Experimental and Computational Mutagenesis to Investigate the Positioning of a General Base within an Enzyme Active Site

Jason P. Schwans[†], Philip Hanoian[‡], Benjamin J. Lengerich[‡], Fanny Sunden[†], Ana Gonzalez[a], Yingssu Tsai[a+], Sharon Hammes-Schiffer[§]*, and Daniel Herschlag[†]*

[†] *Department of Biochemistry, B400 Beckman Center, 279 Campus Drive, Stanford University, Stanford, California, 94305*

[‡] *Department of Chemistry, 104 Chemistry Building, Pennsylvania State University, University Park, Pennsylvania 16802*

[§] *Department of Chemistry, 600 South Mathews Avenue, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*

[a] *Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, California 94025*

[+] *Department of Chemistry, 333 Campus Drive, Stanford University, Stanford, CA 94305*

SUPPORTING INFORMATION

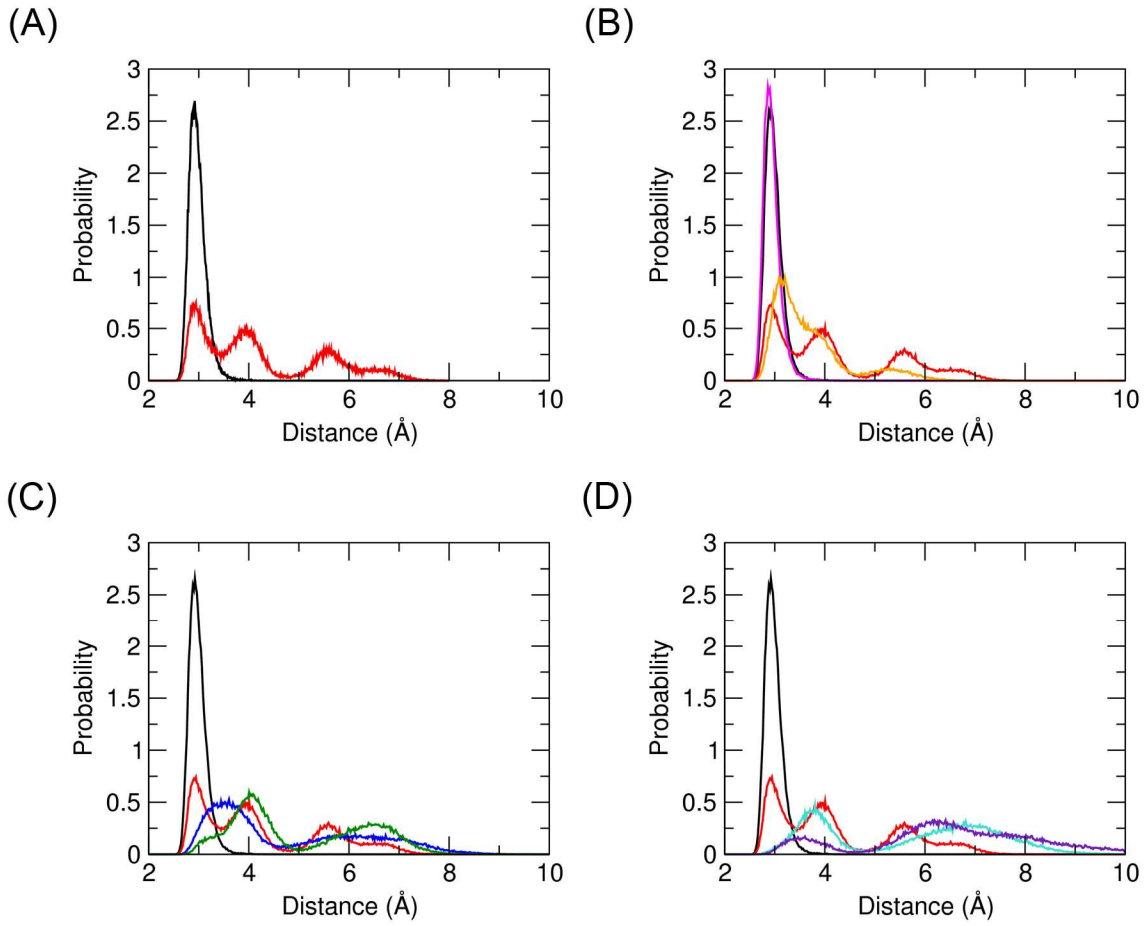

Figure S1. Histograms of distances between C4 of the steroid and the carboxylate oxygen of the general base observed in MD simulations. (A) WT (black) and D38E (red). (B) WT (black), D38E (red), A114G (pink), and D38E/A114G (orange). (C) WT (black), D38E (red), P39G (blue), and D38E/P39G (green). (D) WT (black), D38E (red), P39G/V40G/S42G (cyan), and D38E/P39G/V40G/S42G (purple).

Table S1. Peak maxima (in Å) observed in histograms of donor-acceptor distances between C4 of the steroid and the carboxylate oxygen of the general base in MD simulations (Figure S1).

Enzyme	Peak 1	Peak 2	Peak 3
WT	2.90		
D38E	2.92	4.00	5.58
P39G	3.56	6.08	
D38E/P39G	4.00	6.56	
P39G/V40G/S42G	3.68	6.76	
D38E/P39G/V40G/S42G	3.56	6.28	
A114G	2.84		
D38E/A114G	3.08	5.24	

