Evolutionary Covariance Combined with Molecular Dynamics Predicts a Framework for Allostery in the MutS DNA Mismatch Repair Protein

Bharat Lakhani[1,2], Kelly M. Thayer[3,4], Manju M. Hingorani[1, 2] and David L. Beveridge[2,3]*

[1]Molecular Biology and Biochemistry Department, [2]Molecular Biophysics Program,
[3]Chemistry Department, [4]Computer Science Department,
Wesleyan University, Middletown CT 06459

Supplementary Figure Legends

Figure S1. SCA sector of the MutS protein family. (A) Sectors are defined as a group of co-evolved residues identified by the top significant eigenmode(s) of the SCA correlation matrix. In the case of MutS, one top eigenmode (bottom panel, λ_1 in inset) is well separated from statistical noise (top panel, MSA randomized in 100 trials). (B) Histograms showing the contribution of residues to the top eigenvector of the SCA matrix for the MutS MSA (bottom panel) and randomized MSA (top panel). The sector is defined by ~22% of residues in the tail, corresponding to a cutoff value of 0.04, which selects positions that are outside the expectation for a randomized alignment [42]. (C) A recent study suggests that for several protein family alignments the top eigenvector components can correlate strongly with conservation (although different residues are identified by SCA versus conservation analysis) [80]. A 2 X 2 contingency matrix shows the relationship between sector and conserved residues in the MutS family. The definition of conservation is based on relative entropy: Kullback-Leibler divergence $D_i=\sum_a f_i(a)\log\frac{f_i(a)}{q(a)}$, where $f_i(a)$ is the frequency at which amino acid a occurs in column i of the multiple sequence alignment and $q(a)$ is the background frequency for amino acid a [41]. A Chi-Squared test yields a relatively high p-value of 0.05 for conserved positions with $D_i > 2$; thus, SCA sector and conserved residues do not appear to be strongly correlated in the MutS family of proteins.

Figure S2. MD calculated distance between critical residue pairs in the 6.8 Å cutoff sector contact map. Contact distance distributions between residue pairs (A) A146 (Cα)–Y244 (Cε) and (B) M250 (Cε)–A597 (carbonyl O) calculated from a 15 ns wild type MutS MD trajectory. A red line highlights the mean distance, and the minimum distances are 5.7 Å and 5.0 Å for the two pairs, respectively.

Figure S3. Domain-wise 1D-RMSD comparison between wild type and sector residue alanine mutants of MutS. The 1D-RMSD values calculated as a function of MD time are shown for mismatch binding (I), clamp (IV) and ATPase (V) domains of S1 and S2 subunits. (A) L17A, (B) F39A, (C) E41A, (D) R76A, (E) E99A, (F) P100A, (G) S151A, (H) Q232A, (I) P391A, (J) Q468A, (K) L491A, (L) E500A, (M) V560A, (N) V561A, (O) F567A, (P) L598A, (Q) E699A, (R) F724A and (S) L759A. Also shown are 50 ns data for (T) wild type, (U) R172A and (V) I553A; see also Table I.

Figure S4. Domain-wise 1D-RMSD comparison between wild type and non-sector residue alanine mutants of MutS. The 1D-RMSD values calculated as a function of MD time are shown for mismatch binding (I), clamp (IV) and ATPase (V) domains of S1 and S2 subunits. (A) F78A, (B) K161A, (C) L191A, (D) E204A, (E) P206A, (F) P209A, (G) L341A, (H) I358A and (I) R550A; see also Table I.

Figure S5. Overlay of wild type and alanine mutant contacts between R76 and DNA and K589 and ATP for sector residues. All atom local alignment of R76 and guanine 1546 (left panel), and K589 and ATP (right panel), between wild type (green) and mutant (magenta) proteins shows H-bond disruption at one or both active sites for sector residue mutants. (A) L17A, (B) F39A, (C) E41A, (D) R76A, (E) E99A, (F) P100A, (G) S151A,

(H) Q232A, (I) P391A, (J) Q468A, (K) L491A, (L) E500A, (M) V560A, (N) V561A, (O) F567A, (P) L598A, (Q) E699A, (R) F724A and (S) L759A; see also Table I.

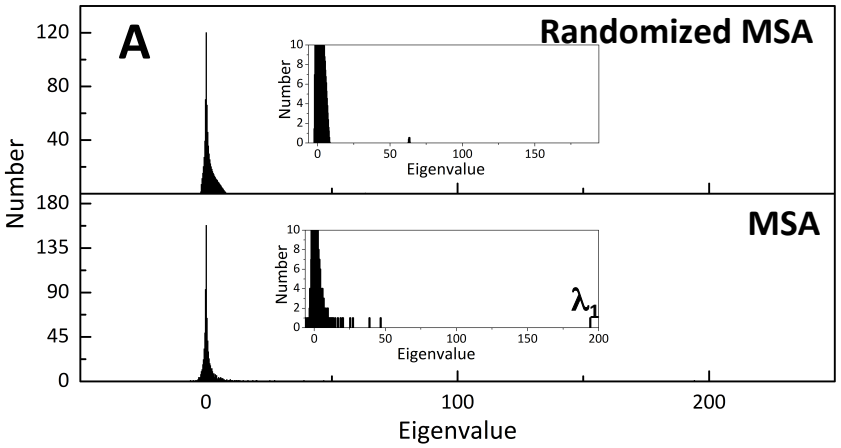
Figure S6. Overlay of wild type and alanine mutant contacts between R76 and DNA and K589 and ATP for non-sector residues. All atom local alignment of R76 and guanine 1546 (left panel), and K589 and ATP (right panel), between wild type (green) and mutant (magenta) proteins shows H-bond disruption at one active site for non-sector residue mutants (A) F78A and (B) K161A, and no disruption for (C) L191A, (D) E204A, (E) P206A, (F) P209A, (G) L341A, (H) I358A and (I) R550A; see also Table I.

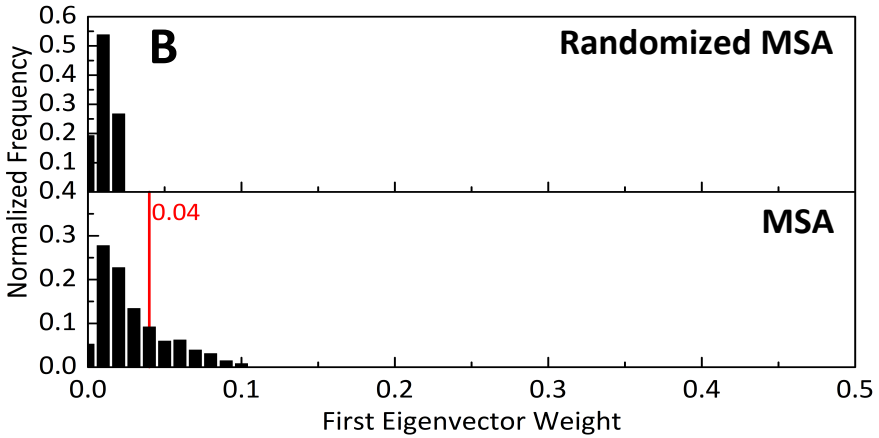
Table S1 (Excel file). Structure Based Alignment of Taq MutS with human MSH2 and MSH6. 160 sector residue positions are noted in bold red. Positions associated with Lynch Syndrome (InSiGHT database) are highlighted in yellow. Residues in the postulated "transmitter" of signals between the DNA binding and ATPase sites (Obmolova et al., 2000), formed by the junction of domains II, III and V and an α-helix in domain IV are marked with an asterisk.

Table S2 (Excel file). MD-calculated domain-wise RMSD for all 31 mutants. Average domain-wise RMSDs for the mismatch binding (I), clamp (IV), and ATPase domains (V) of both subunits in wild type and mutant MD trajectories were calculated over five 3 ns windows of the total 15 ns simulation by global fitting onto the initial reference structure. A difference was considered significant if it varied at least 2σ from the mean (highlgithed in red) of the wild type within the 3 ns window. Label color is keyed to that of MutS domains as shown in Figure 1.

Table S3. Effects of sector (R172A, I553A) and non-sector (Y167A) mutations on ADP-bound forms of MutS. Percent hydrogen bond retention in MD trajectory snapshots between reporter R76 $N\eta^2$ and guanine 1546 backbone oxygen in DNA, and reporter K589 N_ζ and ATP/ADP P^β, in the S1 subunit, as well as the domains destabilized by mutations in S1 and S2 subunits are listed.

