Supporting Information for:

Identification and Characterization of an Allosteric Inhibitory Site on Dihydropteroate Synthase

Dalia I. Hammoudeh[†], *Mihir Daté*[†#], *Mi-Kyung Yun*[†], *Weixing Zhang*[†], *Vincent A. Boyd*[‡], *Ariele Viacava Follis*[†], *Elizabeth Griffith*[‡], *Richard, E. Lee*[‡], *Donald Bashford*[†*], *and Stephen W. White*[†§*]

Departments of [†]Structural Biology and [‡]Chemical Biology and Therapeutics, St. Jude Children's Research Hospital, Memphis, TN, 38105

[§]Department of Microbiology, Immunology and Biochemistry, University of Tennessee Health Science Center, Memphis, TN, 38163

[#]Current address: Infection - iMED, AstraZeneca, Waltham, Ma, 02452

[*]Corresponding Authors:

TABLE OF CONTENTS

Methods

NMR Spectroscopy - 2D and 3D Experiments. All 2D and 3D NMR experiments were collected at 298K on a Bruker 800 MHz spectrometer equipped with a ^{1}H and ^{13}C detect, TCI triple-resonance, cryogenic single-axis gradient probe using standard Bruker pulse programs. Two-dimensional ^{1}H-^{15}N HSQC-TROSY spectra were recorded on 0.34 mM ^{1}H,^{15}N-labeled perdeuterated BaDHPS in 25 mM sodium phosphate, pH 7.0, 5 mM deuterated DTT (Cambridge) and 10% D_2O. To monitor the effects of added substrates, ^{1}H-^{15}N HSQC-TROSY spectra were recorded as described above, first with increasing concentrations of PtPP and then with increasing concentrations of pABA to achieve a final 1:1.1:1.1 (BaDHPS:PtPP:pABA) ternary complex. Finally, ^{1}H-^{15}N HSQC-TROSY spectra were collected from this ternary complex in the presence of 1 mM **11**. For all 2D experiments, 32 scans, 2048×200 complex points, and spectral windows of 14 ppm × 45 ppm in the ^{1}H and ^{15}N dimensions, respectively, were collected. All spectra were processed with TopSpin (Bruker Biospin) and analyzed with CARA.[1] Backbone resonances were assigned using the standard triple-resonance experiments HNCA and HN(CO)CA on a 0.7 mM ^{1}H,^{15}N,^{13}C-labeled BaDHPS sample. For HNCA experiments, 64 scans, 2048×80×40 complex points, and spectral windows of 14, 27, and 30 ppm in the ^{1}H, ^{13}C and ^{15}N dimensions, respectively, were acquired. For HN(CO)CA, data collection parameters were, 48 scans, 2048×40×80 complex points, and spectral windows of 14, 45, and 27 ppm in the ^{1}H, ^{13}C and ^{15}N dimensions, respectively.

Computational Methods. The starting structure for the inhibitor-bound enzyme (IE) was based on the crystallographic structure of the complex with compound **11**. The

strategy for choosing initial coordinates for the unresolved loops 1 and 2 was to model them from the YpDHPS intermediate structure, PDB code 3TYZ,[2] so as to place them in the most functionally relevant conformation while maintaining the BaDHPS sequence. For the starting models of the enzyme-substrate (ES) and inhibitor-enzyme-substrate (IES) complexes, the substrates pABA and DHPP and Mg^{2+} were placed inside the active site of each monomer using previously published PDB structures: 3TYZ for pABA and 3TZF for DHPP and Mg^{2+}. The *apo*-enzyme structure (E) was generated simply by removing the inhibitor from the IE structure. His256 which coordinates the β-phosphate, was considered to be protonated.[2]