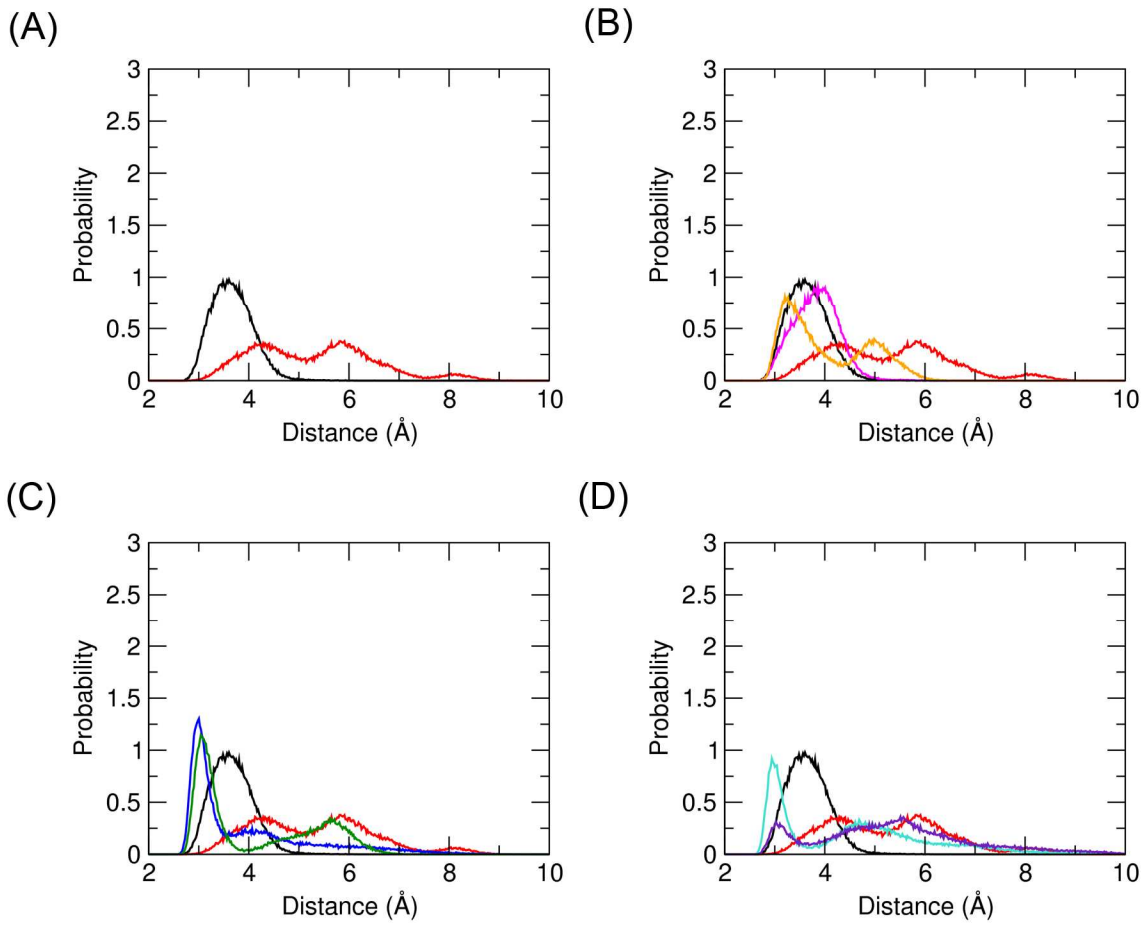

Figure S2. Histograms of distances between C6 of the steroid and the carboxylate oxygen of the general base observed in MD simulations. (A) WT (black) and D38E (red). (B) WT (black), D38E (red), A114G (pink), and D38E/A114G (orange). (C) WT (black), D38E (red), P39G (blue), and D38E/P39G (green). (D) WT (black), D38E (red), P39G/V40G/S42G (cyan), and D38E/P39G/V40G/S42G (purple).

Table S2. Peak maxima (in Å) observed in histograms of donor-acceptor distances between C6 of the steroid and the carboxylate oxygen of the general base in MD simulations (Figure S2).

Enzyme	Peak 1	Peak 2	Peak 3
WT	3.52		
D38E	4.12	5.80	8.00
P39G	2.96	4.04	
D38E/P39G	3.04	5.68	
P39G/V40G/S42G	2.92	4.68	
D38E/P39G/V40G/S42G	3.00	5.48	
A114G	3.96		
D38E/A114G	3.20	4.96	

Table S3. Percent occupancy values of hydrogen bonds in the oxyanion hole.[a]

Enzyme	Tyr14	Asp99	Asp38(Glu)	Tyr14-Tyr55
WT	96.8 (0.4)	87.3 (1.8)	0.05 (0.1)	72.2 (2.2)
D38E	90.7 (1.7)	85.3 (4.6)	21.55 (7.1)	53.5 (4.1)
P39G	94.3 (2.6)	87.5 (6.3)	0.0 (0.0)	68.8 (9.2)
D38E/P39G	95.1 (2.6)	91.7 (4.0)	0.0 (0.0)	59.2 (22.3)
P39G/V40G/S42G	92.1 (7.8)	86.1 (4.9)	0.0 (0.0)	66.1 (15.5)
D38E/P39G/V40G/S42G	95.6 (0.8)	91.7 (2.6)	0.0 (0.0)	66.5 (13.9)
A114G	96.2 (0.7)	91.4 (2.2)	0.23 (0.3)	70.2 (1.6)
D38E/A114G	93.0 (0.5)	90.5 (0.9)	26.28 (28.0)	64.5 (14.4)
S42G	95.1 (3.3)	85.3 (4.1)	0.05 (0.1)	63.7 (11.9)
V40G/S42G	94.0 (1.5)	85.8 (2.3)	0.05 (0.1)	69.3 (1.9)

[a] Hydrogen bonds are determined with the criteria of an O-O distance within 3.5 Å and an O-H-O angle within 30° of linearity. Occupancies are shown for hydrogen bonds of Tyr14, Asp99, and Asp38(Glu) with the steroid oxygen as well as the hydrogen bond between Tyr14 and Tyr55. Results obtained by averaging over two 20-ns MD trajectories and two active sites. Standard deviations among data sets are shown in parentheses.

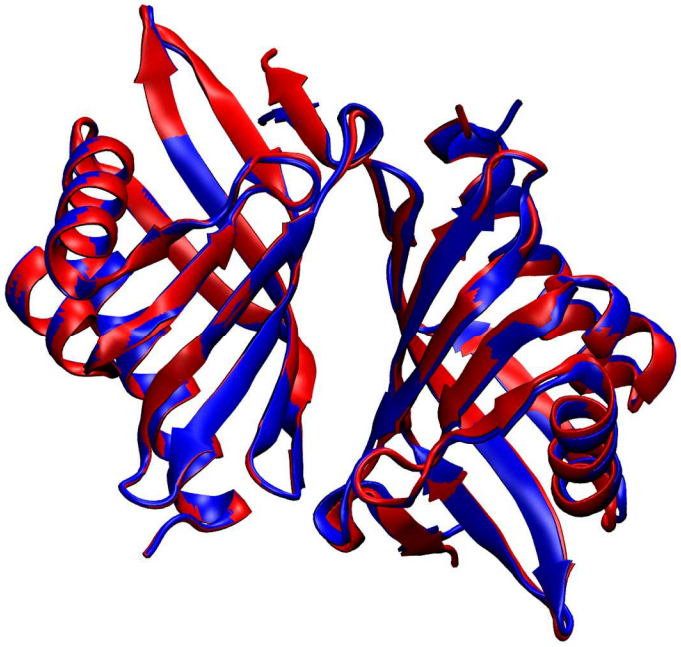

Figure S3. Overlay of the average structures from MD simulations of WT KSI (blue) and A114G KSI (red) illustrating that the average structures are very similar overall with a C_α RMSD of only 0.84 Å for the dimer.