Figure S1





C

		sector		
		Yes	No	
conservation ($D_i > 2$)	Yes	38	96	134
	No	122	483	605
		160	579	739

Figure S2

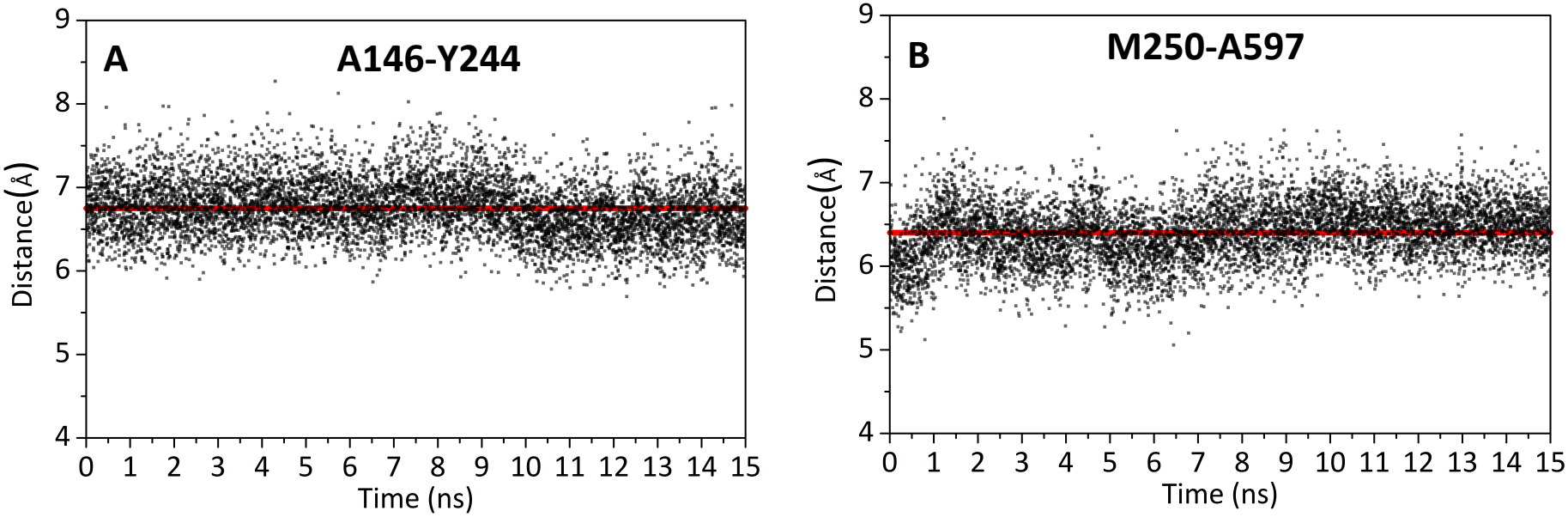


Figure S3

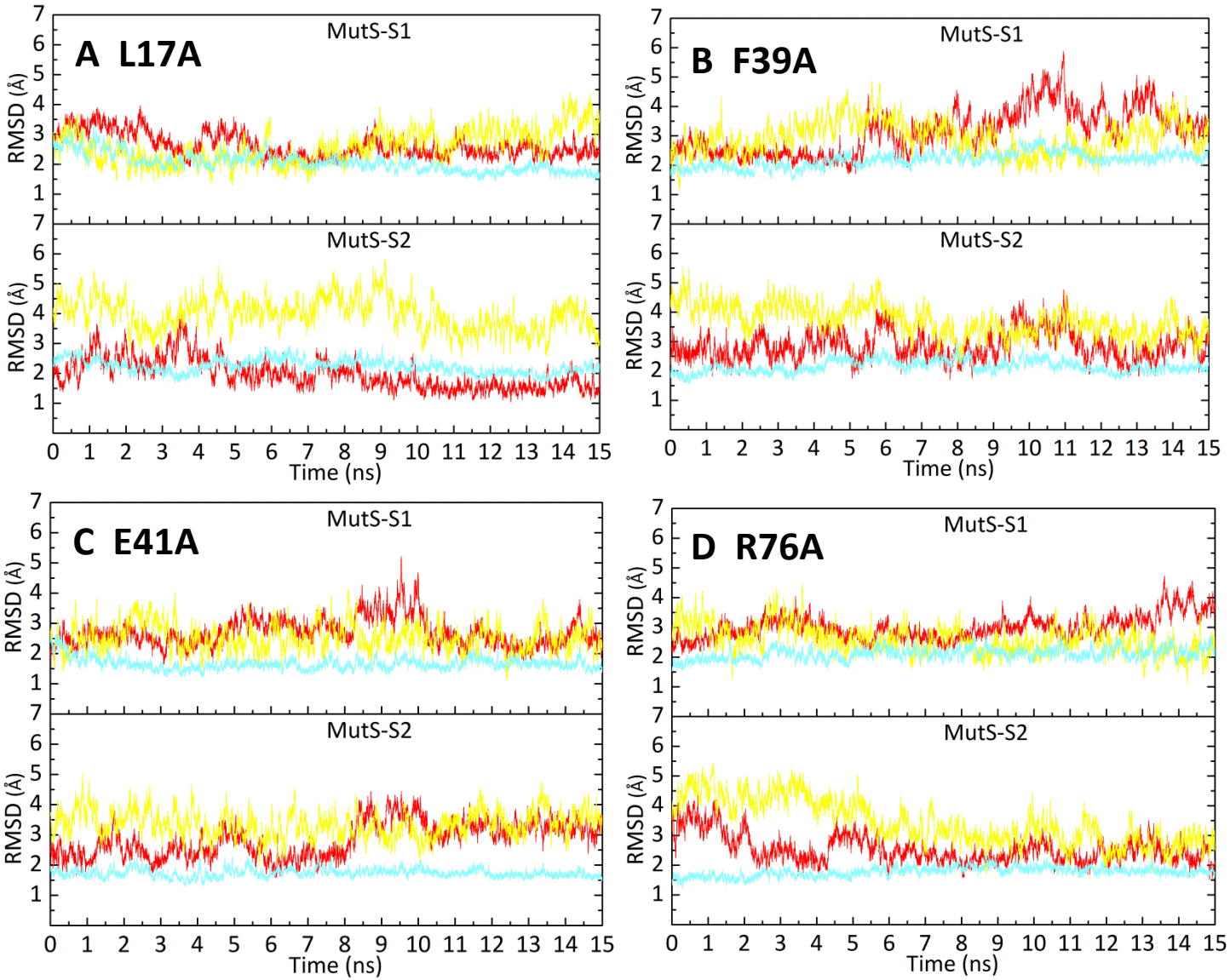


Figure S3

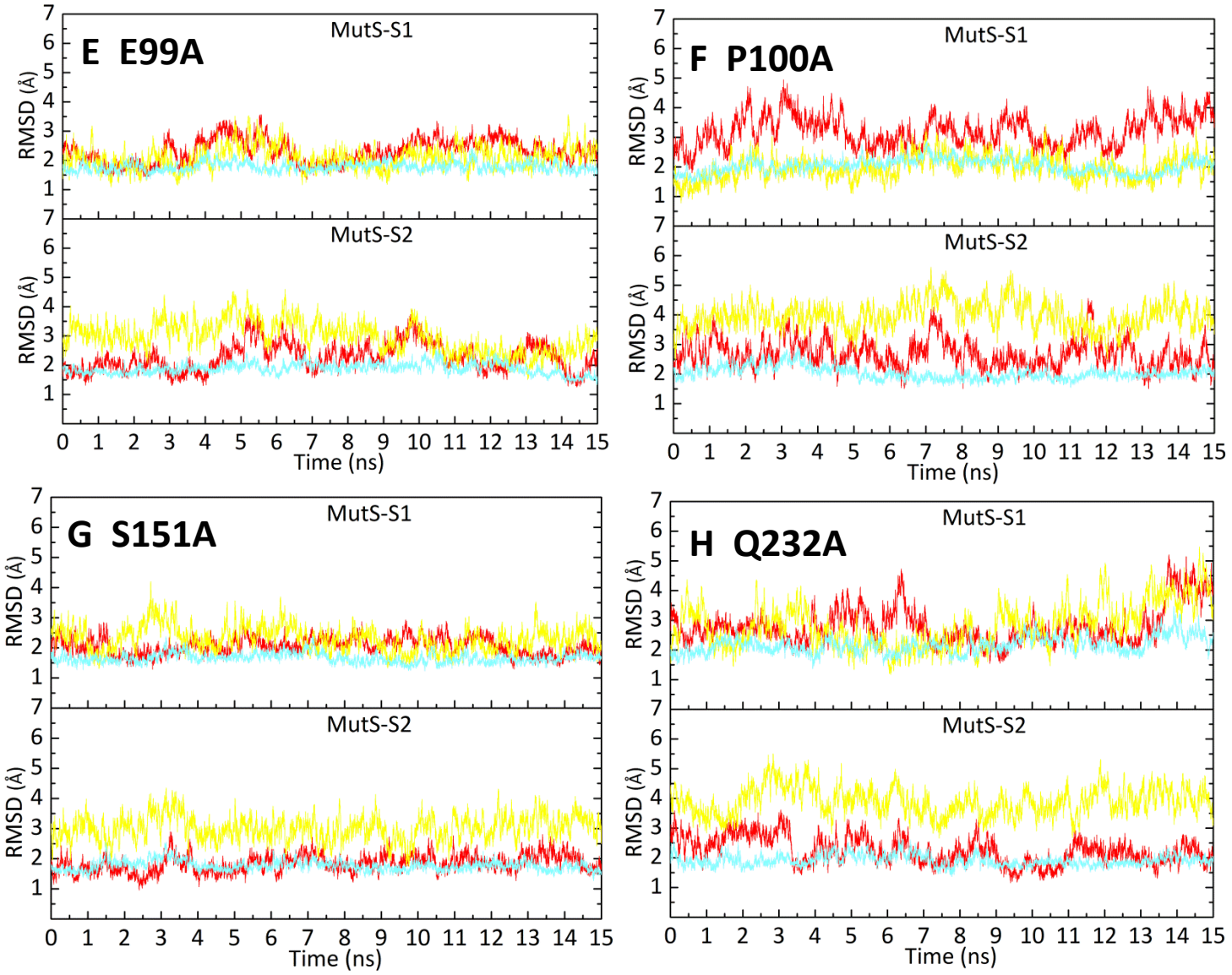

E E99A
MutS-S1
MutS-S2
F P100A
MutS-S1
MutS-S2
G S151A
