* POP has a stable, closed structure at low temperature, and that the high barrier to substrate binding arises from a high-barrier conformational change between open- and closed-states (Fig. 10A).[24] The crystal structure of *Pfu* POP shows that an open form of the enzyme is readily accessible, that the enzyme possesses several Cl^- binding sites, and that a dipeptide can bind within the open active site (Fig. 3C). MD simulations allowed for observation of the open/closed transition, and the same dynamics are observed in simulations of WT, a model of closed WT, and IHBT (Figs. 4 and S3). These findings are consistent with low barrier equilibria involving substrate, Cl^-, and the open and closed forms of *Pfu* POP, in which formation of substrate/Cl^--bound open (observed by X-ray crystallography) and closed (observed in MD simulations) forms of the enzyme can be generated. Rate limiting formation of the ES complex below 40 °C could result not from domain opening but from conversion of a closed, substrate-bound form of the enzyme to one in which the catalytic triad is fully formed and poised for attack of substrate (Fig. 10B)[42]. While such a state is beyond the ability of classical MD simulations to model, conformations resembling those of closed structures with covalent inhibitors bound were observed in simulations of dynamics at 358.15 K but not at 300 K, consistent with the proposal of rate limiting ES formation at low temperatures.

The *Pfu* POP crystal structures and MD simulations also provide insight into substrate entry into the enzyme (Fig. 10C). Specifically, the 219 ± 74 $Å^2$ opening at the narrowest region between blade 3 and the peptidase domain can accommodate a peptide chain (WT Replica 2 in Table S4). The loop lining the edge of this opening within the peptidase domain (D505-Y522) exhibits large

fluctuations in the MD simulations (Fig. S5). Fluctuation of this loop could act to facilitate substrate entry into the POP enzyme interior. While a small opening (71-112 $Å^2$ at 358.15K, Table S5) is observed at the pore lined by the seven β-blades in the propeller domain, this would only be sufficient for entry of peptides containing small residues, contrary to the reported substrate scope. No opening is observed between the propeller blades 1 and 7 (b1/b7), so substrate entry from this region is also unlikely. Entry through the much wider opening between the propeller and peptidase domains therefore seems most reasonable for *Pfu* POP.

Finally, the observation of three distinct halide binding sites between WT and S477C *Pfu* POP structures (Fig. 2) is qualitatively consistent with the aforementioned activation by halide.[22] Halide binding to R476 and R600 in particular provides a structural rationale for the impact of halide concentration on POP activity (Fig. 10D). Forming the catalytic triad in *Pfu* POP requires a large conformational change (Fig. 6) to position H592 between D560 and S477, which move independent of one another. The presence of R476 immediately before to S477 and of R600 in the helix adjacent to the loop containing H592 provides a means to orient the secondary structures containing these residues upon halide binding. By securing the orientation of these secondary structures, the salt bridge helps position S477 and H592 to form the catalytic triad. This is consistent with the results of MD simulations showing shorter H592-S477 distances and smaller inter-domain angles (Figs. 3, 6, and S3) when chloride is bound versus when it is not. This mechanistic proposal, which involves active site pre-organization to facilitate substrate binding, would be reflected in K_M, consistent with the kinetic analyses noted above and reminiscent of previous work on angiotensin converting enzyme[43].

The model of *Pfu* POP peptidase catalysis outlined above resolves previous ambiguities in the literature by considering how the conformational dynamics of the enzyme are perturbed by

temperature and halide binding. Conformational dynamics have been linked to the activity of a number of enzymes,[5] but no examples have been reported in which the conformational dynamics of an enzyme impacts its activity toward non-native catalysis upon installation of synthetic catalytic components.[9] Given the importance of conformational dynamics on native enzyme catalysis, identifying and engineering proteins whose native dynamics can be used to influence non-native catalysis would greatly improve our ability to design and create artificial enzymes, including ArMs, with properties analogous to natural enzymes.

