Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble

Andrew R. Buller[†‡*], Paul van Roye[‡], Jackson K. B. Cahn[§], Remkes A. Scheele[‡], Michael Herger[‡], Frances H. Arnold[‡]

[†]Department of Chemistry

University of Wisconsin-Madison

Madison, WI 53706

[‡]Division of Chemistry and Chemical Engineering

California Institute of Technology

Pasadena, CA 91125

§ Institute of Microbiology

Eidgenössische Technische Hochschule (ETH) Zürich

Zurich 8093, Switzerland

Supporting Information

SI Table of Contents

Experimental Procedures:

General

Chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, VWR, Chem-Impex International, Alfa Aesar) and used without further purification. Multitron shakers (Infors) were used for cell growth. UV-vis spectra were collected on a UV1800 Shimadzu spectrophotometer (Shimadzu). Liquid chromatography/mass spectrometry (LCMS) was performed with an Agilent 1290 UPLC in line with a 6140 MS Single Quad MS detector. Samples were analyzed on a 2.1 × 50 mm C-18 silica column using acetonitrile as the strong solvent and 0.1% (v/v) acetic acid/water as the weak solvent.

Cloning, heterologous expression, and protein purification

The gene encoding *Pf*TrpB (UNIPROT ID Q8U093) was previously codon-optimized for *Escherichia coli* and cloned into pET22(b)+ with a C-terminal 6-His-tag.[1] Expression and purification protocols for the variants reported here were similar to those reported previously.[1] Briefly, a single colony of *E.coli* BL21 *E. cloni* Express cells (Lucigen) harboring the *Pf*TrpB

plasmid was used to inoculate a 5-mL culture of Terrific Broth with 100 µg/mL ampicillin (TB_{amp}) and incubated overnight at 37 °C and 250 rpm. This culture was used to inoculate a 500-mL TB_{amp} expression culture, which was incubated at 250 rpm and 37 °C for ~ 3 h or until an OD_{600} of 0.6-0.8 was reached. Cultures were chilled on ice for 20-30 min and expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells continued to grow shaking at 250 rpm and 20 °C for another 20 h. Cells were harvested by centrifugation at 4 °C and 5,000*g* for 10 min; the pellets were frozen at -20 °C until further use.

Frozen cell pellets were thawed at room temperature and resuspended in 50 mM phosphate buffer, pH 8.0, with 20 mM imidazole and 100 mM NaCl (buffer A), with 200 µM PLP, 1 mg/mL hen egg white lysozyme, and 0.02 mg/mL DNAse. After vortexing, cells were lysed with BugBuster (Novagen) at one half of the manufacturer's recommended concentration. Lysates were aliquoted into 1.6-mL tubes and centrifuged at 20,000g and 4 °C for 10 min and then incubated at 75 °C for 10 min, centrifuged again as described above. The cleared, heat-treated lysate was applied to a 1-mL Histrap HP column (GE Healthcare). The purification was performed with an AKTA purifier FPLC system (GE Healthcare). *Pf*TrpB (and variants thereof) eluted during a linear gradient from buffer A to buffer B (50 mM phosphate buffer with 500 mM imidazole and 100 mM NaCl, pH 8.0) at ≈ 140 mM imidazole. Purified enzyme was dialyzed into 50 mM phosphate buffer, pH 8.0, frozen in liquid N_2, and stored at -80 °C until further use.

A similar procedure was used to prepare the *Pf*TrpS complex with the following exceptions. *Pf*TrpA was expressed without the 6-his tag and initially purified as a heat-treated lysate without the addition of PLP. The amount of TrpA and TrpB present in the lysate was estimated by SDS-PAGE, and the heteromeric TrpS complex was formed by mixing *C*-His-TrpB with an

approximate two-fold molar excess of native TrpA. The resultant complex was purified via pull-down by Ni-affinity chromatography on an AKTA purifier FPLC system (GE Healthcare). TrpS eluted during a linear gradient from buffer A (50 mM phosphate buffer with 20 mM imidazole and 100 mM NaCl, pH 8.0) to buffer B (same as buffer A, but with 500 mM imidazole) at ≈ 140 mM imidazole.

Purified protein concentrations were determined via the Bradford assay (Bio-Rad).

Kinetics

Tryptophan formation was monitored at 75 °C over 1 min at 290 nm using $\Delta\varepsilon_{290}$ = 1.89 mM^{-1} cm^{-1}.[2] The assay buffer contained 200 mM potassium phosphate pH 8.0. Previous studies included additional equivalents of PLP cofactor, however this was found to have no measurable effect on activity and was removed from the assays reported here. Michaelis constants (K_Ms) for indole were determined using a concentration range of 10–300 μM indole with the concentration of L-serine fixed at 20 mM. The K_Ms for L-serine were determined using a concentration range of 0.3–20 mM L-serine with the concentration of indole fixed at 250 μM. Initial velocities were measured in duplicate or triplicate for at least six different substrate concentrations for each enzyme. The data were fit using an in-house non-linear least squares algorithm to the Michaelis-Menten equation implemented in MatLab. The k_{cat}s for the overall reaction (Table 1) were determined from the titration of indole (above).

Kinetic isotope effects (KIE) were measured using d_3-L-Ser as described in the main text. Direct competition experiments for *Pf*TrpB were conducted in triplicate using 10 mM (d_3)-Ser, 1 mM indole, and 10 μM *Pf*TrpB. Reactions were conducted in 0.2 M K phosphate pH 8.0 and incubated at 75 °C for 2 hours and quenched with an equal volume of 100% acetonitrile. The

ratio of L-Trp to d_2-L-Trp was measured by LC-MS and integration of the extracted ion chromatogram at the appropriate m/z, indicating a KIE of 2.61 ± 0.08. KIEs were also determined by comparing initial velocities of reactions conducted with 10 mM (d_3)-Ser, 0.25 mM indole, and 0.5 μM of each *Pf*TrpB variant or *Pf*TrpS. Values are reported in Table 1.