MutS-S1
MutS-S2
H Q232A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)

Figure S3

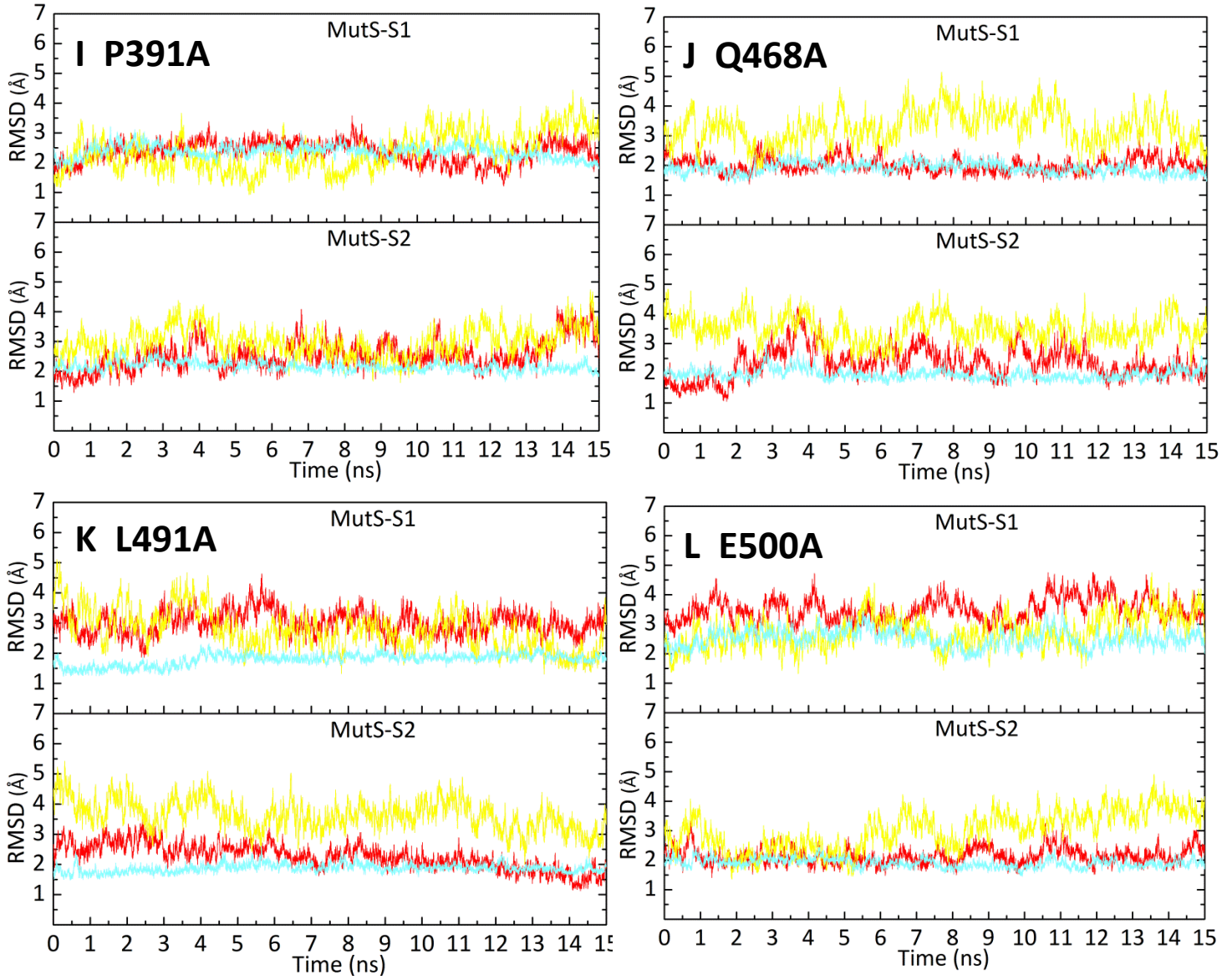


Figure S3

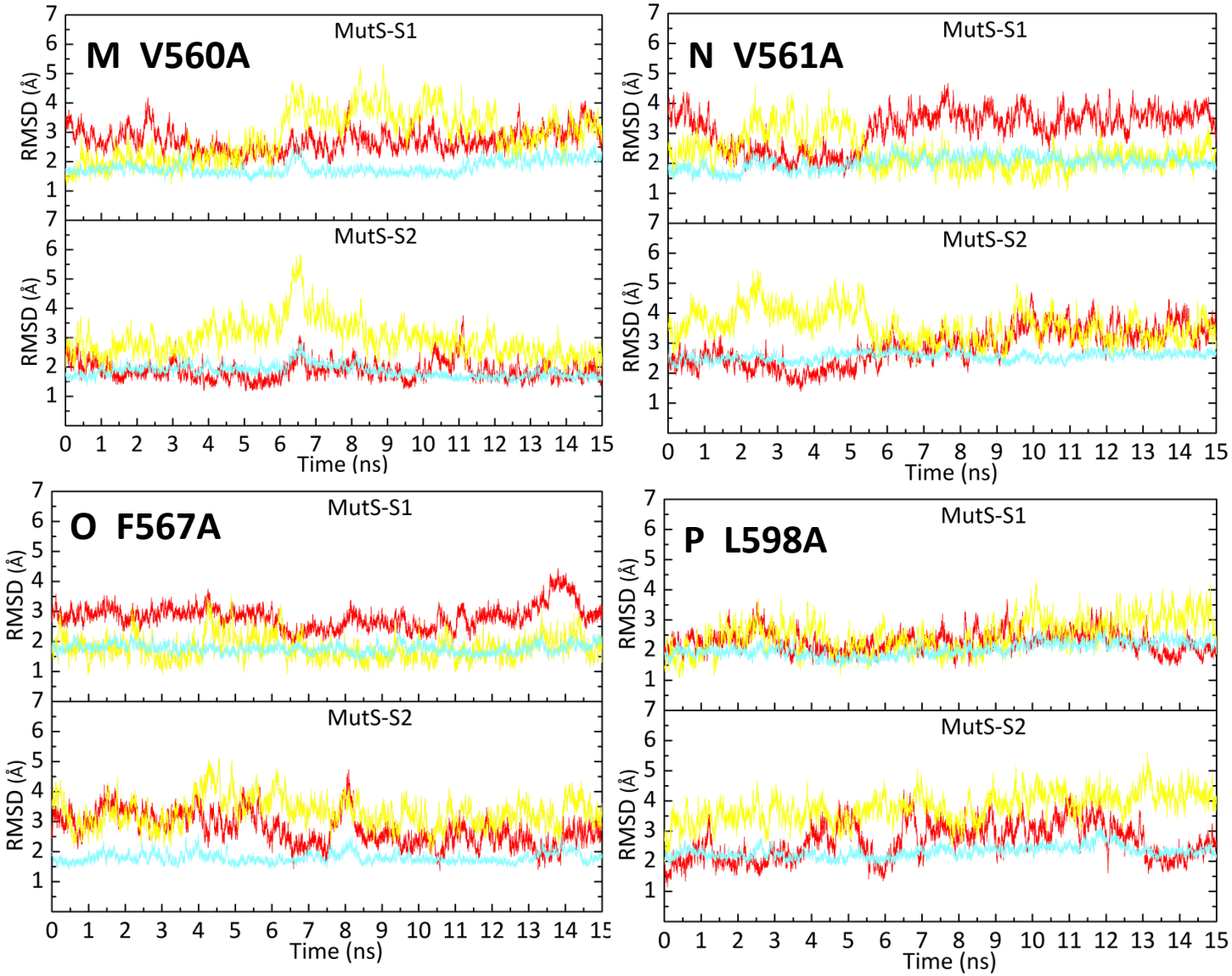

M V560A
MutS-S1
MutS-S2
N V561A
MutS-S1
MutS-S2
O F567A
MutS-S1
MutS-S2
P L598A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)