The starting structures were placed at the center of a periodically repeating box filled with TIP3P water molecules,[3] and a neutralizing number of Na^+ counter-ions. The simulation cell size was chosen such that the distance between the sides of the box and the surface of the solute was at least 12 Å. The system was heated from 0K to 300K in six steps of 50K per 50 ps each followed by equilibration at 300K, using the Langevin method, for 400 ps at a constant pressure of 1atm maintained using isotropic position scaling. In order to remain close to the starting crystal structure during heating and equilibration, weak (0.5 kcal mol^{-1} $Å^{-2}$) center of mass harmonic restraints were applied to all backbone atoms of the protein and carbon atoms of the substrates or compound **11**. With no restraints, another constant pressure equilibration was performed for 400 ps before final constant volume equilibration of 400 ps. Production simulations were run for of 16 ns at constant volume and temperature. A time step of 2 fs was used and the SHAKE algorithm was applied to constrain bonds involving hydrogen atoms. Electrostatic interactions were treated with the particle mesh Ewald method[4] with a

direct-space sum limit of 10 Å. The ff99SB force field[5] was used for protein residues and ions. Previously published sets of partial charges[2] were chosen for substrate atoms while other parameters were generated using Antechamber for substrate molecules.

Normal Mode Analysis (NMA) or quasi-harmonic analysis of the trajectories was performed using the cpptraj program[6] within Amber12.[7] Briefly, the method of NMA is based on the diagonalization of the covariance matrix of atomic displacements observed in the ensemble, which yields a set of eigenvectors and eigenvalues. The eigenvectors represent the directions in 3N dimensional space (where N is number of atoms) describing correlated displacements of atoms. The eigenvalues represent the mean square fluctuation of the total displacement along the corresponding eigenvector. For the analysis of correlations between residues, only α-carbon coordinates were analyzed because more meaningful statistics can be obtained by restricting the number of atoms included in the analysis. Correspondences between quasi-harmonic modes of two different simulations were established by a dot-product test. For animations, all atoms were used.

Surface Plasmon Resonance (SPR) Experiments. Experiments were conducted at 20° C using a Biacore T200 optical biosensor (GE Healthcare). Poly-His tagged BaDHPS was immobilized on polycarboxylate hydrogel-coated gold chips preimmobilized with nitrilotriacetic acid (NiHC200m chips; Xantec Bioanalytics) by capture-coupling, a hybrid method of capture and amine coupling chemistry.[8] The chip was primed in chelating buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 50 μM EDTA, 0.005% Tween20) and was preconditioned at 10 μL/min with three 60s injections of wash buffer (10 mM HEPES pH 8.3, 150 mM NaCl, 350 mM EDTA, 0.05% Tween20)

and one 60 s injection of chelating buffer before being charged with a 60 s injection of 500 μM $NiCl_2$ in chelating buffer. After priming into immobilization buffer (20 mM HEPES pH 7.6, 150 mM NaCl, 1 mM TCEP, 0.005% Tween20, 5% glycerol), carboxyl groups on the hydrogel were activated with *N*-ethyl-*N′*-(3-dimethylaminopropyl) carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS), and DHPS was injected until immobilization of ~3100 RU was achieved. Any remaining active sites were blocked by Tris molecules in the binding analysis buffer (20 mM Tris pH 7.6, 150 mM NaCl, 1 mM TCEP, 0.01% Tween20, 5% glycerol, 5% DMSO). One flow cell on the chip was charged with Ni^{2+} and activated with EDC/NHS without adding protein to be used as a reference cell.

Compound **1** was prepared in binding analysis buffer as a 3-fold dilution series starting at 200 μM and was injected at a flow rate of 100 μL/min. A series of buffer-only (blank) injections was included throughout the experiment to account for instrumental noise. The data were processed, double-referenced, solvent corrected[9] and analyzed using the software package Scrubber2 (version 2.0c, BioLogic Software). The equilibrium dissociation constant (K_D) was determined by fitting the data to a 1:1 interaction model.