UV-Vis spectroscopy

Spectra were collected between 550 and 250 nm on a UV1800 Shimadzu spectrophotometer (Shimadzu) using 5–60 μM of enzyme in 200 mM potassium phosphate pH 8.0 in a quartz cuvette. Samples were incubated at 75 °C for > 3 min to ensure a stable temperature was reached prior to addition of substrates. For steady-state spectra, enzyme and 1 mM indole were pre-incubated and the reaction initiated by addition of 20 mM L-serine. The initial spectrum was recorded in < 15 s, prior to the consumption of indole. A second spectrum was collected after 1 min, whereupon consumption of indole shifted the steady state to reflect only the Stage I intermediates of the catalytic cycle. These latter spectra matched those previously reported in the absence of indole or tryptophan.

Tryptophan binding affinities were determined by titration in the spectrophotometer and monitoring the change in absorbance at 475 nm, corresponding to population of the E(Q_3) species, using 20 μM enzyme and fit to a single site binding isotherm. To increase the signal for *Pf*TrpB, 60 μM was used. For *Pf*TrpB^{2B9}, experiments were performed at both 10 and 20 μM, which ensured that the fit value of the K_D for Trp (12 μM) was not an artifact of the lower bound of the assay, which equals half the protein concentration.

Protein Crystallography

*Pf*TrpB^{WT} crystallization conditions were identified previously,[1] and crystals routinely grew in sitting drops against a 1-mL reservoir of 15-20% PEG3350 and 0.1 M Na HEPES pH 7.85 with mother liquor comprised of 1.5 µL of 8.0 mg/mL *Pf*TrpB^{WT} and 1.5 µL of well solution. Crystals of *Pf*TrpB^{4D11} were inconsistently grown in similar conditions. To produce crystals for *Pf*TrpB^{2B9} (and for *Pf*TrpB^{4D11}), crystallization was induced by cross-seeding with pre-formed seeds of *Pf*TrpB^{WT} crystals prepared using the Seed Beed^{TM} kit (Hampton Research). Well solution contained 12-16% PEG3350 and 0.1 M Na HEPES pH 7.85. The mother liquor consisted of 1.5 µL of 10 mg/mL *Pf*TrpB variant and 1.5 µL of seed solution. Crystals grew over the course of three days and were stable in the dark over several months. Ser-bound crystals were prepared by soaking pre-formed crystals with 250 mM L-serine dissolved in well solution for 5-10 min. Trp-bound crystals were prepared by addition of solid Trp (~ 1-2 mg) directly onto drops containing pre-formed crystals and allowed to equilibrate overnight. Crystals were cryo-protected through oil immersion in Fomblin Y (Sigma) and flash frozen in liquid N_2 until diffraction. Data were collected remotely at the Stanford Synchrotron Radiation Laboratories on beamline 12-2. Crystals routinely diffracted at or below 2.0 Å, and the data were integrated and scaled using XDS[3] and AIMLESS.[4] Structures were solved using molecular replacement with PHASER, as implemented in CCP4.[5,6] The search model comprised of a single monomer of *Pf*TrpB^{WT} (PDB ID: 5DVZ), *Pf*TrpB^{WT}-Ser (PDB ID: 5DW0), or *Pf*TrpB^{WT}-Trp (PDB ID: 5DW3) where each of the mutated residues was deleted, the ligands and waters removed, and the whole model subject to ten cycles of geometric idealization in Refmac5. Model building was performed in Coot[7] beginning with data processed at 2.4 Å, followed by subsequent inclusion of increasingly higher resolution shells of data with relaxed geometric constraints. Refinement was

performed using REFMAC5[8] and resolution cutoffs were applied when incorporation of additional data caused R_{free} in the final bin to exceed 0.4; this criteria conformed to the crystallographic cutoff where $CC_{1/2} > 0.3$ along the strongest axis of diffraction.[4,9] The MolProbity server was used to identify rotamer flips, clashes, and calculate Ramachandran statistics.[10] After the protein, ligand, and solvent atoms were built, TLS operators were added to refinement, which resulted in substantial improvements in R_{free} for the final models. Crystallographic and refinement statistics are reported in Table S1.