Figure S3

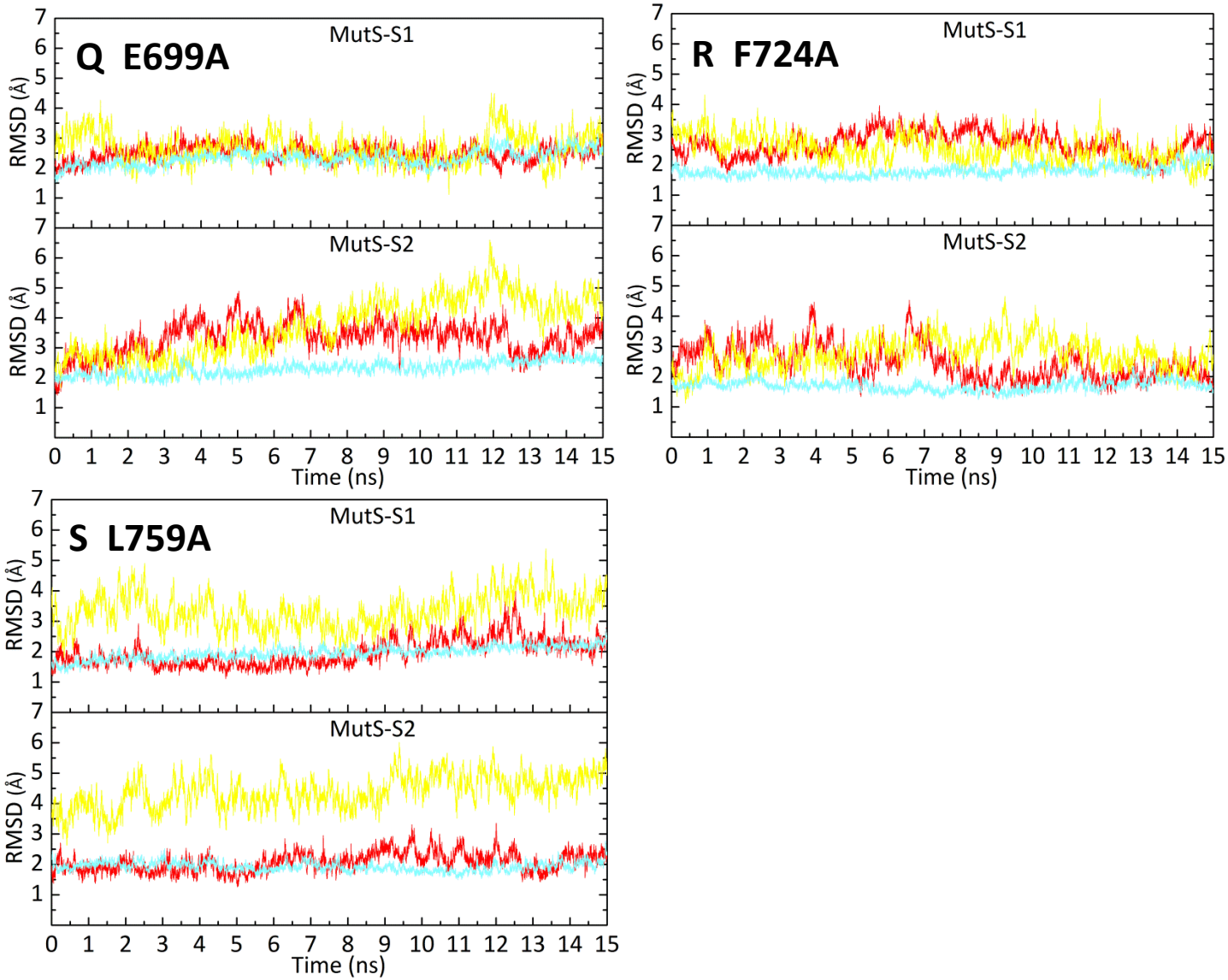

Q E699A
MutS-S1
MutS-S2
R F724A
MutS-S1
MutS-S2
S L759A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)

Figure S3

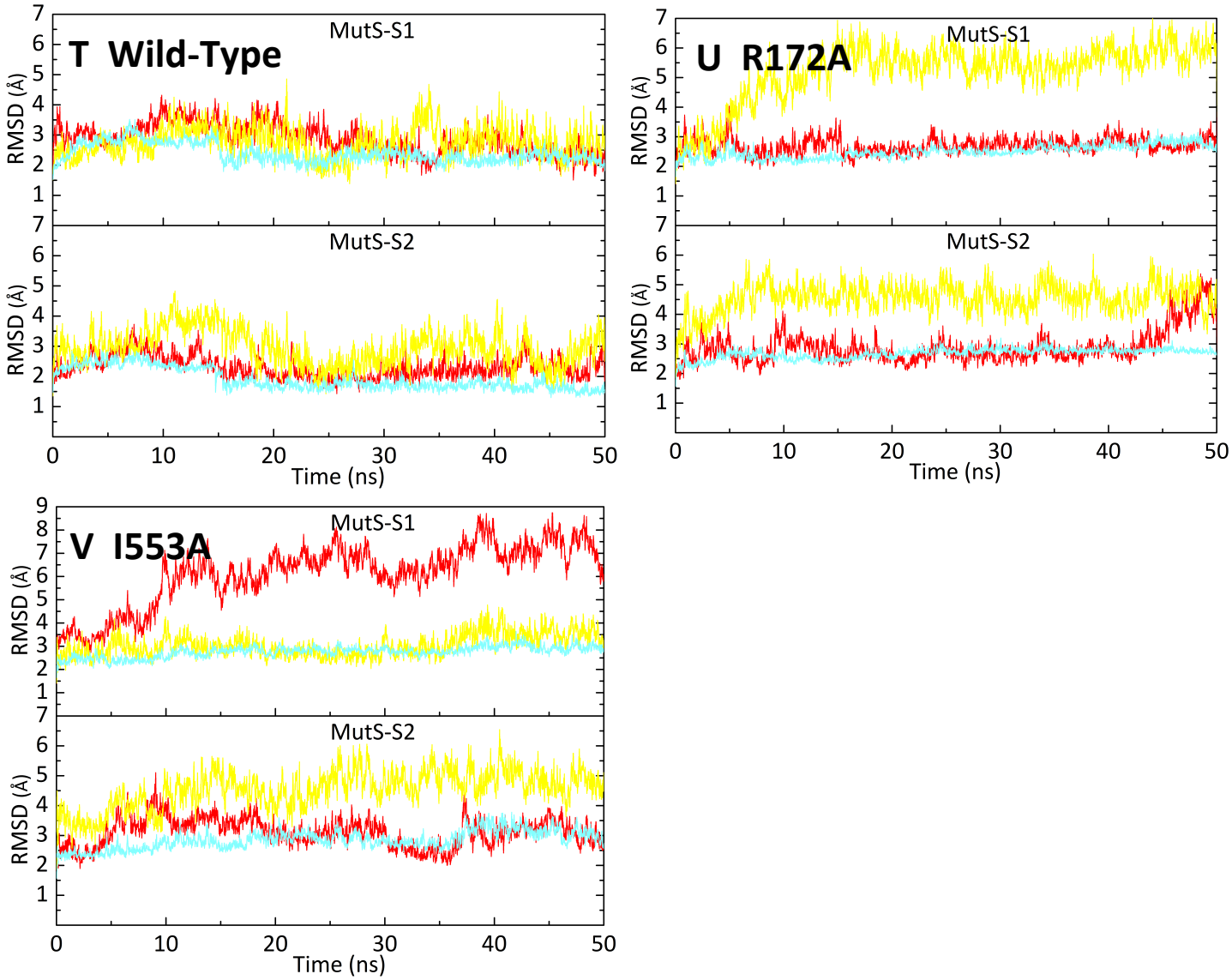

T Wild-Type
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)
U R172A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)
V I553A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)