Table S1. X-ray crystallography data collection statistics

Parameter	Compound 4	Compound 5	Compound 6	Compound 11
Space group	$P6_222$	$P6_222$	$P6_222$	$P6_222$
Cell dimensions (Å)	98.5, 98.5, 262	99.7, 99.7, 263.2	98.3,98.3,263.4	98.8, 98.8, 263.3
Resolution range (Å)[a]	50.0-2.0 (2.07-2.0)	50.0-2.18 (2.26-2.18)	50.0-1.77 (1.83-1.77)	30.0-2.3 (2.38-2.3)
Rsym[a,b]	0.078(0.35)	0.074 (0.414)	0.089 (0.597)	0.174 (0.356)
I/σ[a]	27.9(4.4)	50.0(3.1)	25.9(2.2)	8.0(3.0)
Completeness (%)[a]	90.5(48.1)	86.3(65.5)	94.8(71.6)	90.5(54.0)
Redundancy[a]	11.5(6.7)	11.0(6.3)	11.2(6.2)	11.1(7.4)
Unique reflections	47,488	35,835	70,993	31,293

[a] Values in parentheses refer to the highest resolution shell.
[b] $R_{sym} = \Sigma \left| (I - \langle I \rangle) \right| / \Sigma(I)$, where I is the observed intensity.

Table S2. X-ray crystallography refinement statistics

Parameter	Compound 4	Compound 5	Compound 6	Compound 11
Resolution range (Å)	32.68 - 1.99	49.92 - 2.18	32.19 - 1.77	29.94 - 2.30
No. of reflections	47,064	35,313	70,904	29407
No. in test set (5%)	2397	1772	3575	1565
R_{work}	0.219	0.226	0.192	0.207
R_{free}[a]	0.250	0.280	0.220	0.253
Rmsd from ideal values				
Bond lengths (Å)	0.007	0.007	0.006	0.011
Bond angles (°)	1.050	1.065	1.062	1.448
Ramachandran plot				
Favored (%)	98.32	97.93	98.69	96.93
Allowed (%)	1.31	1.69	1.12	2.69
Outliers (%)	0.37	0.38	0.19	0.38
PDB code	**4NHV**	**4NIL**	**4NIR**	**4NL1**

[a] R_{free} is the R value obtained for a test set of reflections consisting of randomly selected 5% subset of the data set excluded from refinement.

Table S3. NMR ^{1}H and ^{15}N assigned chemical shifts on BaDHPS

Residue	^{1}H (ppm)	^{15}N (ppm)	Residue	^{1}H (ppm)	^{15}N (ppm)
Ser -18	7.53	117.369	Gly 64	8.327	110.528
Ser -17	8.409	118.897	Glu 65	8.402	121.088
His -16	7.809	123.745	Ser 66	7.959	118.232
His -14	9.117	127.265	Thr 67	8.468	118.963
His -13	8.749	127.464	Arg 68	8.177	118.722
His -12	8.129	118.365	Gly 70	7.816	114.713
His -11	8.136	121.022	Phe 71	8.027	120.158
Ser -10	8.096	120.884	Ala 72	7.734	124.874
Ser -9	7.614	119.067	Lys 73	8.572	121.354
Gly -8	8.329	109.668	Val 74	7.278	123.014
Leu -7	8.415	122.018	Ser 75	7.877	111.458
Val -6	8.061	122.616	Val 76	8.715	113.052
Arg -4	8.054	119.295	Glu 77	8.558	128.792
Gly -3	8.388	111.259	Glu 78	9.403	125.339
Ser -2	6.931	118.299	Glu 79	9.008	126.401
His -1	8.865	128.062	Ile 80	8.263	117.037
Met 1	8.013	122.151	Lys 81	8.429	122.274
Lys 2	9.382	127.132	Arg 82	8.063	128.859
Trp 3	7.932	122.35	Trp 123	8.102	124.622
Asp 31	8.395	124.276	Gly 124	7.863	113.185
Ser 32	8.354	120.026	Ala 125	8.552	123.294
Phe 33	8.477	120.756	Lys 126	7.83	127.862
Ser 34	8.531	118.432	Ala 127	7.83	121.952
Asp 35	7.905	123.878	Asn 147	7.755	118.697
Gly 36	8.231	109.731	Arg 148	9.362	122.682
Gly 37	8.129	109.134	Asp 149	7.04	130.121
Ser 38	7.83	119.361	Asn 150	8.885	114.845
Tyr 39	7.251	121.952	Gly 186	7.802	114.513
Asn 40	7.183	110.661	Ile 187	7.618	123.811
Glu 41	8.313	122.151	Gly 188	7.802	103.754
Asp 43	7.884	122.483	Phe 189	8.32	119.229
Ala 44	7.012	121.487	Ala 190	8.238	123.944
Ala 45	8.143	119.229	Lys 266	7.973	119.229
Val 46	8.449	125.14	Met 267	8.265	116.838
Arg 47	8.245	125.206	Met 268	7.945	119.959
His 48	8.626	112.322	Asp 269	8.177	121.088
Ala 49	8.429	117.103	Ala 270	8.197	123.147
Glu 55	8.667	117.967	Met 271	7.911	122.35
Gly 56	7.108	103.953	Ile 272	115.31	115.31
Ala 57	7.916	123.474	Gly 273	115.853	115.853
His 58	6.972	117.927	Lys 274	115.775	115.775
Ile 59	8.279	122.948	Gly 275	112.786	112.786
Asp 61	7.921	116.771	Val 276	118.697	118.697
Ile 62	8.388	123.274	Lys 277	130.32	130.32
Gly 63	8.531	112.454			