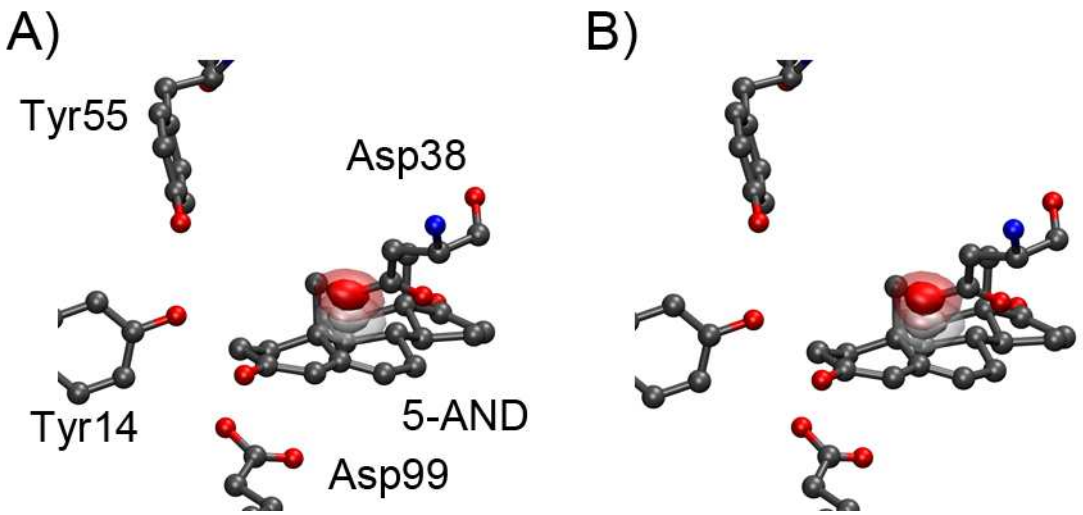

Figure S4. Atomic isodensity surfaces of the protonated catalytic base carboxylate group from the MD simulations of WT KSI (A) and the S42G mutant (B). Oxygen density is shown in red and hydrogen density is shown in white. In each case, the most probable 90% of bins are drawn in a transparent isosurface, and the most probable 50% of bins are drawn in an opaque isosurface. The MD average structures of the heavy atoms of the steroid intermediate (5-AND), the catalytic base Asp38, and the active site residues Tyr14, Tyr55, and Asp99 are shown.

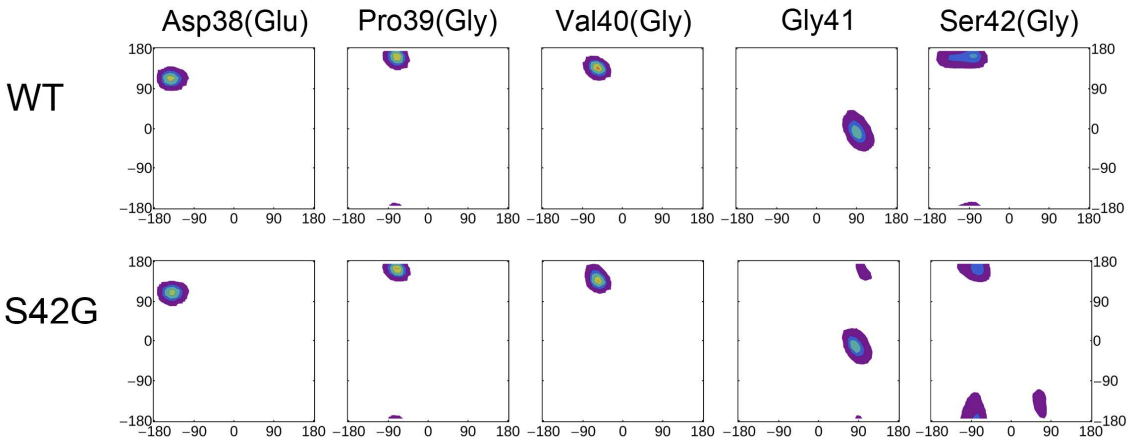

Figure S5. Ramachandran plots of the catalytic base loop residues Asp38, Pro39(Gly), Val(40)Gly, Gly41, and Ser42(Gly) obtained from MD trajectories by calculating the probability distribution function averaged over all data sets. The x-axis corresponds to the angle φ, and the y-axis corresponds to the angle ψ, both expressed in units of degrees. Each row corresponds to the specified mutant, and each column corresponds to the specified residue in the catalytic base loop.

Table S4. Crystallographic data collection and refinement statistics.

Data Set	**P39A**	**P39G V40G S42G**	**D38E P39G V40G S42G**
PDB ID	3MHE	3OV4	3NM2
Resolution Range (Å)	29.7-1.72 (1.77-1.72)[a]	37.5-1.83 (1.87-1.82)[a]	29.7-1.89 (1.94-1.89)[a]
Space Group	$P2_12_12_1$	$P6_122$	$P6_522$
a, Å	40.8	64.4	59.4
b, Å	66.0	64.4	59.4
c, Å	86.9	505.6	142.0
α, °	90.0	90.0	90.0
β, °	90.0	90.0	90.0
γ, °	90.0	120.0	120.0
Number Unique Reflections	25533 (38331)	56752 (2680)	12563 (1741)
Completeness	99.8 (96.8)	98.6 (86.8)	99.5 (94.5)
Multiplicity	13.6 (10.6)	13.1 (17.2)	13.2 (10.8)
R_{merge}, %	5.1	6.7	4.5
Mean I/σ	4.4	10.8	0.8
Refinement Statistics			
No. Residues	248	496	124
No. Waters	156	502	85
R_{work}, %	18.7 (34.4)	18.7 (31.9)	20.2 (29.1)
R_{free}, %	23.3 (34.4)	22.8 (36.9)	25.4 (35.0)
rmsd bond, Å	0.026	0.022	0.026
rmsd angle, °	2.215	1.857	2.137

$R_{merge}=\Sigma_{hkl}\Sigma_i|I(hkl)_i-\{I(hkl)\}|/\Sigma_{hkl}\Sigma_i I(hkl)_i$

$R_{work}=\Sigma_{hkl}|F(hkl)_o-\{F(hkl)_c\}|/\Sigma_{hkl}F(hkl)_o$

R_{free} was calculated exactly as R_{work} where $F(hkl)_o$ were taken from 10% of the data not included in refinement.