Supplemental Figures

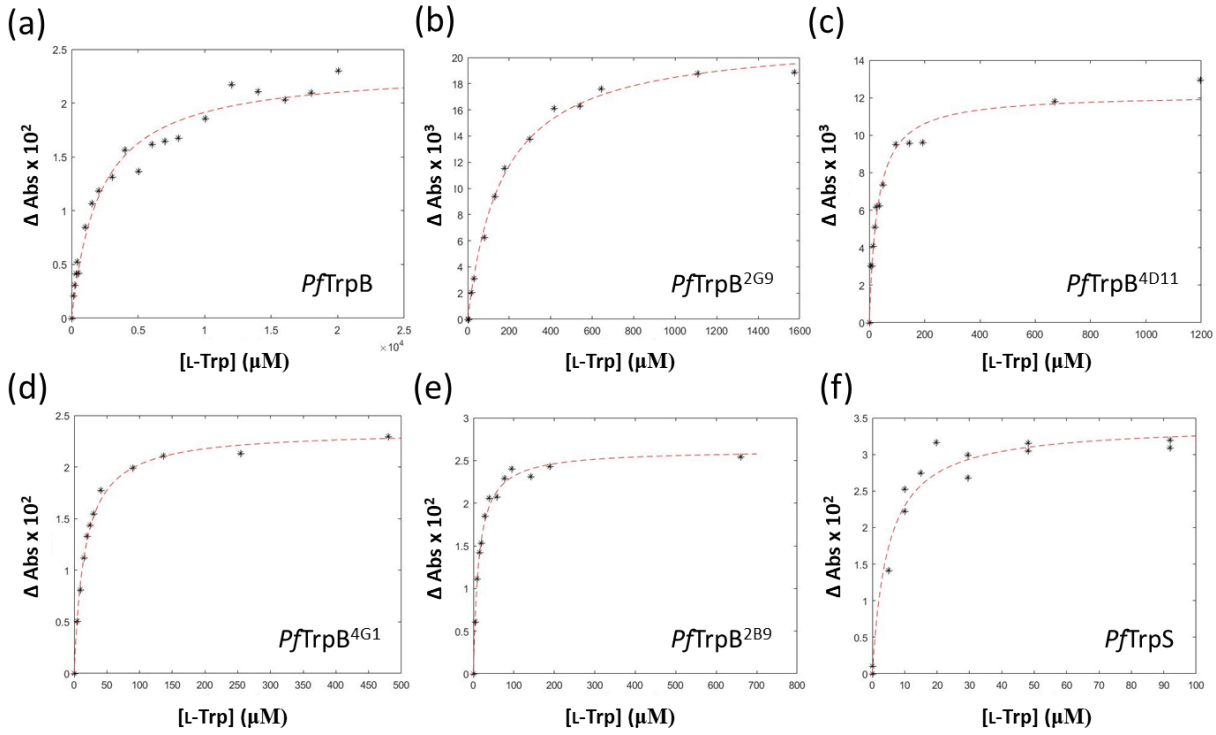

Figure S1. Trp binding isotherms for TrpB enzymes. Representative titration of Trp with each enzyme. The change in absorbance was recorded at 475 nm, corresponding to the $E(Q_3)$ intermediate. A non-linear fit to a single-site binding model for all titrations were performed in Matlab.

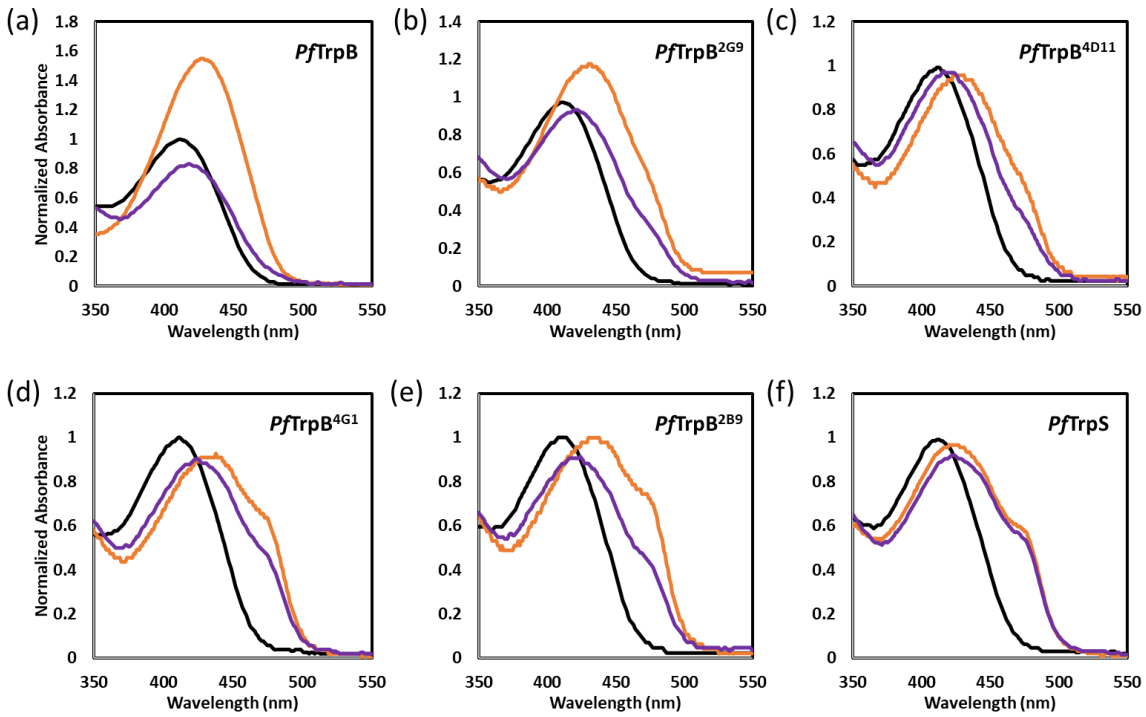

Figure S2. Comparison of steady-state and Trp-bound spectra of TrpB enzymes. Spectra were collected with 20-60 μM enzyme and are normalized with the absorbance of E(Ain) set at 1.0 for easy comparison (black). The steady-state spectra are shown in orange and are identical to those shown in Figure 2. The spectra of the Trp-bound protein at equilibrium are shown in purple and are identical to those shown in Figure 3. For *Pf*TrpB and *Pf*TrpB^{2G9}, the peak at steady state has a significant population of E(Aex$_1$) that convolutes direct comparison.