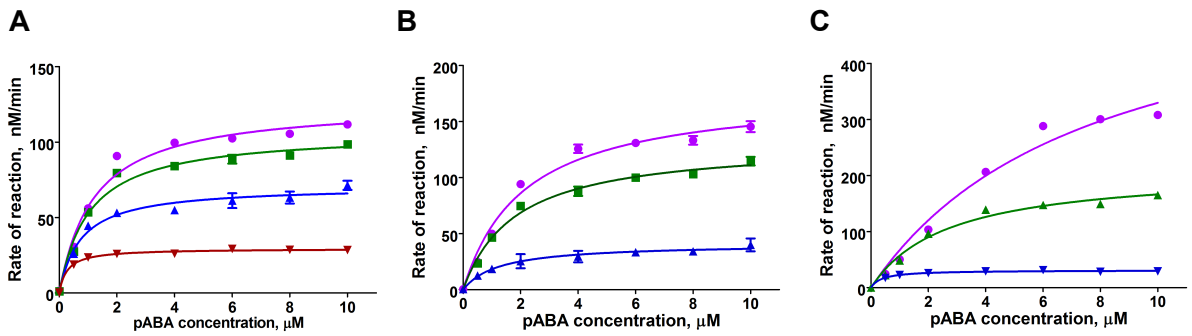

Figure S1. Michaelis-Menten kinetics. Initial rates of 5 nM (A) BaDHPS, (B) YpDHPS, and (C) SaDHPS catalysis were measured at increasing concentrations of pABA (0, 0.5, 1, 2, 4, 6, 8, 10 μM), a fixed DHPP concentration of 20 μM at increasing concentrations of **11**: 0 μM (magenta curves), 50 μM (green), 100 μM (blue) and 200 μM (red – only measured for the BaDHPS enzyme (A)). Refer to Table **1** for the kinetic parameters K_m and V_{max}. Data were analyzed by using GraphPad Prism software.