[a] Highest resolution shell is in parentheses.

Table S5. Normalized[a] B-factors for the catalytic base Asp38 from crystal structures of WT KSI and loop mutants.[b]

System	WT	P39A	P39G/V40G/S42G	D38E/P39G/V40G/S42G
PDB ID	8CHO	3MHE	3OV4	3NM2
N	0.06	1.22	1.72	-0.15
C	0.10	1.34	2.09	0.15
O	0.77	1.53	2.24	0.09
C_α	0.10	1.33	1.98	0.09
C_β	0.22	1.60	2.05	0.30
C_γ	0.16	2.04	2.43	0.79
C_δ[c]	NA	NA	NA	1.44
$O_\delta1/O_\varepsilon1$	0.10	2.27	2.77	2.45
$O_\delta2/O_\varepsilon2$	0.71	2.56	2.69	1.75

[a] The normalized B-factors are expressed in terms of the number of standard deviations above or below the mean value of the structure in question as described in Carugo, O. and Argos, P. 1997 *Protein Engineering* 10, 7, 777-787.

[b] Values are averaged over all chains in the asymmetric unit.

[c] Note C_δ is only present in mutants containing a Glu as position 38.

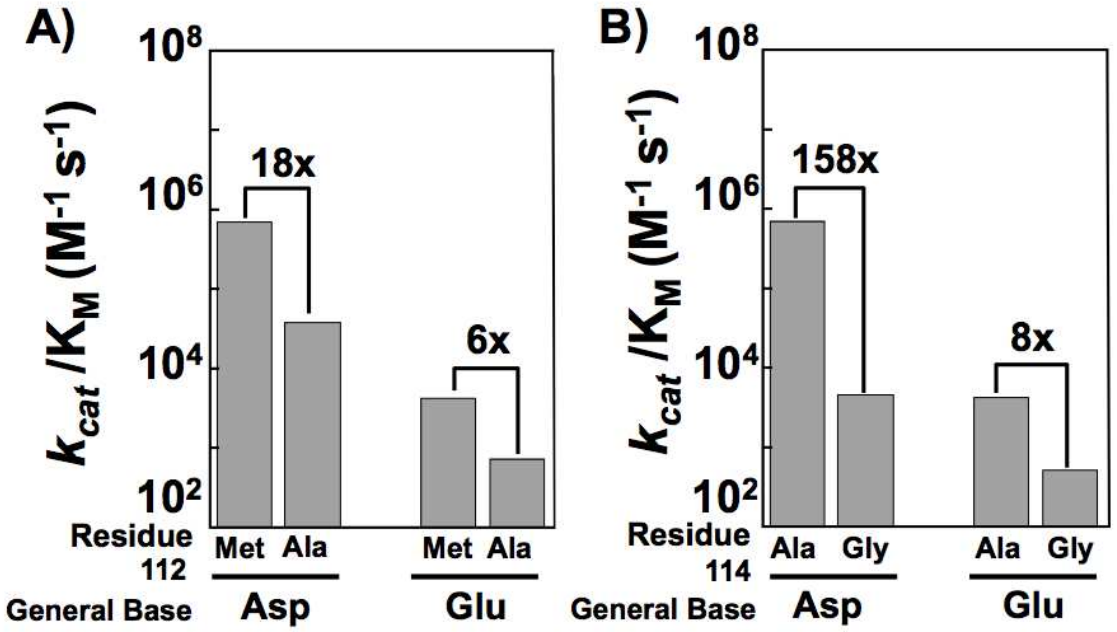

Figure S6. Effects on k_{cat}/K_M for mutation of (A) Ala114 to Gly and (**B**) Met112 to Ala in KSI with a positioned Asp general base and a mispositioned D38E mutant general base. Values are from Table 3.