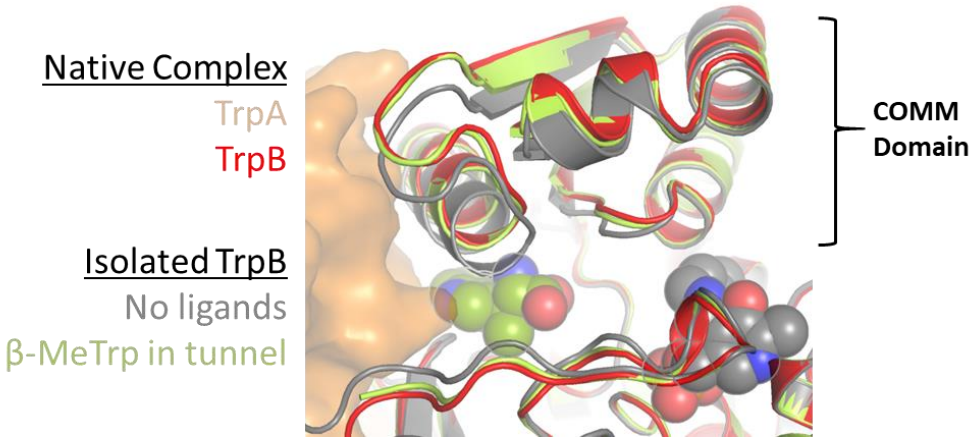

Figure S3. Comparison of native TrpBs in the E(Ain) state. In the native TrpS complex, TrpB (red) is in an extended open conformation. This is closely matched by the conformation of isolated TrpB with β-MeTrp (green spheres) in a tunnel connecting the two subunits, with an RMSD = 0.36 Å^2 between the backbone atoms of TrpB. In contrast, isolated TrpB without any ligands bound forms a non-extended 'open' conformation (gray).

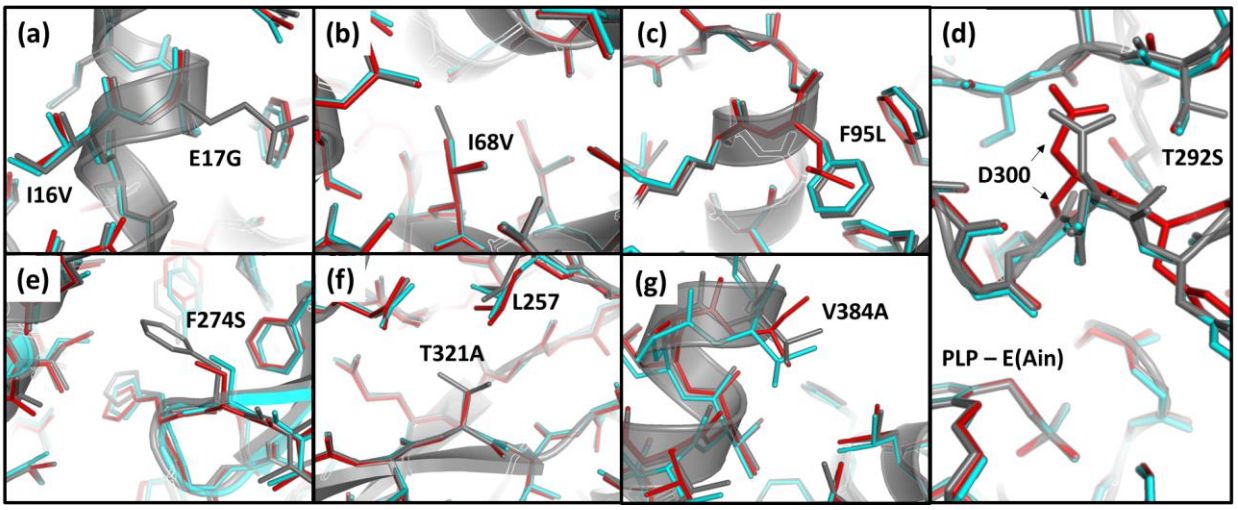

Figure S4. Structural comparison of *Pf*TrpB (gray), *Pf*TrpB^{4D11} (blue, 4D11), and *Pf*TrpB^{2B9} (red, 2B9) in the E(Ain) state. (a) The I16V mutation is present in 2B9, and E17G is in both 4D11 and 2B9. (b) I68V is present in both 4D11 and 2B9. (c) F95L is present in 2B9. (d) T292S is present in both 4D11 and 2B9. This residue hydrogen bonds with D300 in the 2B9 structure. See main text for further detail. (e) F274 is present in 4D11 and 2B9. (f) T321A is present in 4D11 and 2B9. This mutation is associated with a small readjustment of the L257 residue, but no further changes are observable within the region. (g) V384A is present in 2B9 and is located near the C-terminus of the protein where there is generally high conformational variability. Minor variation is observed between protomers of a given protein and a single, representative chain is shown here for clarity.