Figure S4

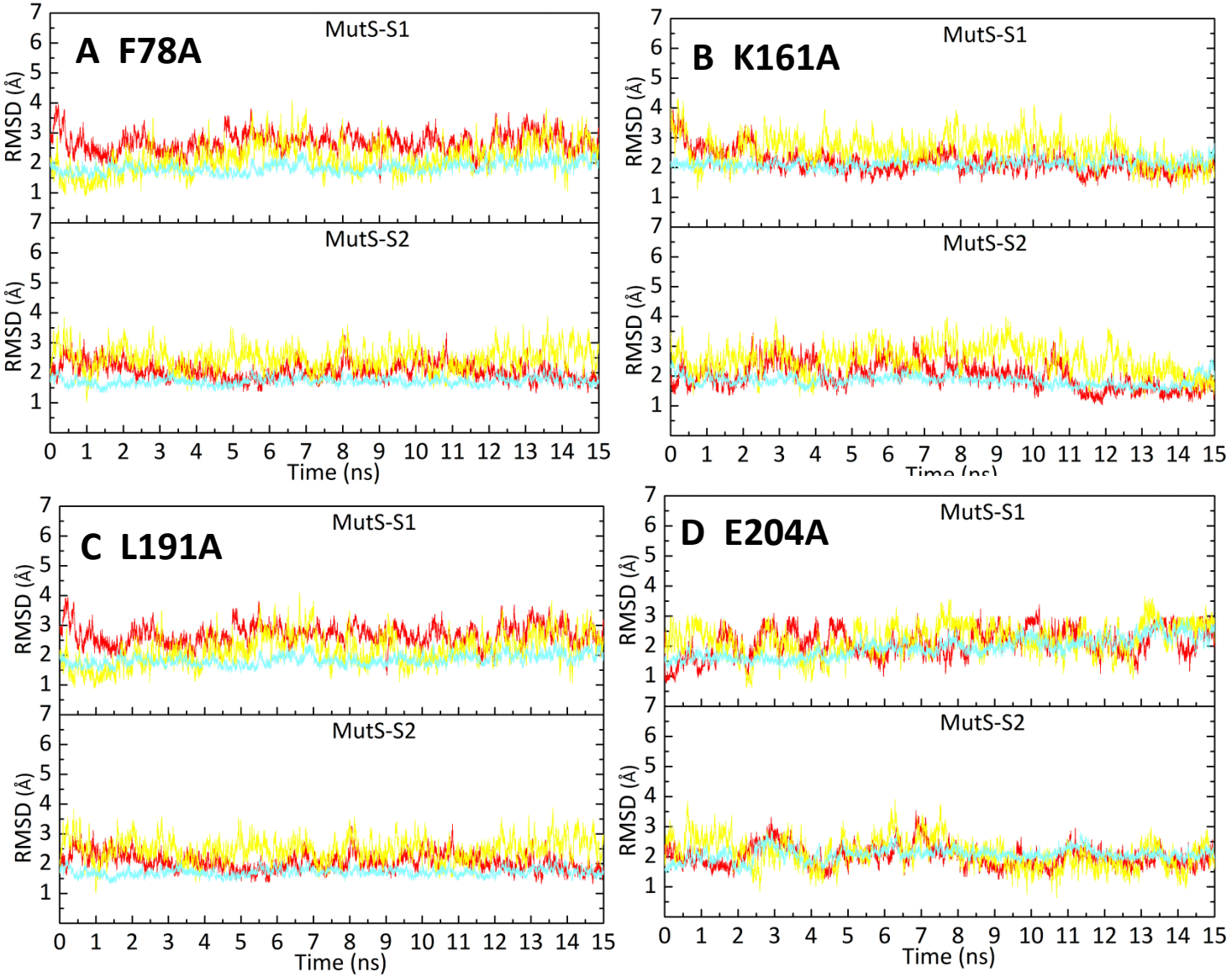

A F78A
MutS-S1
MutS-S2
B K161A
MutS-S1
MutS-S2
C L191A
MutS-S1
MutS-S2
D E204A
MutS-S1
MutS-S2
RMSD (Å)
Time (ns)