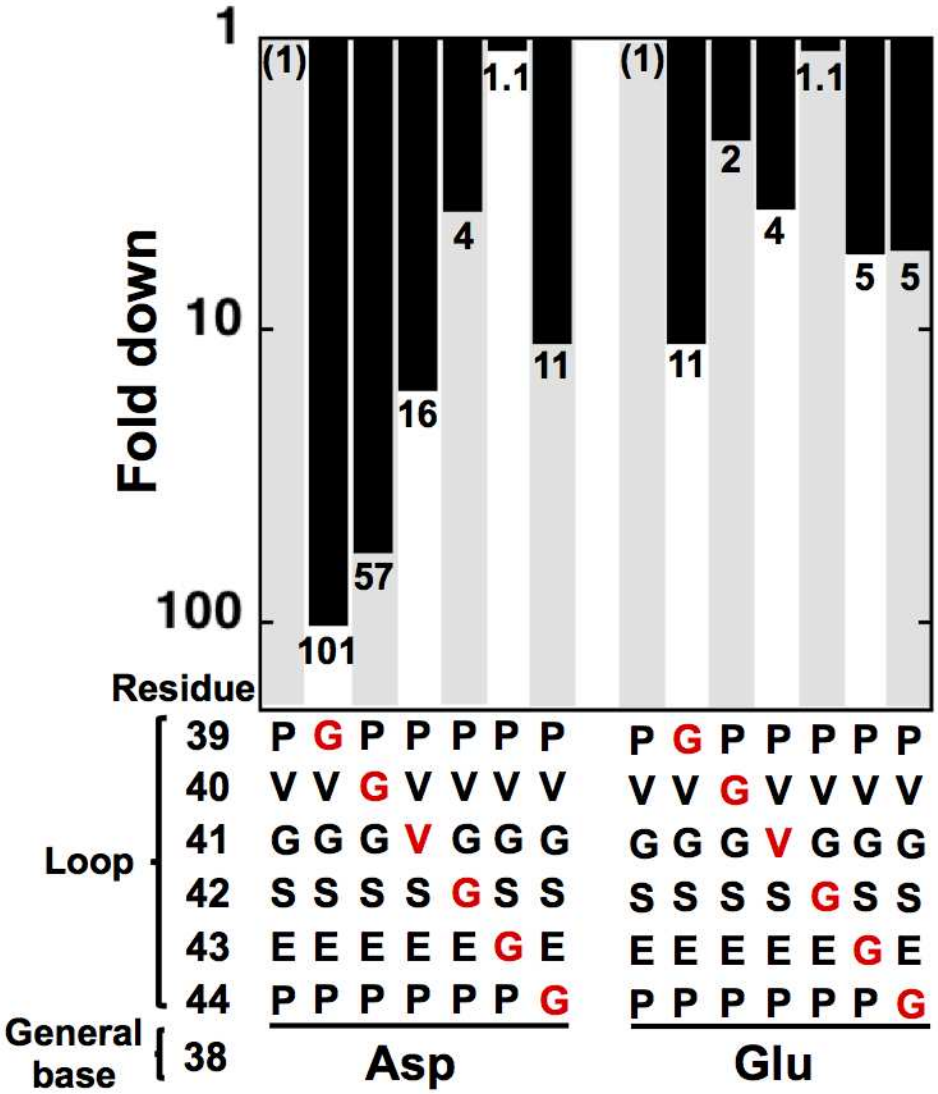

Figure S7. Effects on k_{cat}/K_M for mutation of loop residues in KSI with a positioned Asp38 general base and a mispositioned D38E mutant general base. Values are plotted relative to WT KSI on the left hand side and relative to the D38E mutant on the right hand side. Data are from Table 4.

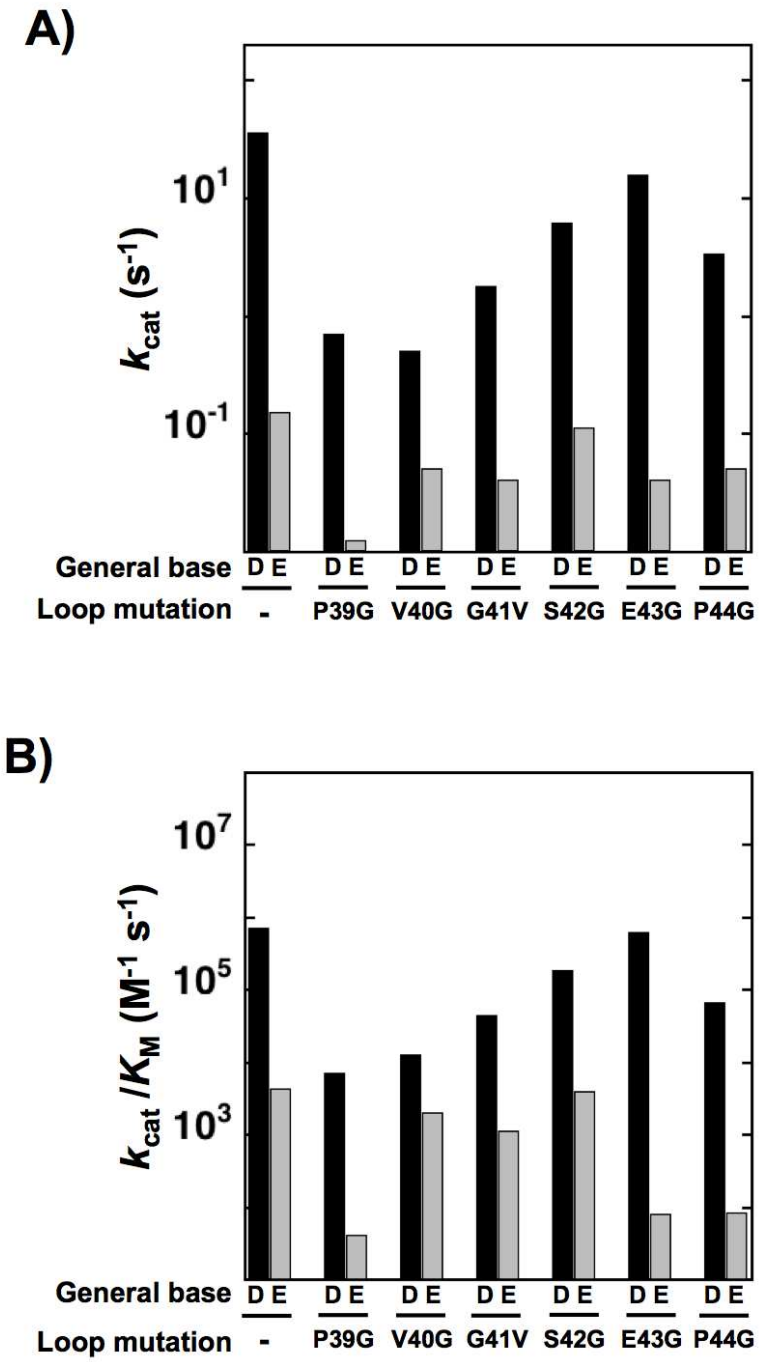

Figure S8. Effects of loop mutations on k_{cat} (A) and k_{cat}/K_M (B). Values are from Table 4. The designations 'D' and 'E' denote the identity of the catalytic base as Asp and Glu, respectively, and the '-' indicates no loop residues are mutated.