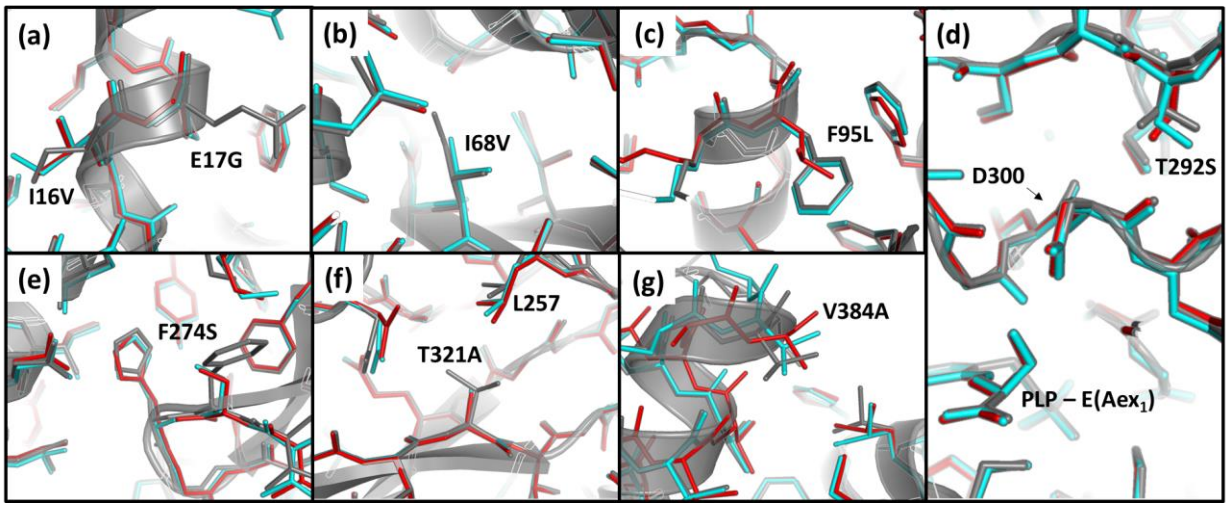

Figure S5. Structural comparison of *Pf*TrpB (gray), *Pf*TrpB^{4D11} (blue, 4D11), and *Pf*TrpB^{2B9} (red, 2B9) with Ser bound in the E(Aex$_1$) state. (a) The I16V mutation is present in 2B9, and E17G is in both 4D11 and 2B9. (b) I68V is present in both 4D11 and 2B9. (c) F95L is present in 2B9. (d) T292S is present in both 4D11 and 2B9. (e) F274 is present in 4D11 and 2B9. (f) T321A is present in 4D11 and 2B9. This mutation is associated with a small readjustment of the L257 residue, but no additional change is observable within the region. (g) V384A is present in 2B9 and is located near the C-terminus of the protein where there is generally high conformational variability. Minor variation is observed between protomers of a given protein and a single, representative chain is shown here for clarity.

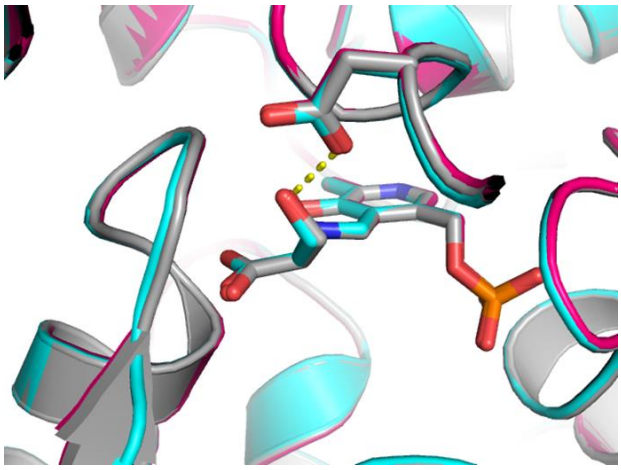

Figure S6. Structural comparison of *Pf*TrpB (gray), *Pf*TrpB^{4D11} (blue) and *Pf*TrpB^{2B9} (magenta) in the E(Aex$_1$) state shows no observable differences. PDB IDs: 5DW0, 6AMH, and 5VM5, respectively

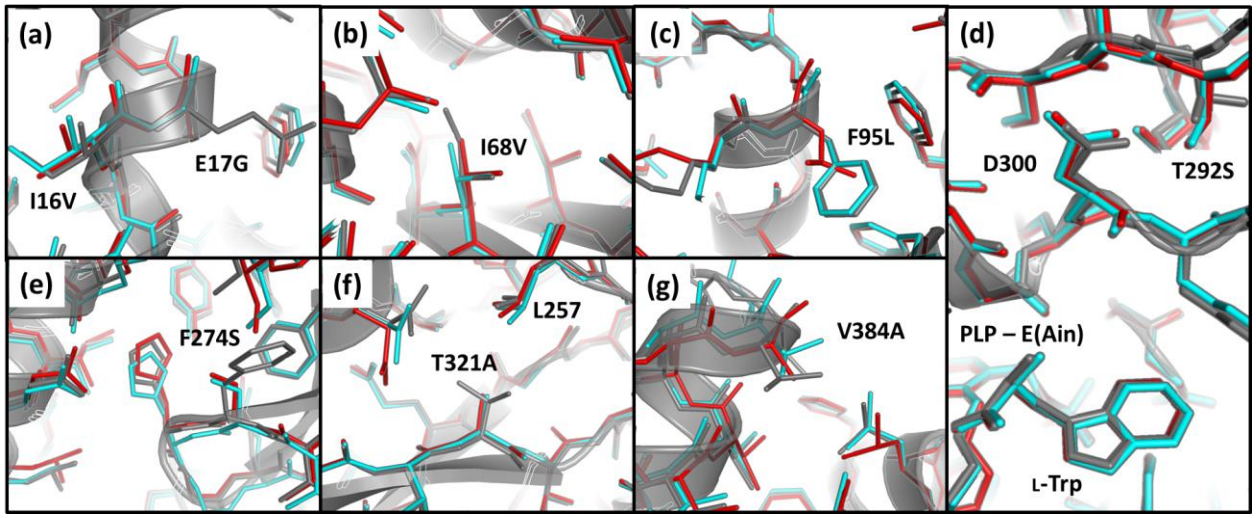

Figure S7. Structural comparison of *Pf*TrpB (gray), *Pf*TrpB^{4D11} (blue, 4D11), and *Pf*TrpB^{2B9} (red, 2B9) in the E(Ain) state with Trp non-covalently bound in the active site. (a) The I16V mutation is present in 2B9, and E17G is in both 4D11 and 2B9. (b) I68V is present in both 4D11 and 2B9. (c) F95L is present in 2B9. (d) T292S is present in both 4D11 and 2B9. This residue hydrogen bonds with D300 in each structure (e) F274 is present in 4D11 and 2B9. (f) T321A is present in 4D11 and 2B9. In this state, the N267 side chain showed multiple conformations, but no additional perturbations in the neighboring residues. (g) V384A is present in 2B9 and is located near the C-terminus of the protein where there is generally high conformational variability. Minor variation is observed between protomers of a given protein and a single, representative chain is shown here for clarity.