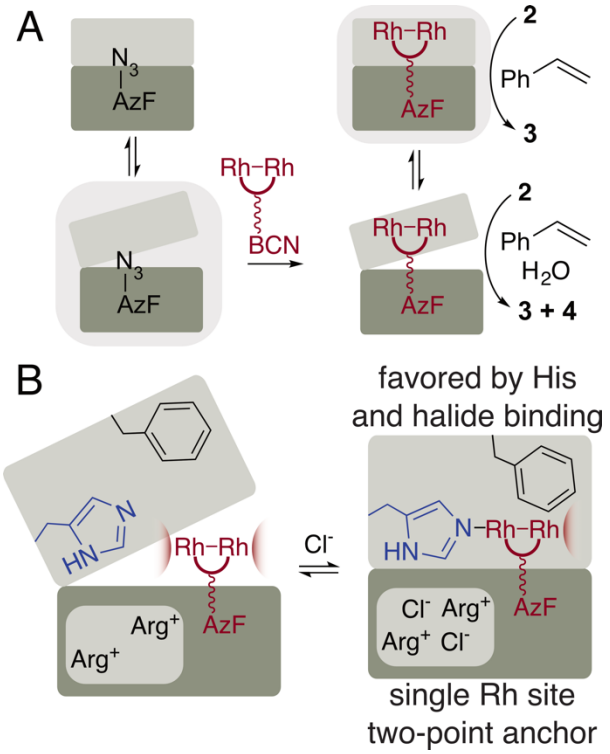

Figure 11. A model for the effects of conformational dynamics on POP bioconjugation and catalysis. A) bioconjugation can readily occur with the open conformation, while exclusion of water from the closed conformation allows for selective reaction of carbene precursor 2 with styrene to give cyclopropane 3. B) The closed form of the ArM is favored by both halide and Rh binding within the active site.

Efficient ArM formation via covalent cofactor linkage requires a sterically accessible linkage site within the host, but complete encapsulation of the cofactor provides the greatest potential for a host to influence cofactor/substrate reactivity (Fig. 11A). These conflicting requirements can be addressed if the host can sample an open conformation for bioconjugation and a closed conformation for catalysis. The open structure of *Pfu* POP shows that the enzyme possesses a cavity large enough (~10^4 Å^3)[44] to host a wide range of bulky metal complexes, including **1** (~10^3 Å^3)45. MD simulations of the apo structures show that *Pfu* POP can readily access open conformations with large θ values that would allow facile cofactor access to the active site. In the open conformation of *Pfu* POP, however, both rhodium centers in **1** could catalyze carbene insertion reactions into olefin substrates or water with little influence from the protein host (Fig. 11B). As previously noted, MD simulations of the dirhodium ArM constructed from WT *Pfu* POP show that similar θ values, consistent with a structure open to bulk water, are observed for both systems. Consistent with these simulations, carbene insertion into water to give compound **4** is primarily observed for reactions of the WT POP ArM, and low enantioselectivity is observed for the olefin insertion product.

A key finding of our previous rational design[27] and directed evolution[47] efforts is the importance of metal-binding residues in the propeller domain of *Pfu* POP for ArM activity. Histidine binding to one of rhodium atoms in **1** creates a single site catalyst, which eliminates problems arising from differential activity of two rhodium sites in different regions of the host (Fig. 11B). More importantly, because **1** is linked to the peptidase domain, metal binding to the propeller domain cross-links the two domains via a coordinate bond within the POP active site. This cross-link would favor the closed form of the host, which would create a more ordered, hydrophobic environment that could increase the local concentration of olefin relative to water and

better orientation cofactor and substrate for selective carbene insertion. Indeed, greatly increased activity, selectivity, and specificity is observed for dirhodium ArMs containing metal binding residues at site 326 or 328. POP variant ZA4 (L328), for example, provided **3** in 25% yield and 38% ee with a **3**/**4** ratio of 0.6, while the corresponding L328H variant gave **3** in 61% yield and 85% ee with a **3**/**4** ratio of 1.6.