Figure S4

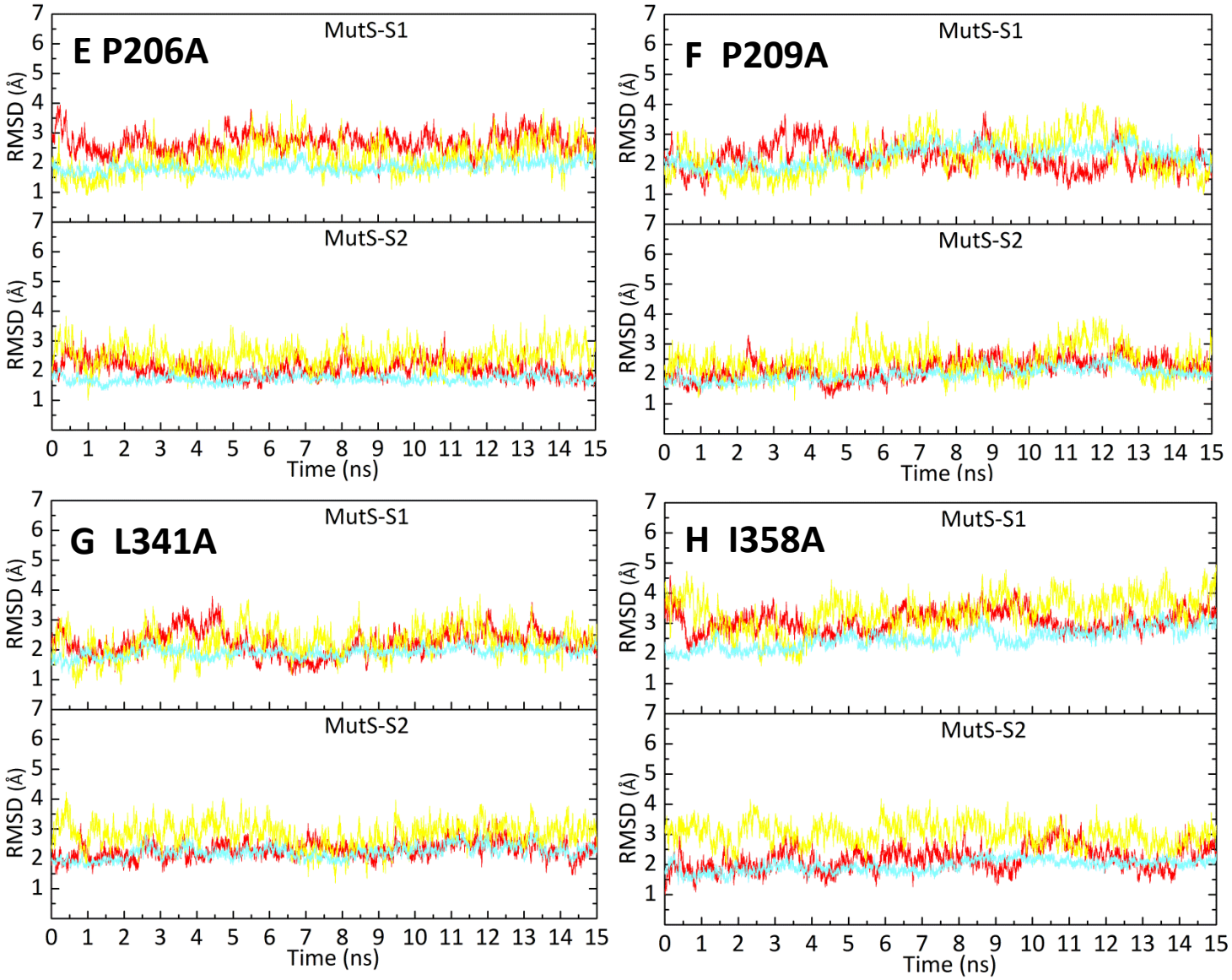


Figure S4

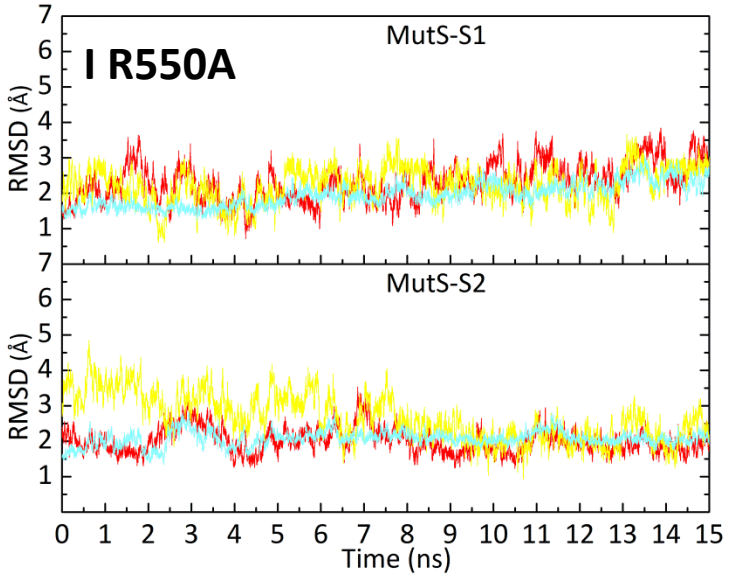


Figure S5

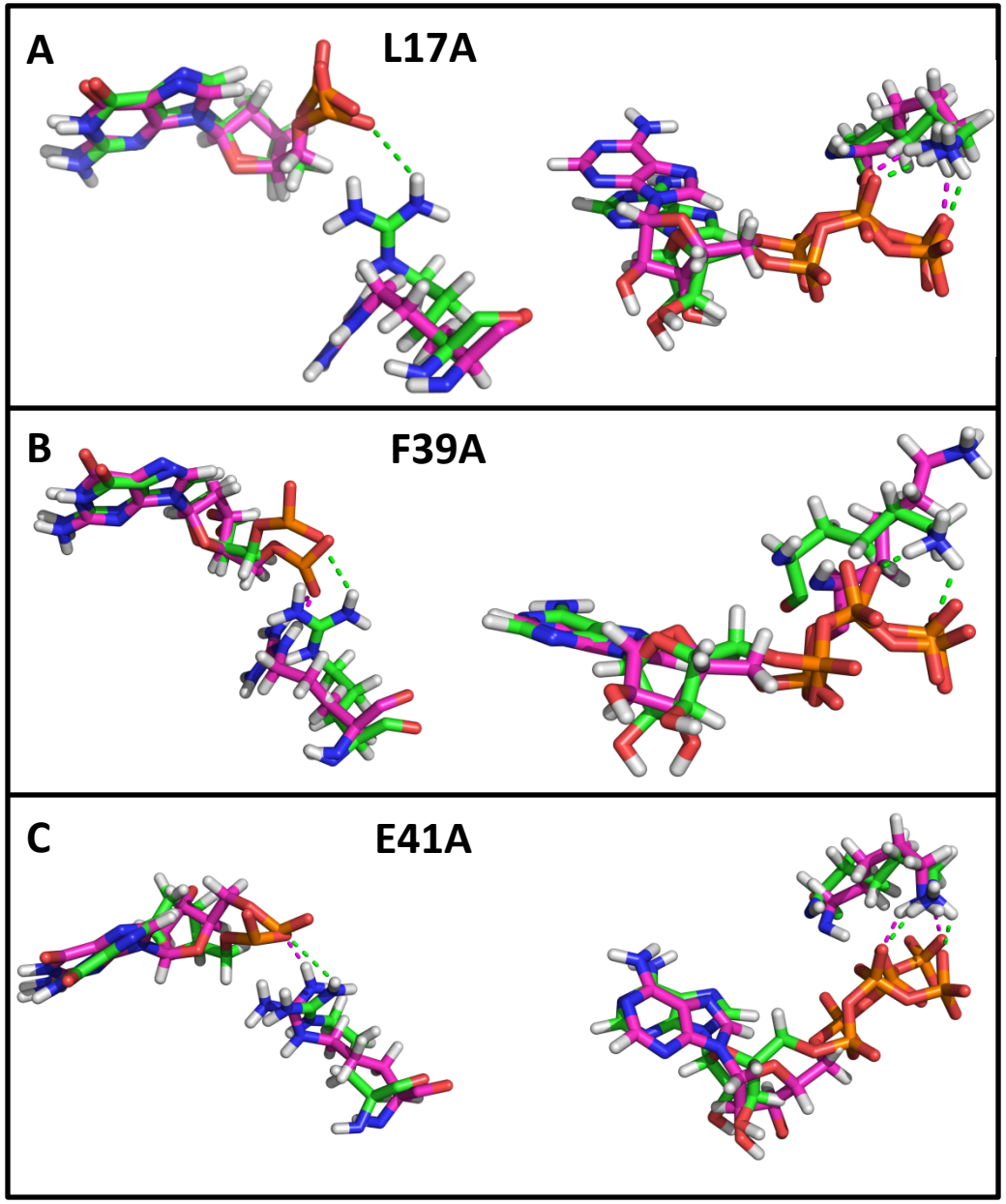