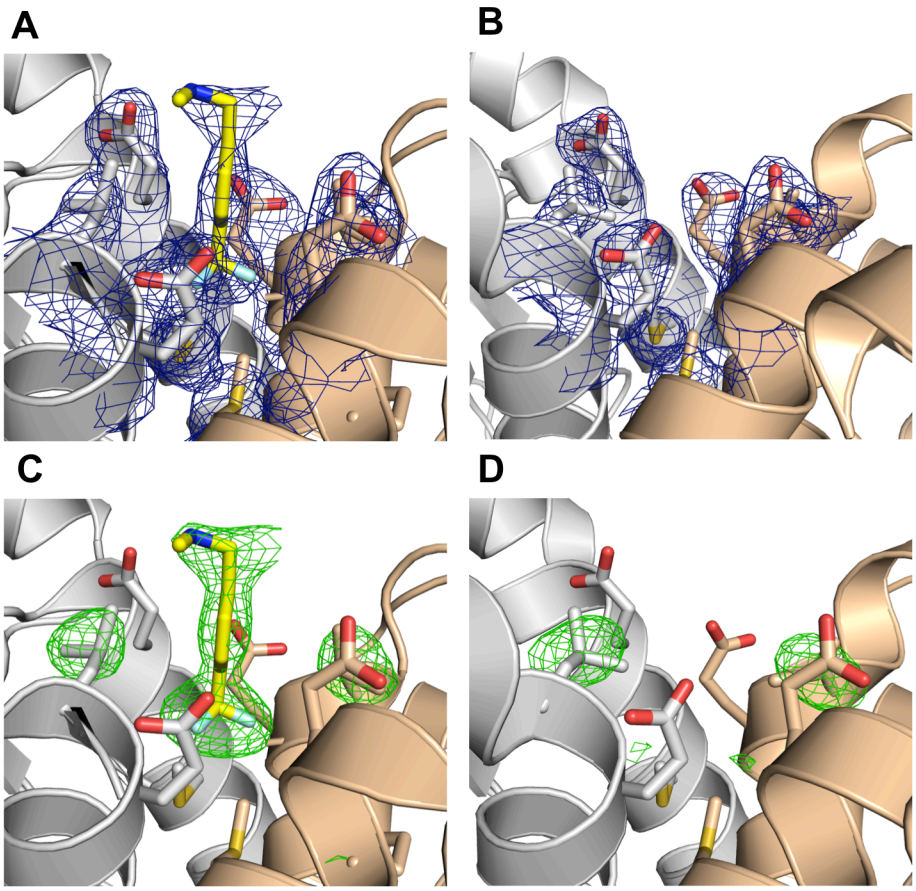

Figure S2. Clear electron density at the twofold symmetry axis is only present in DHPS crystals soaked with dimer interface inhibitors. (A) The 2Fo-Fc density map (blue mesh) at the dimer interface showing compound 11 and neighboring residues. (B) The identical map shown in (A) for a previously published structure (pdb code 4D9P) containing a pterin pocket inhibitor[10]. (C) Simulated annealing (SA) omit map (green mesh) of the compound **11** structure in which the compound is omitted together with the side chain of Leu235 as an internal reference. (D) The identical map shown in (C) for the structure shown in (B). The 2Fo-Fc maps are contoured at 1 σ and the SA omit maps are contoured at 3 σ.