Based on the enhanced peptidase activity of *Pfu* POP in the presence of high concentrations of [NaX], the corresponding dirhodium ArM was evaluated under similarly high salt concentrations. Significant increases in ArM enantioselectivity, specificity, and yield were observed up to 1.75 M NaX, and NaBr provided slightly higher performance than NaCl.[27] For example, while POP variant ZA4 provided 3 in 19% yield and 11% ee in the absence of salt, the 25% yield and 38% ee noted above was observed in the presence of 1.75 M NaBr. These improvements are consistent with both the active site pre-organization and smaller inter-domain angles observed in simulations in the apo system. That is, if these same features manifest in the ArM, the more organized active site could provide a more rigid anchor site for synthetic cofactors covalently linked within the POP active site. The smaller inter-domain angle also implies that the host is spending more time in a closed conformation, which would provide a more favorable environment for the desired olefin insertion chemistry.

Conclusion

The crystal structures of WT *Pfu* POP and its S477C mutant reveal that both enzymes adopt an open structure with an intermediate inter-domain angle relative to other POP enzymes reported to date. MD simulations show that the conformation observed in the crystal structures is indeed intermediate between much larger extremes that can be sampled by the enzyme. The observed

dynamics show that substrates can access the *Pfu* POP active site via the inter-domain opening, which lacks a "latch" present in other structures, while other previously proposed openings are much smaller and show no significant opening during MD simulations.

The *Pfu* POP crystal structures provide rare glimpses of the loop containing the catalytic histidine (H592) in an open POP structure. MD simulations show that H592 can access a conformations analogous to those observed in previously reported structures of POP-inhibitor complexes. The structure of WT *Pfu* POP also contains a bound prolylproline ligand, showing that substrate-like compounds remain bound in the open for of the enzyme. Up to three bound chloride ions were observed in the different *Pfu* POP structures, and those bound to R476 and R600 provide a rationale for previously observed halide activation involving active site preorganization. MD simulations show that chloride binding at these sites alters both histidine loop conformation and the inter-domain angle to favor a more closed form of the enzyme.

Together, these results resolve a number of questions in the literature regarding native *Pfu* POP structure, inter-domain dynamics, substrate entry, and activation by halide ions. Specifically, our analysis reveals the importance of facile inter-domain opening, rate-limiting EA formation or hydrolysis, and halide-induced active site preorganization during the *Pfu* POP catalytic cycle. This study also sheds light on the efficacy of *Pfu* POP as a scaffold for artificial metalloenzyme formation. Inter-domain dynamics allow for incorporation of bulky cofactors within the POP active site. Both intramolecular domain cross-linking via rhodium coordination and increased halide concentration can favor the closed form of the enzyme required for selective catalysis. While conformational dynamics are widely understood to impact different aspects of enzyme catalysis, the finding that the conformational dynamics of *Pfu* POP can influence two completely different catalytic activities attests to the remarkable utility of enzymes as supramolecular hosts

for catalysis. *Pfu* POP thus constitutes a powerful platform for the development of supramolecular catalysts, including artificial metalloenzymes, with catalytic properties otherwise found only in natural enzymes.

References

1. Brown, C. J.; Toste, F. D.; Bergman, R. G.; Raymond, K. N. Chem. Rev. 2015, 115, 3012–3035.

2. Deraedt, C.; Astruc, D. Coord. Chem. Rev. 2016, 324, 106–122.

3. Breslow, R. Acc. Chem. Res. 1995, 28, 146–153.

4. Bar-Even, A.; Noor, E.; Savir, Y.; Liebermeister, W.; Davidi, D.; Tawfik, D. S.; Milo, R. Biochemistry 2011, 50, 4402–4410.