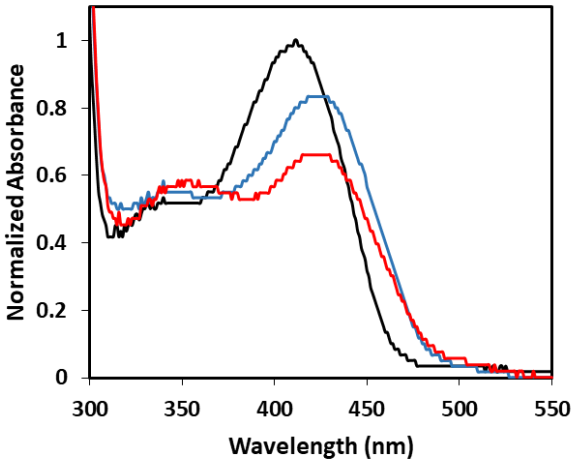

Figure S8. Comparison of the steady-state distribution of intermediates with L-Ser added to *Pf*TrpB^{4D11} (blue) and *Pf*TrpB^{2B9} (red) at 25 °C. Spectra are normalized with the E(Ain) absorbance (black) at 412 nm set to 1.0 for each enzyme.

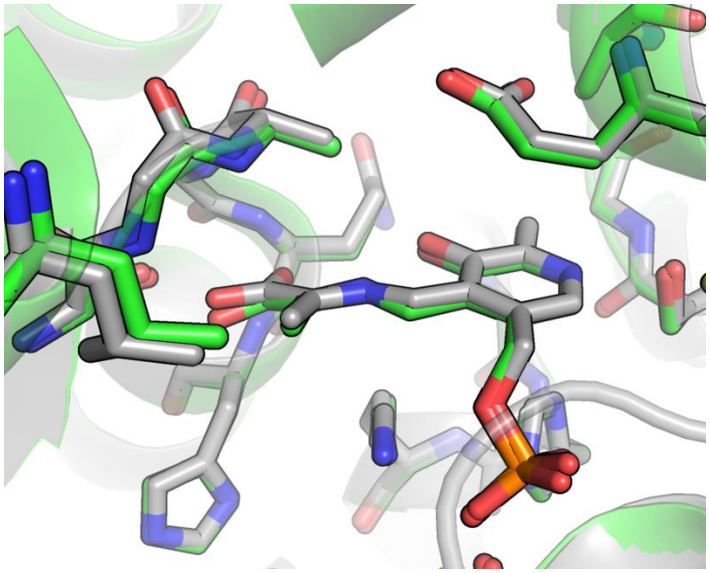

Figure S9. Comparison of the E(A-A) state of *Pf*TrpB^{2B9} (gray, PDB ID: 5VM5) and *St*TrpS (green, PDB ID: 4HN4) shows a nearly identical active site.

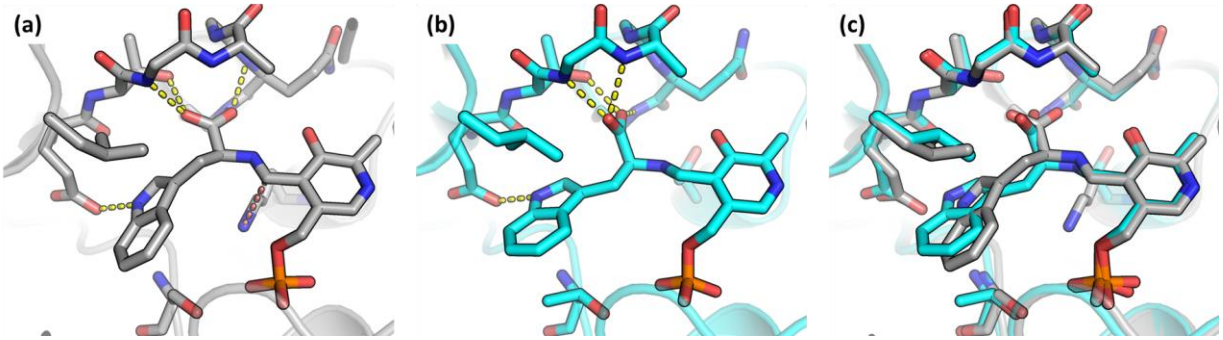

Figure S10. Structural comparison of native *Pf*TrpB^{2B9}-E(Aex$_2$) with *St*TrpS$^{\beta K87T}$-E(Aex$_2$). (a) The active site of *Pf*TrpB^{2B9} with Trp covalently bound as the E(Aex$_2$) intermediate. Hydrogen bonds to the carboxylate and indole N1 of Trp are shown in yellow. L161 is observed in a non-standard eclipsed rotamer. The next step in the catalytic cycle is transimination with the active site lysine, K82, with the trajectory indicated in orange dashes. (b) The active site of the catalytically inactive *St*TrpS$^{\beta K87T}$ with tryptophan covalently bound as the E(Aex$_2$) intermediate and hydrogen bonds indicated (PDB ID: 2TYS). (c) Overlay of the two structures. Introduction of the K87T mutation appears to cause two principle active site distortions. First, the E(Aex$_2$) shifts location and the carboxylate rotates ~ 40°, which alters its hydrogen bond interactions. Second, repositioning of the Trp allows the L161 residue (*Pf* numbering) to relax away from its eclipsed conformation into a standard staggered rotamer.