Figure S5

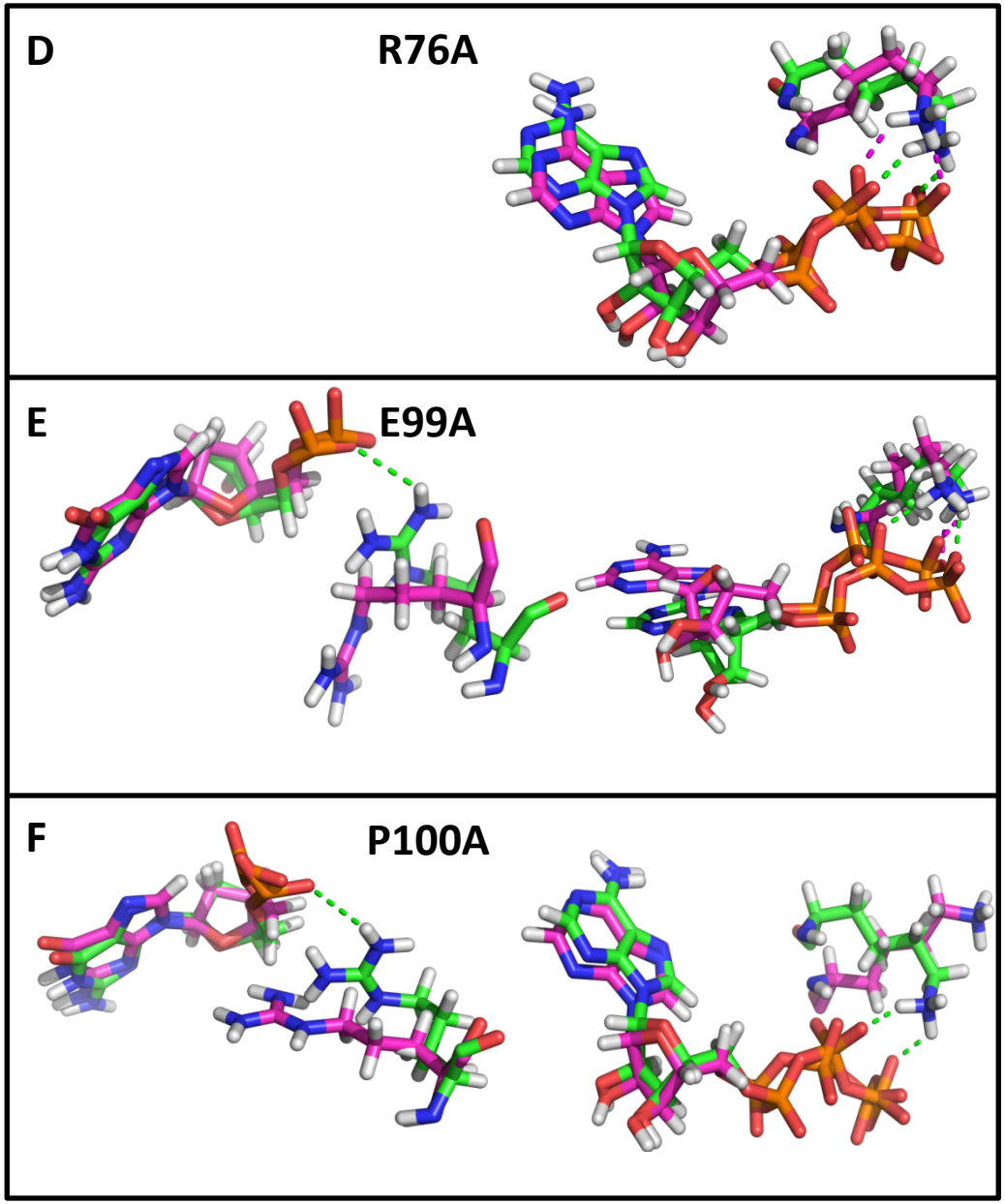


Figure S5

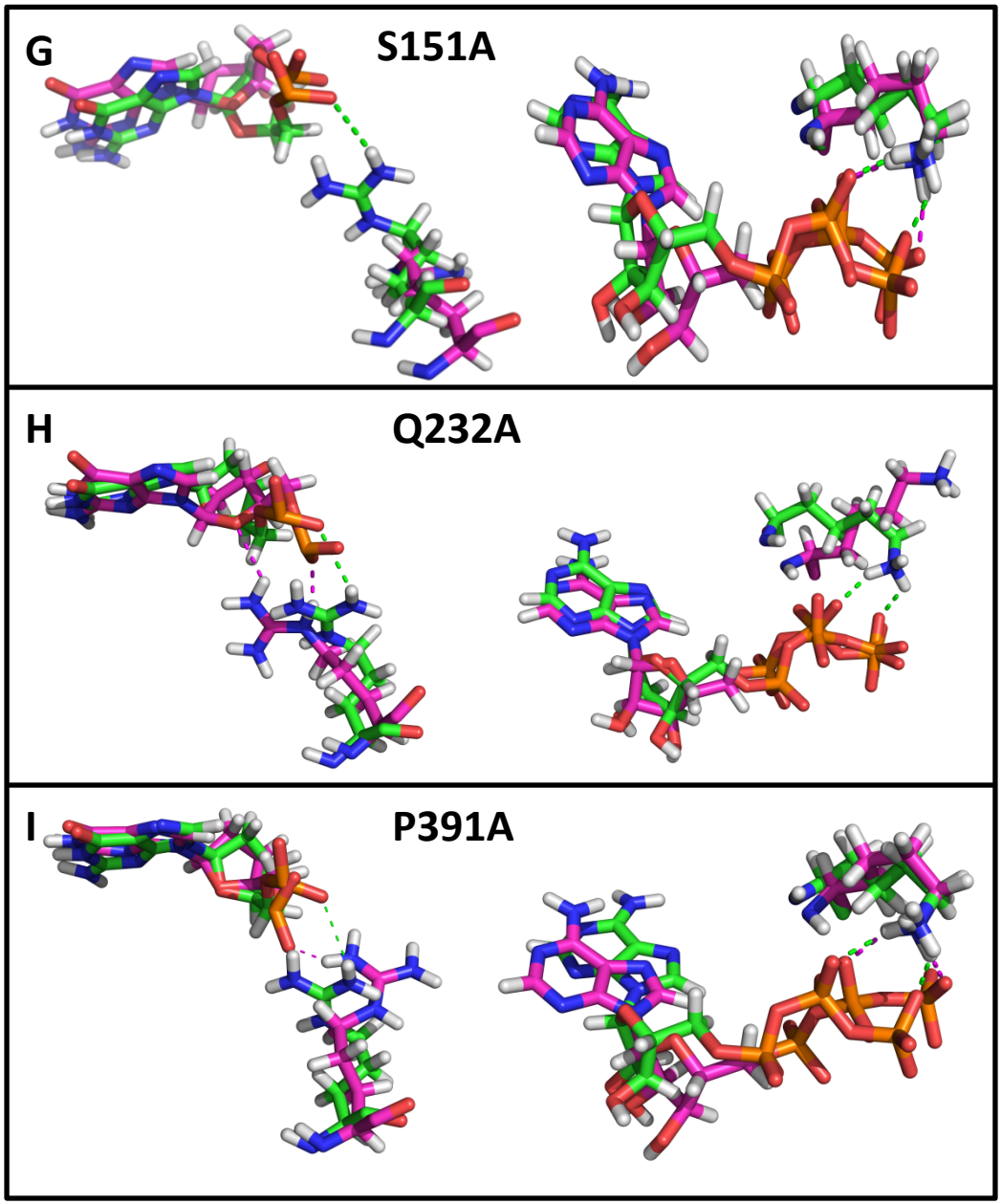


Figure S5

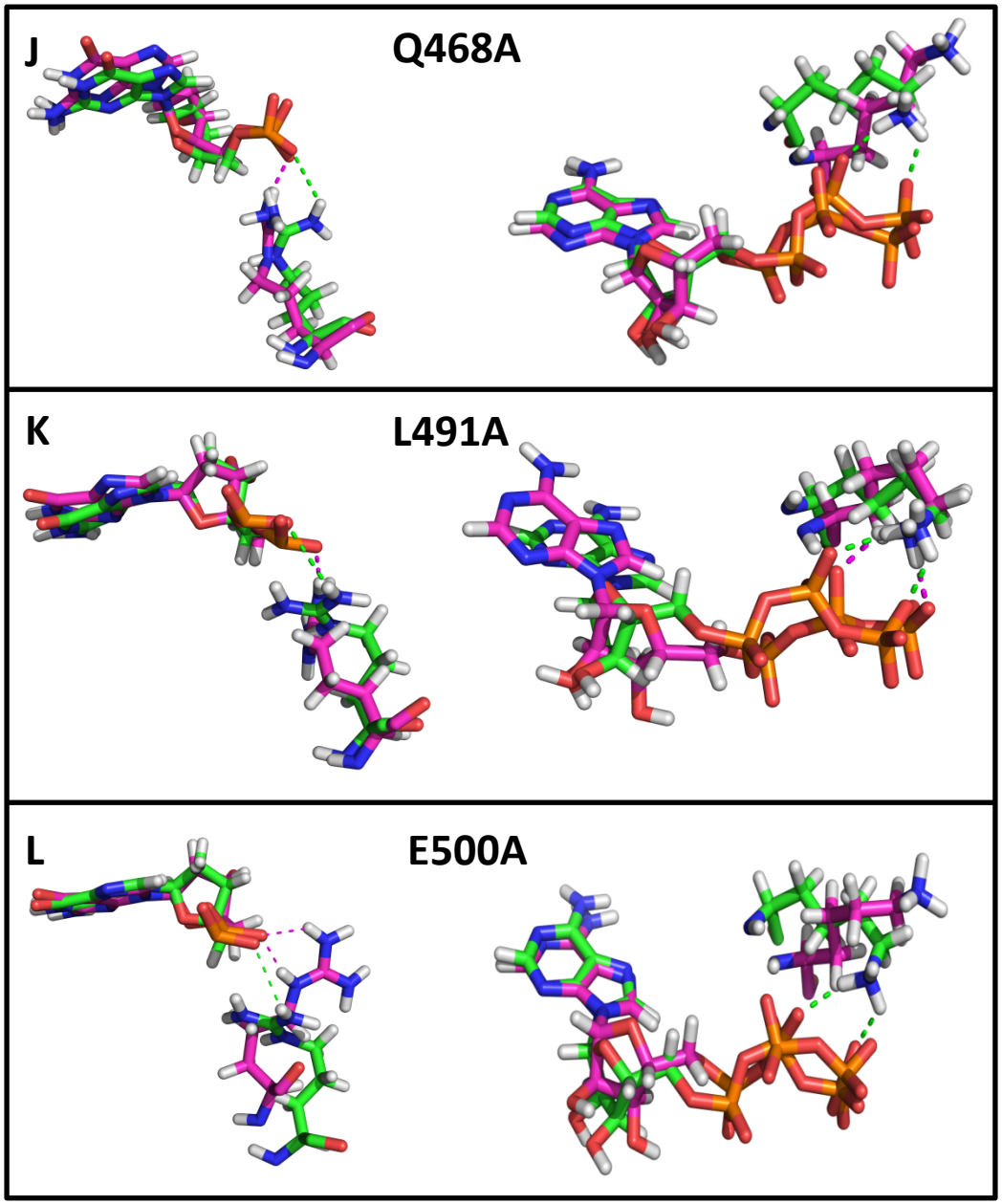


Figure S5