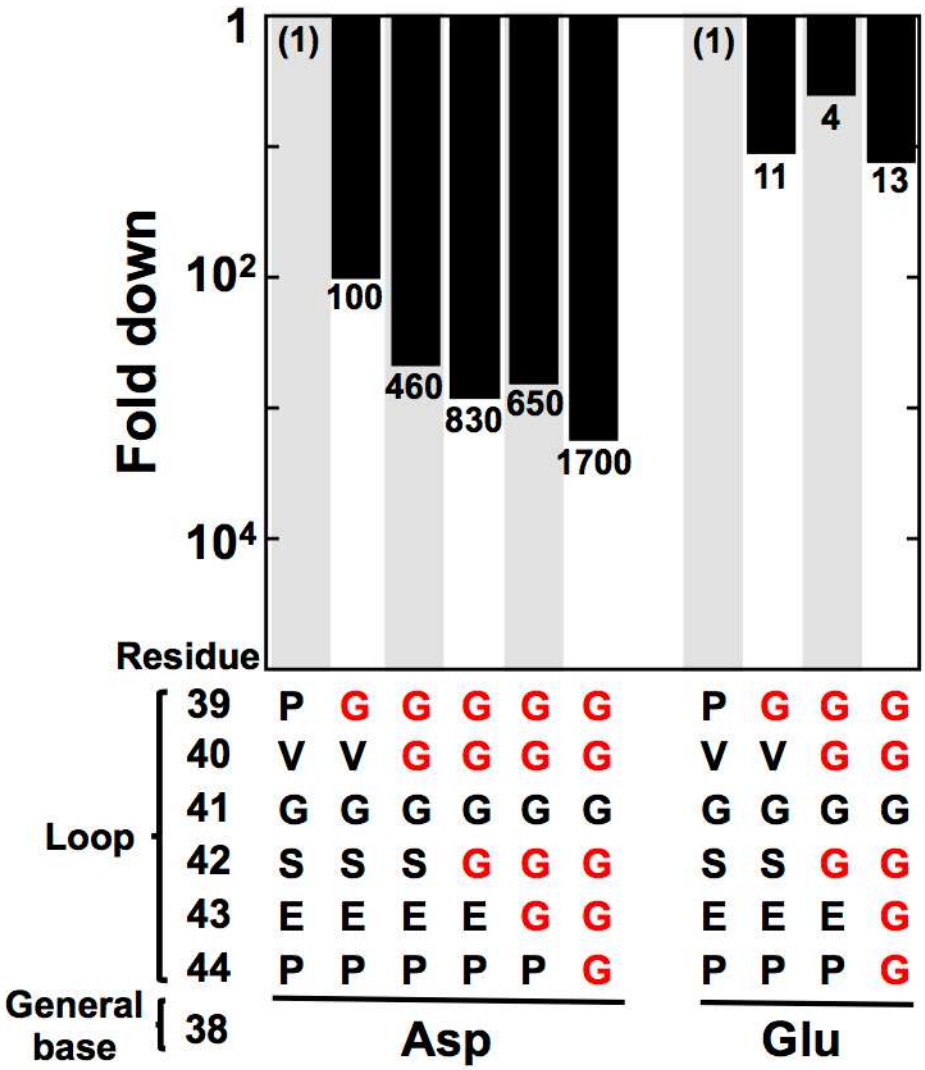

Figure S9. Effects of mutating multiple loop residues to glycine on k_{cat}/K_M in KSI with a positioned Asp38 general base and a mispositioned Glu (D38E mutant) general base are shown. Values are plotted relative to WT KSI on the left hand side and relative to the D38E mutant on the right hand side. Data are from Table 7.

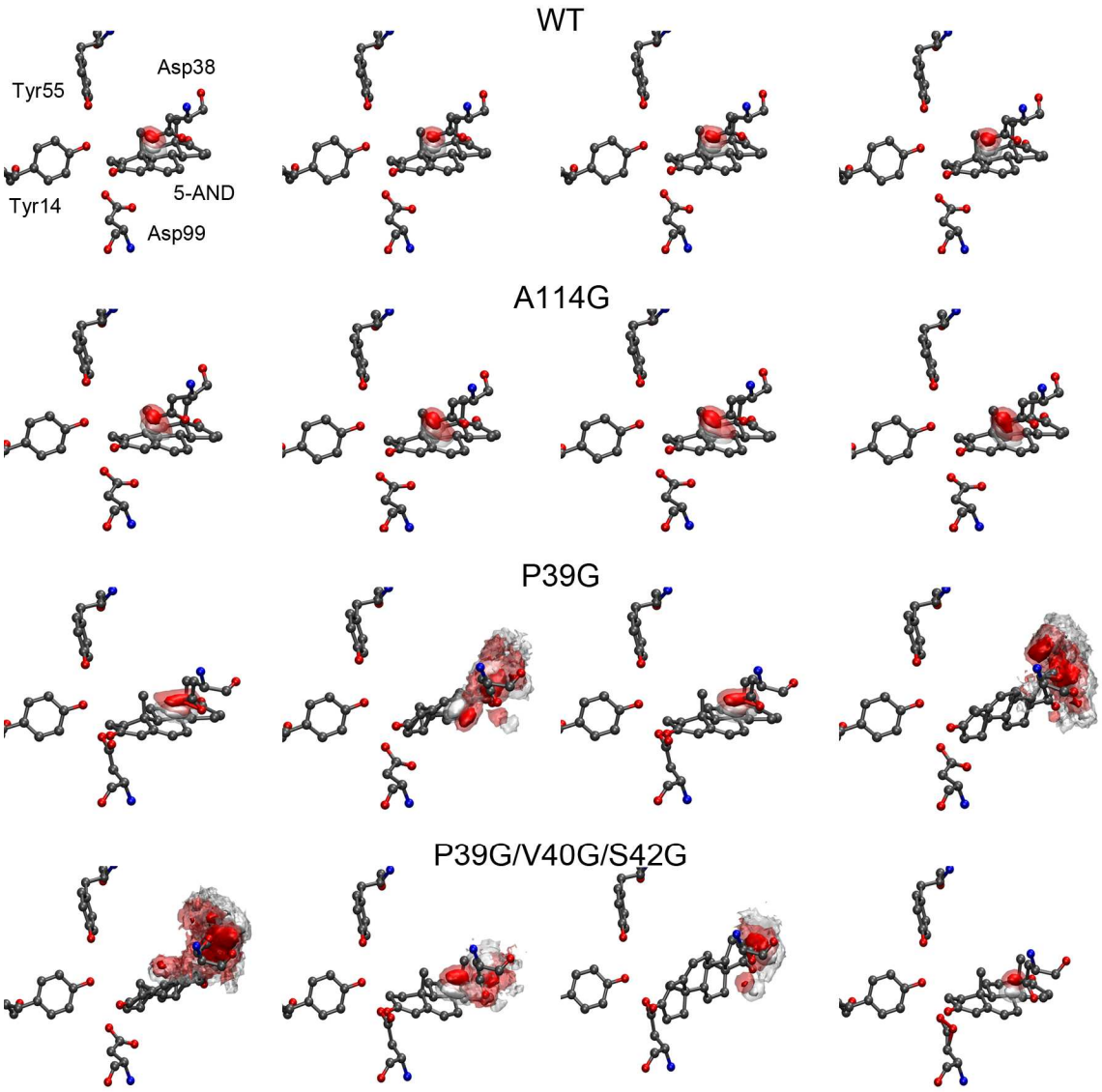

Figure S10. Atomic isodensity surfaces of the protonated catalytic base carboxylate group from independent trajectories and both active sites from the MD simulations. Oxygen density is shown in red and hydrogen density is shown in white. In each case, the most probable 90% of bins are drawn in a transparent isosurface, and the most probable 50% of bins are drawn in an opaque isosurface. The MD average structures of the heavy atoms of the steroid intermediate (5-AND), the catalytic base (Asp38), and the active site residues Tyr14, Tyr55, and Asp99 are shown.