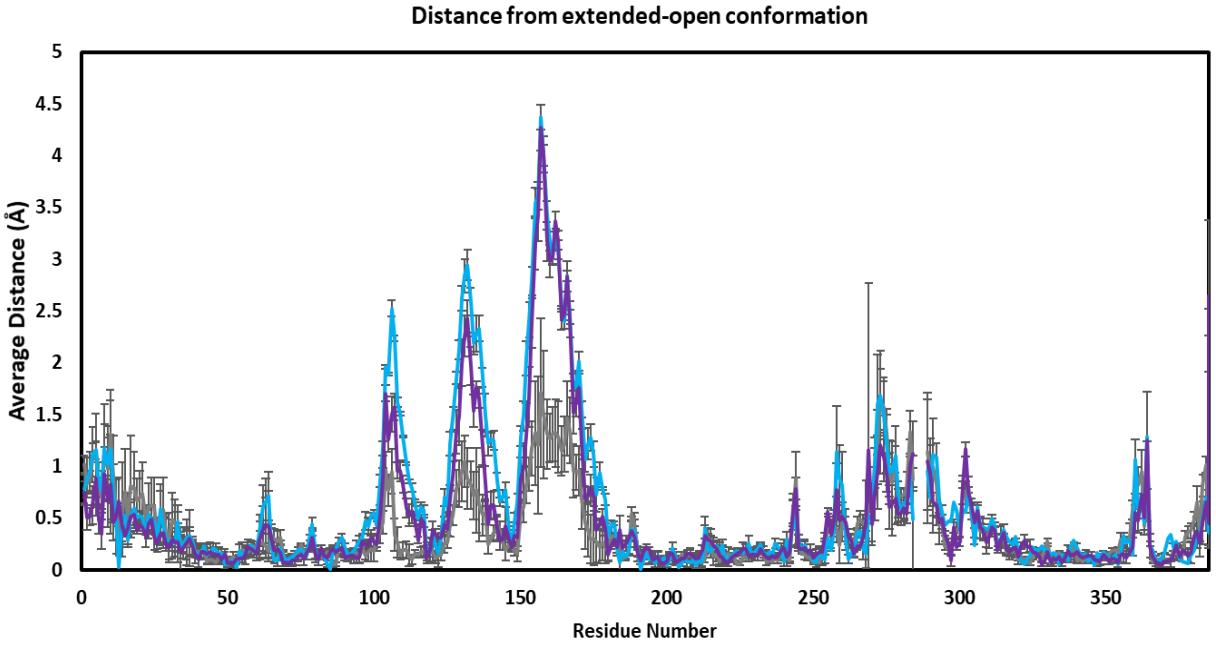

Figure S11. Comparison of substrate and product bound *Pf*TrpB structures. Average distance and standard deviation of each residue of the four protomers of a given structure from the extended-open conformation. *Pf*TrpB, no ligands, E(Ain), grey (PDB ID: 5DVZ). *Pf*TrpB, Ser-bound E(Aex$_1$), cyan (PDB ID: 5DW0). *Pf*TrpB, Trp-bound, E(Ain), purple (PDB ID: 5DW3).

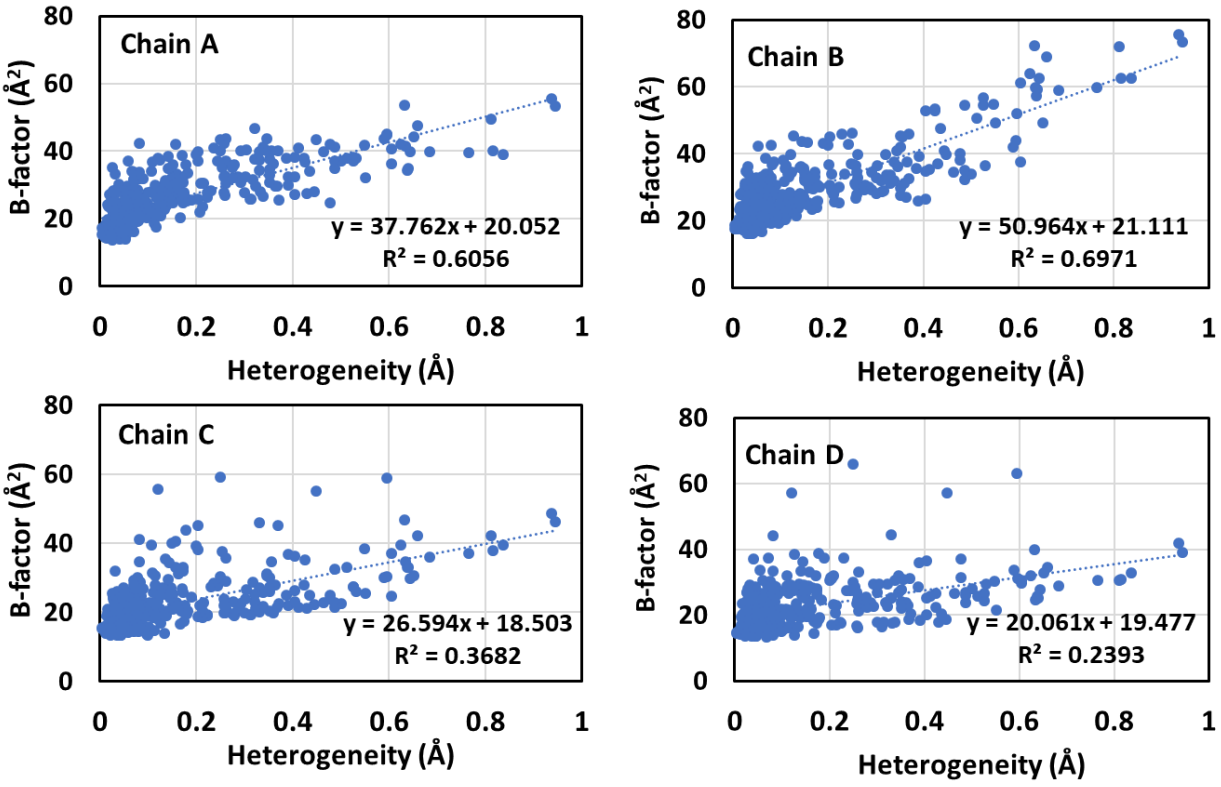

Figure S12. Comparison of B-factor with conformational heterogeneity in *Pf*TrpB[2B9]. The heterogeneity is calculated as the standard deviation of the distance of each reside from the extended-open conformation. As this parameter is an average across the protein, it is the same in each panel. Chains are therefore distinguished by the B-factor at Cα of each residue. The plots show that B-factors are weakly correlated with the observed heterogeneity.