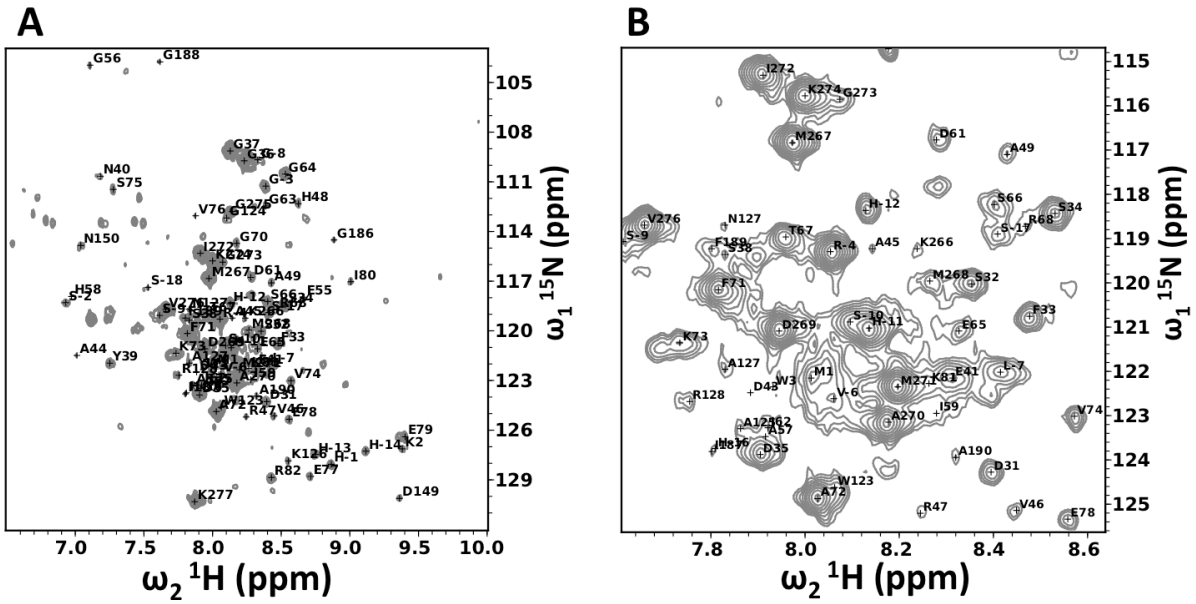

Figure S3. BaDHPS NMR assignments. **(**A) ^{1}H-^{15}N-TROSY-HSQC spectrum of BaDHPS with assigned resonances labeled. (B) Close-up of overlapped region in (A). See Supplementary Table S3 for the ^{1}H and ^{15}N chemical shifts.