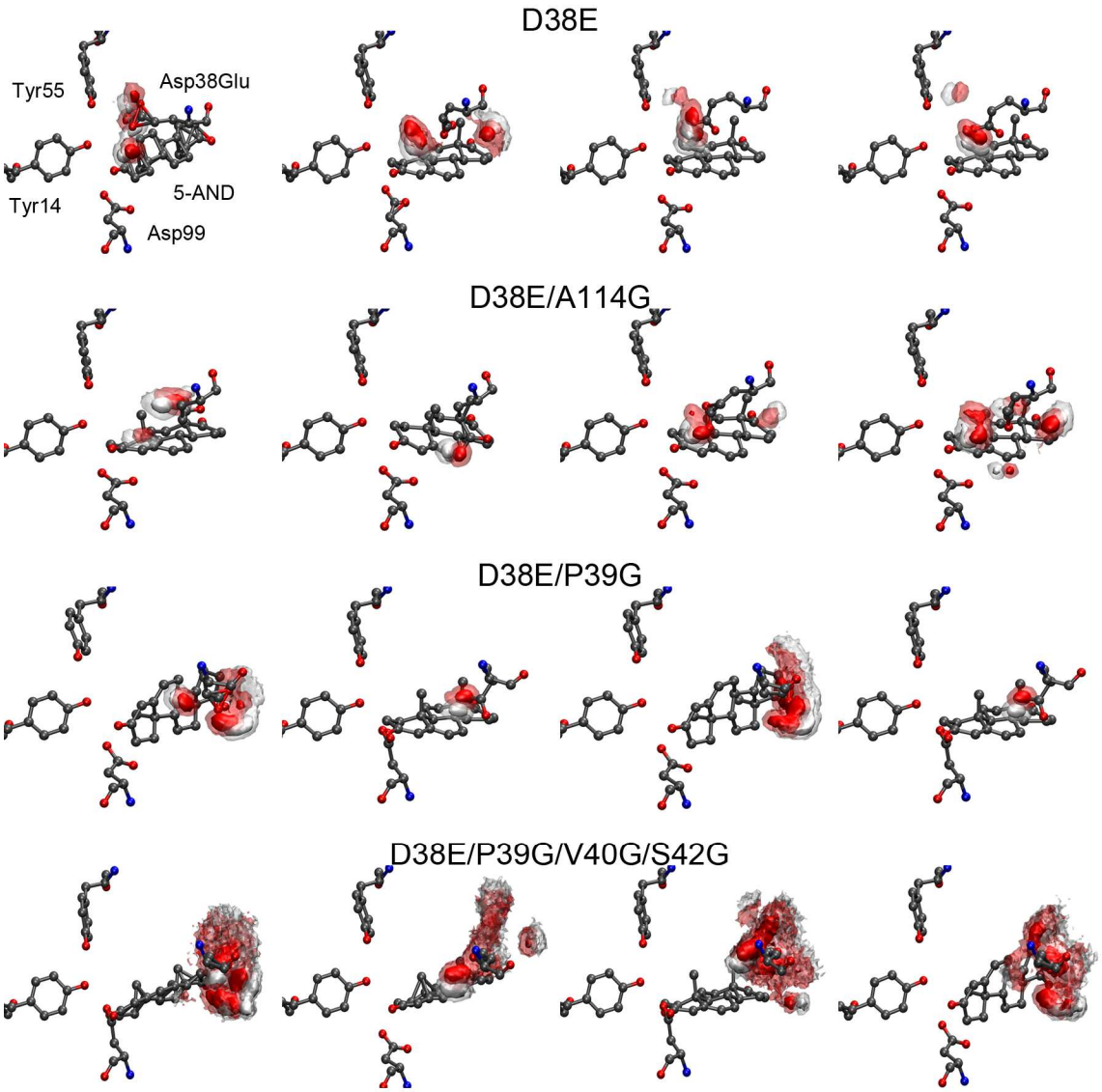

Figure S11. Atomic isodensity surfaces of the protonated catalytic base carboxylate group from independent trajectories and both active sites from the MD simulations. Oxygen density is shown in red and hydrogen density is shown in white. In each case, the most probable 90% of bins are drawn in a transparent isosurface, and the most probable 50% of bins are drawn in an opaque isosurface. The MD average structures of the heavy atoms of the steroid intermediate (5-AND), the catalytic base (Asp38Glu), and the active site residues Tyr14, Tyr55, and Asp99 are shown.