Table S1. X-Ray crystallographic data collection and refinement statistics

PDB ID	6AMC	6AMH	6AMI	6AM7	5VM5	6AM9	6AM8
Protein	*Pf*TrpB^{4D11}	*Pf*TrpB^{4D11}	*Pf*TrpB^{4D11}	*Pf*TrpB^{2B9}	*Pf*TrpB^{2B9}	*Pf*TrpB^{2B9}	*Pf*TrpB^{2B9}
Ligand	N/A	L-Serine	L-Trp	N/A	L-Serine	L-Serine	L-Trp
State of PLP	E(Ain)	E(Aex$_1$)	E(Ain)	E(Ain)	E(Aex$_1$) x3, E(A-A) x1	E(Aex$_1$) x1, E(A-A) x3	E(Ain) x2, E(Aex$_2$) x2
Data Collection							
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions (Å)	a,b,c = 84.4, 110.7, 160.4	a,b,c = 84.8, 109.6, 160.9	a,b,c = 84.7, 109.9, 160.3	a,b,c = 84.0, 107.8, 159.8	a,b,c = 83.1, 108.7, 160.5	a,b,c = 87.5, 109.1, 160.5	a,b,c = 82.2, 106.0, 158.7
Cell angles	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°
Wavelength (Å)	0.9795	0.9795	0.9795	0.9795	0.9795	0.9795	0.9795
Beamline	SSRL 12.2	SSRL 12.2	SSRL 12.2	SSRL 12.2	SSRL 12.2	SSRL 12.2	SSRL 12.2
Resolution (Å)	40 – 1.93	40 – 1.63	40 – 1.97	40 – 1.47	40 – 1.67	40 – 2.09	40 – 1.83
Last bin (Å)	(1.96 – 1.93)	(1.66 – 1.63)	(2.00– 1.97)	(1.49 – 1.47)	(1.70 – 1.67)	(2.13 – 2.09)	(1.86 – 1.83)
No. observations	771,083	1,251,845	715,034	2,033,273	1,686,757	614,681	886,235
Completeness (%)	99.7 (99.61)	99.9 (99.9)	99.9 (99.9)	98.8 (77.3)	99.7 (99.7)	99.6 (99.7)	99.2 (88.3)
R_{pim}	0.026 (0.547)	0.035 (1.191)	0.047 (1.338)	0.031 (0.915)	0.043 (0.894)	0.030 (0.739)	0.036 (0.864)
CC(1/2)	0.998 (0.605)	0.999 (0.327)	0.999 (0.360)	0.999 (0.284)	0.998 (0.638)	0.998 (0.695)	0.999 (0.597)
I/σI	15.4 (1.3)	12.2 (0.7)	11.4 (0.8)	12.1 (0.8)	10.7 (0.9)	13.5 (0.9)	11.5 (0.9)
Redundancy	6.7 (6.5)	6.7 (6.8)	6.7 (6.9)	8.4 (4.8)	10.0 (10.2)	6.7 (6.9)	7.3 (4.6)
Refinement							
Total no. of reflections	108,623	175,962	100,105	229,981	159,409	86,509	115,028
Total no. of atoms	11,776	12,380	12,210	12,528	12,315	11,601	12,210
Final bin (Å)	(1.98 – 1.93)	(1.67 – 1.63)	(2.02 – 1.97)	(1.51 – 1.47)	(1.71 – 1.67)	(2.14-2.09)	(1.88-1.83)
R_{work} (%)	22.6 (35.8)	20.7 (37.3)	22.1 (39.0)	19.9 (36.2)	19.5 (39.7)	21.4 (41.4)	20.2 (38.4)
R_{free} (%)	25.6 (37.9)	23.3 (39.1)	25.9 (39.5)	21.3 (37.2)	22.4 (39.9)	24.4 (39.7)	23.1 (39.8)
Average B factor (Å^2)	50.4	33.4	45.5	24.5	29.8	60.1	39.6
Ramachandran plot Favored, %	98.1	97.9	97.8	98.1	98.5	97.7	97.8
Allowed, %	99.9	99.7	99.9	99.8	99.9	99.8	99.7
Outliers, %	0.1	0.3	0.1	0.2	0.1	0.2	0.3

Values in parenthesis are for the highest resolution shell. R_{merge} is $\Sigma|Io - I| / \Sigma Io$, where Io is the intensity of an individual reflection, and I is the mean intensity for multiply recorded reflections. R_{work} is $\Sigma||Fo - Fc|| / Fo$, where Fo is an observed amplitude and Fc a calculated amplitude. R_{free} is the same statistic calculated with a 5% subset of the data that was included in refinement.

References

(1) Buller, A. R.; Brinkmann-Chen, S.; Romney, D. K.; Herger, M.; Murciano-Calles, J.; Arnold, F. H. *Proc. Natl. Acad. Sci.* **2015**, *112*, 14599.

(2) Dunn, M. F. *Archives of Biochemistry and Biophysics*. **2012**, 519, 154.

(3) Kabsch, W.; IUCr. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125.

(4) Evans, P. R.; Murshudov, G. N. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204.

(5) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.

(6) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **2011**, 67, 235.

(7) Emsley, P.; Cowtan, K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126.

(8) Winn, M. D.; Isupov, M. N.; Murshudov, G. N. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2001**, *57*, 122.

(9) Karplus, P. A.; Diederichs, K. *Science*. **2012**, *336*, 1030.

(10) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12.