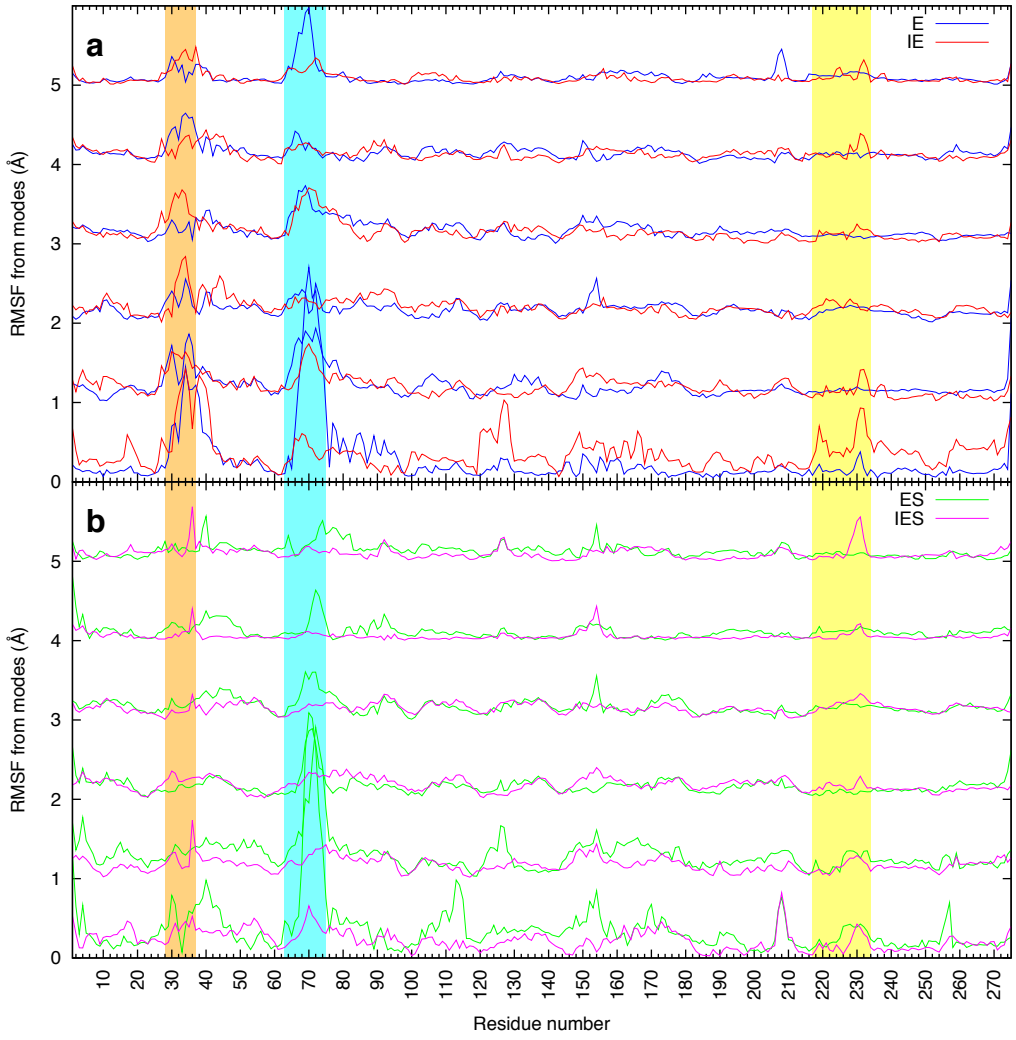

Figure S4. RMSF from eigenvectors: RMSF from first six dominant quasi-harmonic modes. Compared quasi-harmonic modes from different simulations are matched by the dot-product method. Plots for each successive mode are offset by 1.0 Å for clarity. Highlighting of loop regions and plot line colors are as in Figure 6. (A) The simulations without substrates (E and IE). (B) The simulations with substrates (ES and IES).

Figure S5. Relative rotation of the monomers within the YpDHPS dimer upon enzyme catalysis. (Yellow) apo YpDHPS structure (pdb ID: 3TZN), (red) intermediate-state YpDHPS structure (pdb ID: 3TYZ), (magenta) product analog bound YpDHPS structure (pdb ID: 3TYU). In this figure, one monomer in the dimer was fully superimposed to emphasize the relative motion of the second monomer. The rotation (around the axis of helix α7 at the dimer interface) is most obvious within the box and indicated by the arrow.