5. Bhabha, G.; Biel, J. T.; Fraser, J. S. Acc. Chem. Res. 2015, 48, 423–430.

6. Renata, H.; Wang, Z. J.; Arnold, F. H. Angew. Chem. Int. Ed. 2015, 54, 3351–3367.

7. Coelho, P. S.; Arnold, F. H.; Lewis, J. C. Synthetic Biology Approaches for Organic Synthesis; Molander, G. A.; Knochel, P., Eds.; Second Edition. Elsevier Ltd.: Oxford, 2014; Vol. Volume 9: Enabling Technologies for Organic Synthesis, pp. 390–420.

8. Qi, D.; Tann, C.; Haring, D.; Distefano, M. Chem. Rev. 2001, 101, 3081–3111.

9. Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Kohler, V.; Lewis, J. C.; Ward, T. R. Chem. Rev. 2017, acs.chemrev.7b00014–90.

10. Osuna, S.; Jiménez-Osés, G.; Noey, E. L.; Houk, K. N. Acc. Chem. Res. 2015, 48, 1080–1089.

11. Polgár, L. Cell Mol Life Sci 2002, 59, 349–362.

12. Szeltner, Z.; Polgár, L. Curr. Protein Pept. Sci. 2008, 9, 96–107.

13. Rea, D.; Fülöp, V. Cell Biochem. Biophys. 2006, 44, 349–365.

14. Walter, R.; Shlank, H.; Glass, J. D.; Schwartz, I. L.; Kerenyi, T. D. Science 1971, 173, 827–829.

15. Kaushik, S.; Sowdhamini, R. BMC Genomics 2014, 15, 985–13.

16. Fulop, V.; Böcskei, Z.; Polgár, L. Cell 1998, 94, 161–170.

17. Haffner, C. D.; Diaz, C. J.; Miller, A. B.; Reid, R. A.; Madauss, K. P.; Hassell, A.; Hanlon, M. H.; Porter, D. J. T.; Becherer, J. D.; Carter, L. H. Bioorg. Med. Chem. Lett. 2008, 18, 4360–4363.

18. Shan, L.; Mathews, I. I.; Khosla, C. Proc. Natl. Acad. Sci. U.S.A. 2005, 102, 3599–3604.

19. Li, M.; Chen, C.; Davies, D. R.; Chiu, T. K. J. Biol. Chem. 2010, 285, 21487–21495.

20. Harwood, V. J.; Denson, J. D.; Robinson-Bidle, K. A.; Schreier, H. J. J. Bacteriol. 1997, 179, 3613–3618.

21. Harwood, V. J.; Schreier, H. J. Meth. Enzymol. 2001, 330, 445–454.

22. Harris, M. N.; Madura, J. D.; Ming, L.-J.; Harwood, V. J. J. Biol. Chem. 2001, 276, 19310–19317.

23. Juhász, T.; Szeltner, Z.; Polgár, L. FEBS Letters 2006, 580, 3493–3497.

24. Juhász, T.; Szeltner, Z.; Polgár, L. Proteins 2007, 69, 633–643.

25. Kaushik, S.; Etchebest, C.; Sowdhamini, R. Proteins 2014, 82, 1428–1443.

26. Lewis, J. C. ACS Catal. 2013, 3, 2954–2975.

27. Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. Nat. Commun. 2015, 6, 7789.

28. Szeltner, Z.; Juhász, T.; Szamosi, I.; Rea, D.; Fülöp, V.; Módos, K.; Juliano, L.; Polgár, L. BBA - Proteins and Proteomics 2013, 1834, 98–111.

29. Fulop, V.; Böcskei, Z.; Polgár, L. Cell 1998, 94, 161–170.

30. Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J. L.; Verschueren, K. H. G.; Goldman, A. Prot. Eng. Des. Sel. 1992, 5, 197–211.