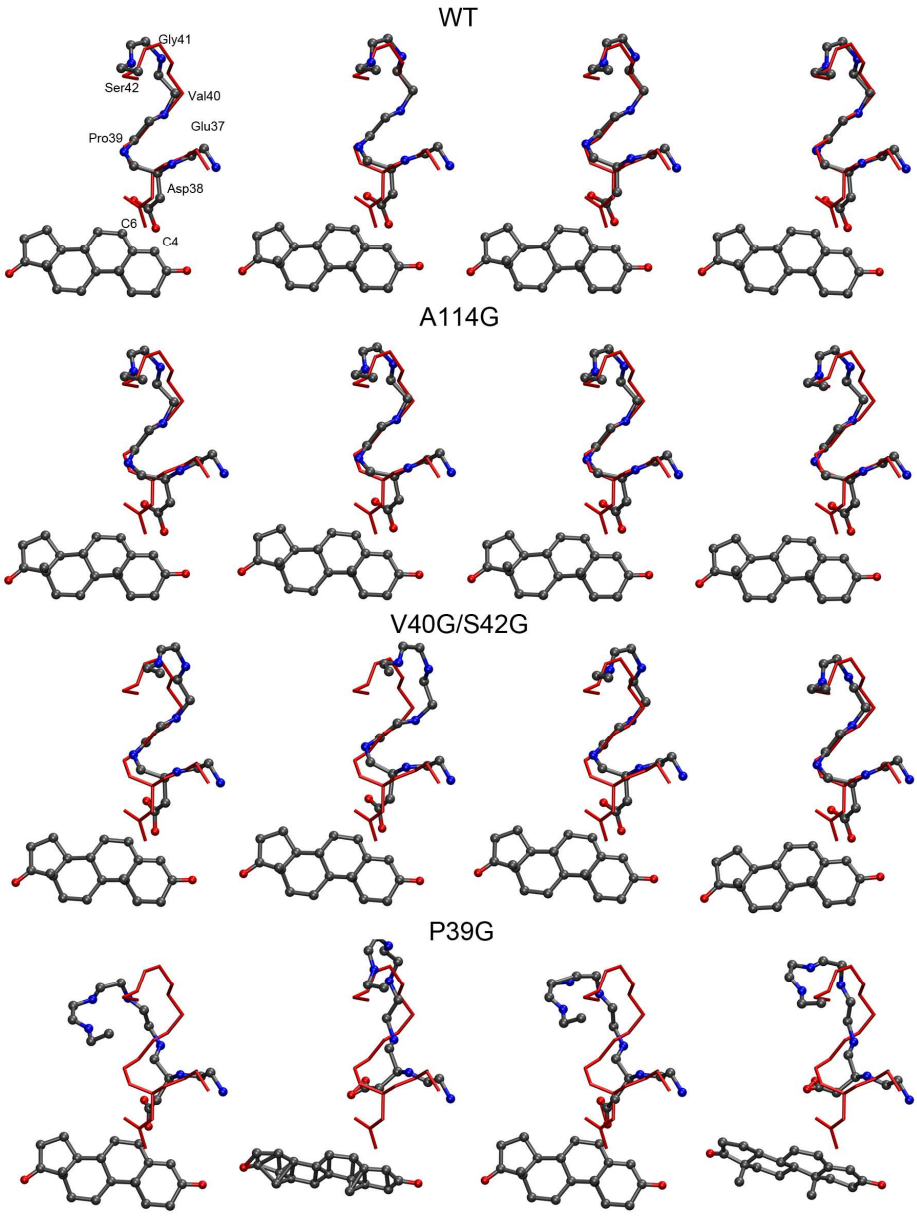

Figure S12. Average structures of WT, A114G, V40G/S42G, and P39G KSI from individual trajectories and active sites from MD simulations. The average structures of residues 37-42 and the catalytic base side chain are depicted as balls and sticks with the corresponding atoms in the D38N-equilenin crystal structure (PDB ID: 1QJG) depicted in red for reference.

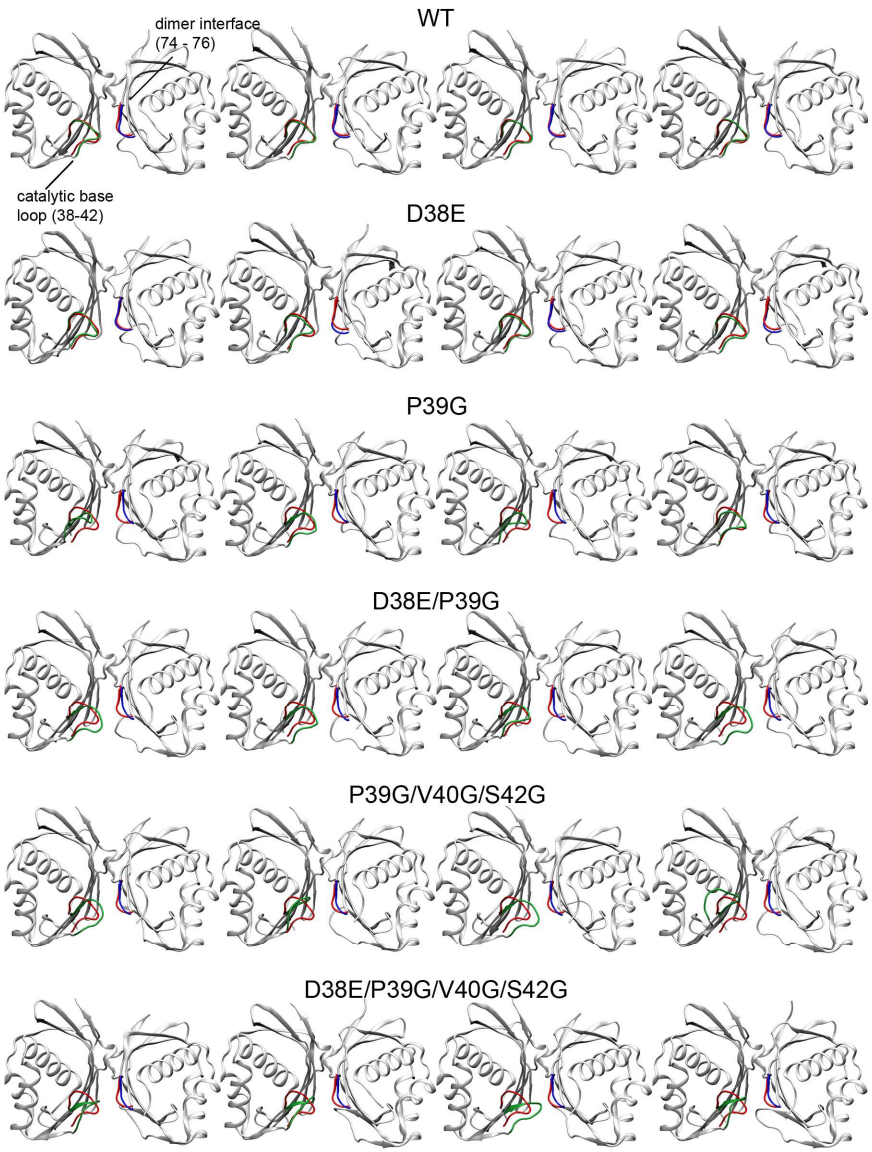

Figure S13. Ribbon diagrams of the average structures of WT KSI and multiple KSI mutants. The average structure of KSI in each case is shown in gray, highlighting the conformational changes in the catalytic base loop residues (38-42, green) and the proximal residues across the dimeric interface (74-76, blue). The backbone structures of these two regions in the D38N-equilenin crystal structure (PDB ID: 1QJG) are shown in red for reference. Structures were aligned by minimization of the RMSD of C_α atoms in the dimer.