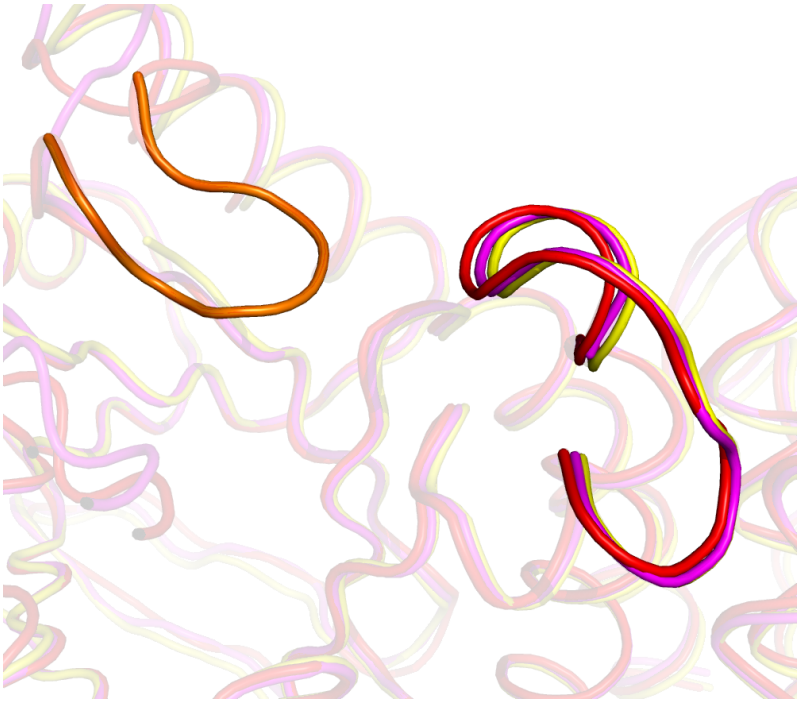

Figure S6. Subtle loop7 movements relative to loop1 in YpDHPS upon catalysis. (Yellow) apo YpDHPS structure (pdb ID: 3TZN), (red) intermediate-state YpDHPS structure (pdb ID: 3TYZ), (magenta) product analog bound YpDHPS structure (pdb ID: 3TYU). In this figure, the cores of one monomer from all three structures were aligned to show the loop7 motion in relation to the body of the protein. The loop1 (orange) structure is taken from the intermediate-state structure.

Movies. Animation movies of the most dominant Quasi-harmonic mode from simulation E (Movie S1), IE (Movie S2), ES (Movie S3) and IES (Movie S4). Movies were made using Chimera[11] and the PCAsuite[12] animation utility to which Amber12 trajectory was fed as input. For each movie, loop1 and loop2 are in orange ribbon and loop7 and helix7 are in blue ribbon representation. Only select enzyme residues, compound **11** and substrates are shown in stick, while Mg^{2+} ions are in ball representation. Carbon atoms are colored tan (enzyme), light green (compound **11**), and dark green (substrates). Mg^{2+} cations and fluorine atoms of **11** are colored in magenta. Movies S1 (simulation E) and S3 (simulation ES), show a breathing motion where monomers move away from each other and come closer again. In Movies S2 (simulation IE) and S4 (simulation IES), this breathing motion is frozen.

Movie S1. Animation movie of the most dominant Quasi-harmonic mode from simulation of apo DHPS (E).

Movie S2. Animation movie of the most dominant Quasi-harmonic mode from simulation of DHPS:**11** (IE).

Movie S3. Animation movie of the most dominant Quasi-harmonic mode from simulation of DHPS:Substrates (ES).

Movie S4. Animation movie of the most dominant Quasi-harmonic mode from simulation of DHPS:**11:**Substrates (IES).

1. Keller, R. L. J. (2005) Optimizing the Process of Nuclear Magnetic Resonance Spectrum Analysis and Computer Aided Resonance Assignment. In *Molecular Biology and Biophysics*, Institute of Molecular Biology and Biophysics at ETH Zurich.
2. Yun, M.-K., Wu, Y., Li, Z., Zhao, Y., Waddell, M. B., Ferreira, A. M., Lee, R. E., Bashford, D., and White, S. W. (2012) Catalysis and Sulfa Drug Resistance in Dihydropteroate Synthase. *Science 335*, 1110–1114.

3. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys 79*, 926–935.
4. Cheatham, T. E., III, Miller, J. L., Fox, T., Darden, T. A., and Kollman, P. A. (1995) Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J Am Chem Soc 117*, 4193–4194.
5. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins 65*, 712–725.
6. Roe, D. R., and Cheatham III., T. E. (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput 9*, 3084-3095.
7. D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Goetz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman. (2012) AMBER 12, University of California, San Francisco.
8. Rich, R. L., Errey, J., Marshall, F., and Myszka, D. G. (2011) Biacore analysis with stabilized G-protein-coupled receptors. *Anal Biochem 409*, 267–272.
9. Myszka, D. G. (1999) Improving biosensor analysis, *J Mol Recognit 12*, 279–284.
10. Zhao, Y., Hammoudeh, D., Yun, M. K., Qi, J., White, S. W., and Lee, R. E. (2012) Structure-based design of novel pyrimido[4,5-c]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem 7*, 861–870.
11. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem 25*, 1605–1612.
12. Meyer, T., Ferrer-Costa, C., Perez, A., Rueda, M., Bidon-Chanal, A., Luque, F.J., Laughton, C.A., and Orozco, M. (2006) Essential dynamics: A tool for efficient trajectory compression and management. *J Chem Theory Comput 2*, 251–258.