31. Fulop, V.; Szeltner, Z.; Renner, V.; Polgár, L. J. Biol. Chem. 2001, 276, 1262–1266.

32. Vieille, C.; Zeikus, G. J. Microbiol Mol Biol R 2001, 65, 1–43.

33. Sterpone, F.; Melchionna, S. Chem. Soc. Rev. 2012, 41, 1665–1676.

34. Kichik, N.; Tarragó, T.; Claasen, B.; Gairí, M.; Millet, O.; Giralt, E. 2011, 12, 2737–2739.

35. TarragÃ³, T.; Claasen, B.; Kichik, N.; Rodriguez-Mias, R. A.; GairÃ, M.; Giralt, E. ChemBioChem 2009, 10, 2736–2739.

36. Kaszuba, K.; Róg, T.; Danne, R.; Canning, P.; Fülöp, V.; Juhász, T.; Szeltner, Z.; Pierre, J. F. S.; García-Horsman, A.; Männistö, P. T.; Karttunen, M.; Hokkanen, J.; Bunker, A. Biochimie 2012, 94, 1398–1411.

37. Kaushik, S.; Sowdhamini, R. PLoS ONE 2011, 6, e26251.

38. St-Pierre, J.-F.; Karttunen, M.; Mousseau, N.; Róg, T.; Bunker, A. J. Chem. Theory Comput. 2011, 7, 1583–1594.

39. Kotev, M.; Lecina, D.; Tarragó, T.; Giralt, E.; Guallar, V. Biophysj 2015, 108, 116–125.

40. López, A.; Herranz-Trillo, F.; Kotev, M.; Gairí, M.; Guallar, V.; Bernadó, P.; Millet, O.; Tarragó, T.; Giralt, E. ChemBioChem 2016, 17, 913–917.

41. Canning, P.; Rea, D.; Morty, R. E.; Fülöp, V. PLoS ONE 2013, 8, e79349.

42. Smith, A. J. T.; Mueller, R.; Toscano, M. D.; Kast, P.; Hellinga, H. W.; Hilvert, D.; Houk, K. N. J. Am. Chem. Soc. 2008, 130, 15361–15373.

43. Liu, X.; Fernandez, M.; Wouters, M. A.; Heyberger, S.; Husain, A. J. Biol. Chem. 2001, 276, 33518–33525.

44. Kozlikova, B.; Sebestova, E.; Sustr, V.; Brezovsky, J.; Strnad, O.; Daniel, L.; Bednar, D.; Pavelka, A.; Manak, M.; Bezdeka, M.; Benes, P.; Kotry, M.; Gora, A.; Damborsky, J.; Sochor, J. Bioinformatics 2014, 30, 2684–2685.

45. Voss, N. R.; Gerstein, M. Nucleic Acids Research 2010, 38, W555–W562.

Acknowledgement

This work was supported by, or in part by, the U.S. Army Research Laboratory and the U. S. Army Research Office under contract/grant numbers W911NF-14-1-0334 and 66796-LS-RIP (to J.C.L.) and W911NF-18-1-0200 (to J.C.L. and B.R.), and the NSF under CAREER Award CHE-1351991 (to J.C.L.), and The David and Lucile Packard Foundation (to J.C.L.). This work is based on research conducted at the APS on the NE-CAT beamlines, which are supported by a grant from the NIGMS (P41 GM103403) from the NIH. Use of the APS, an Office of Science User Facility operated for the U.S. DOE Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. This research was supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute, and in part by the Indiana METACyt Initiative. The Indiana METACyt Initiative at IU was also supported in part by Lilly Endowment, Inc.

Author Information

Corresponding Author

* jcl3@iu.edu

*roux@uchicago.edu

*sukumar@anl.gov

Present Addresses

K.E.G.: University of California - Los Angeles, Boyer Hall 610, 611 Charles E. Young Drive East, Los Angeles, CA 90095

R.B.: National Institute on Alcohol Abuse and Alcoholism (NIAAA), 5625 Fishers Ln room 3N15, Rockville, MD 20852

Author Contributions

‡These authors contributed equally.