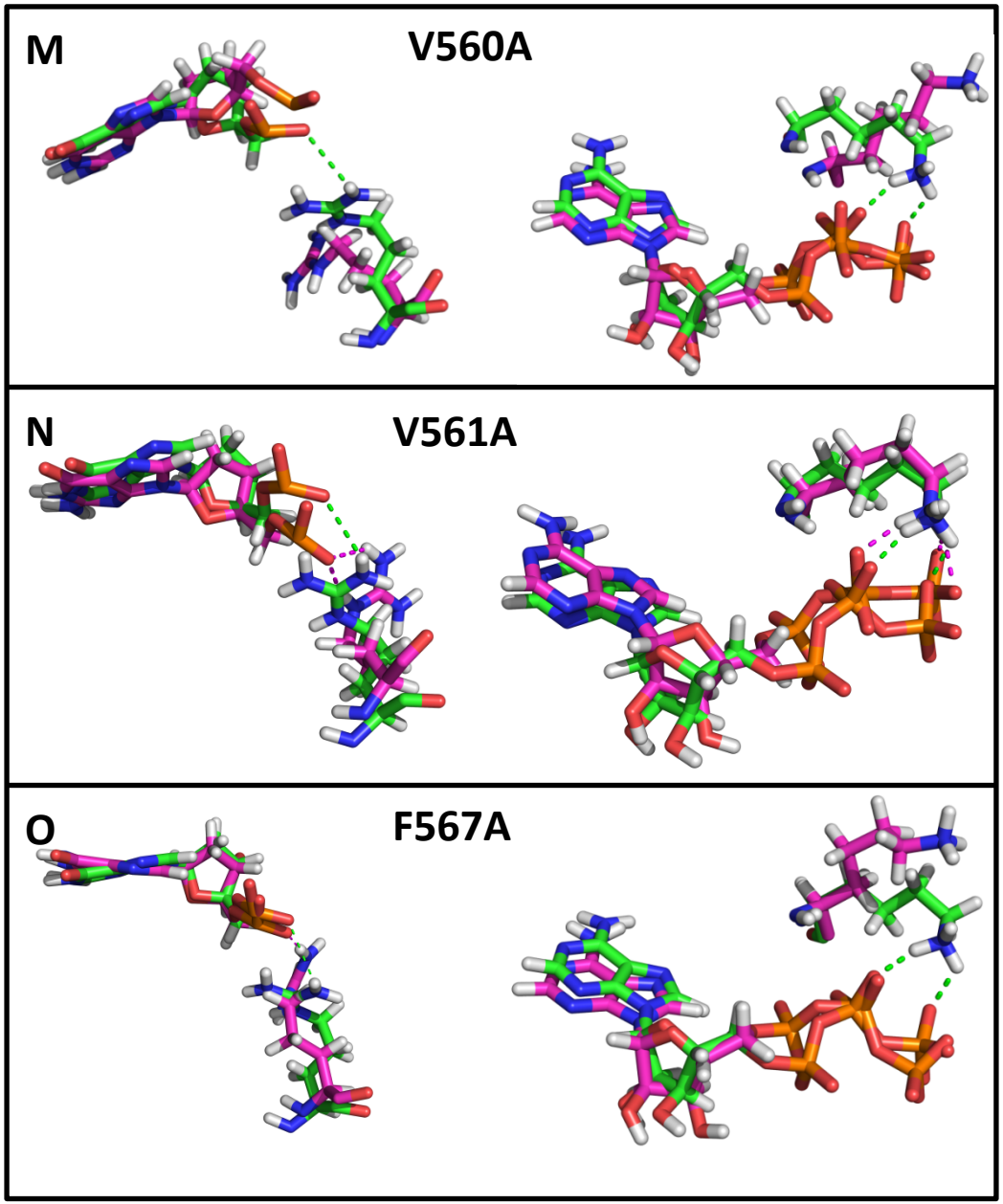


Figure S5

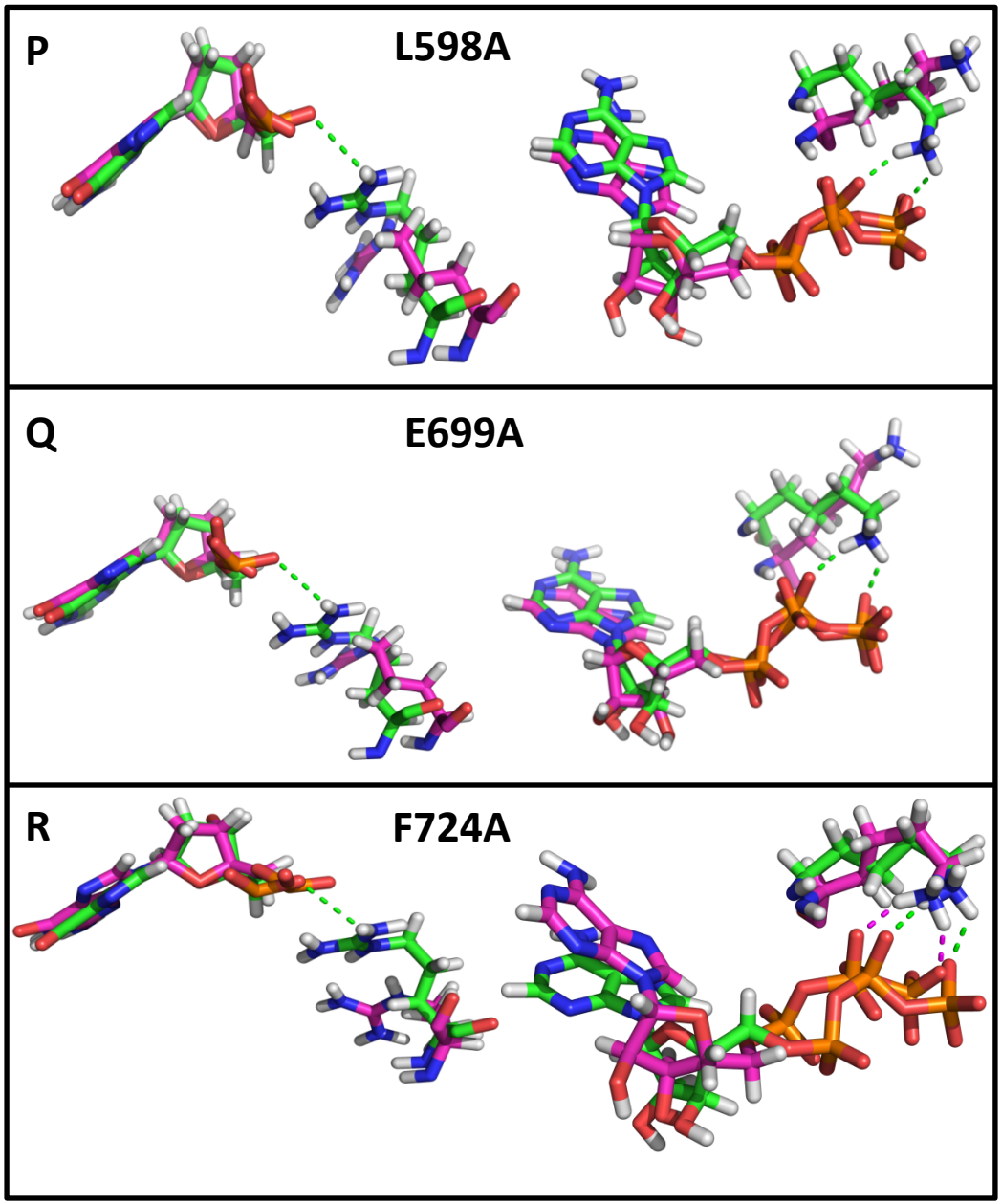


Figure S5

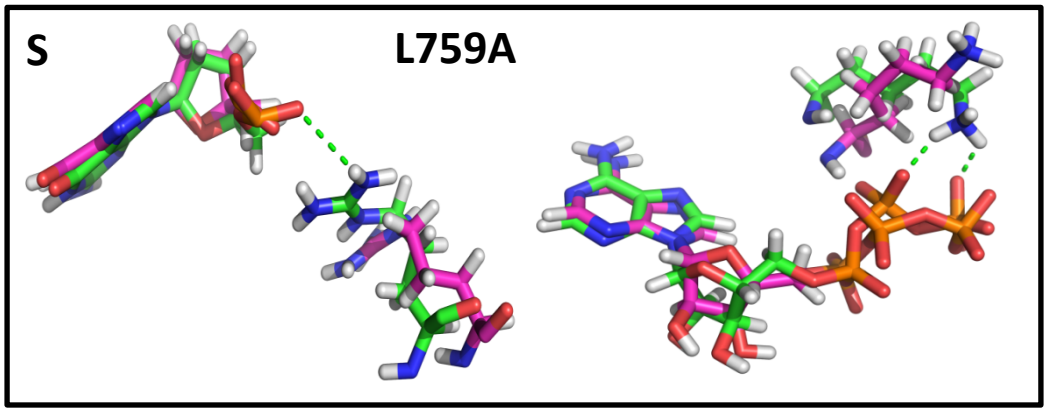


Figure S6

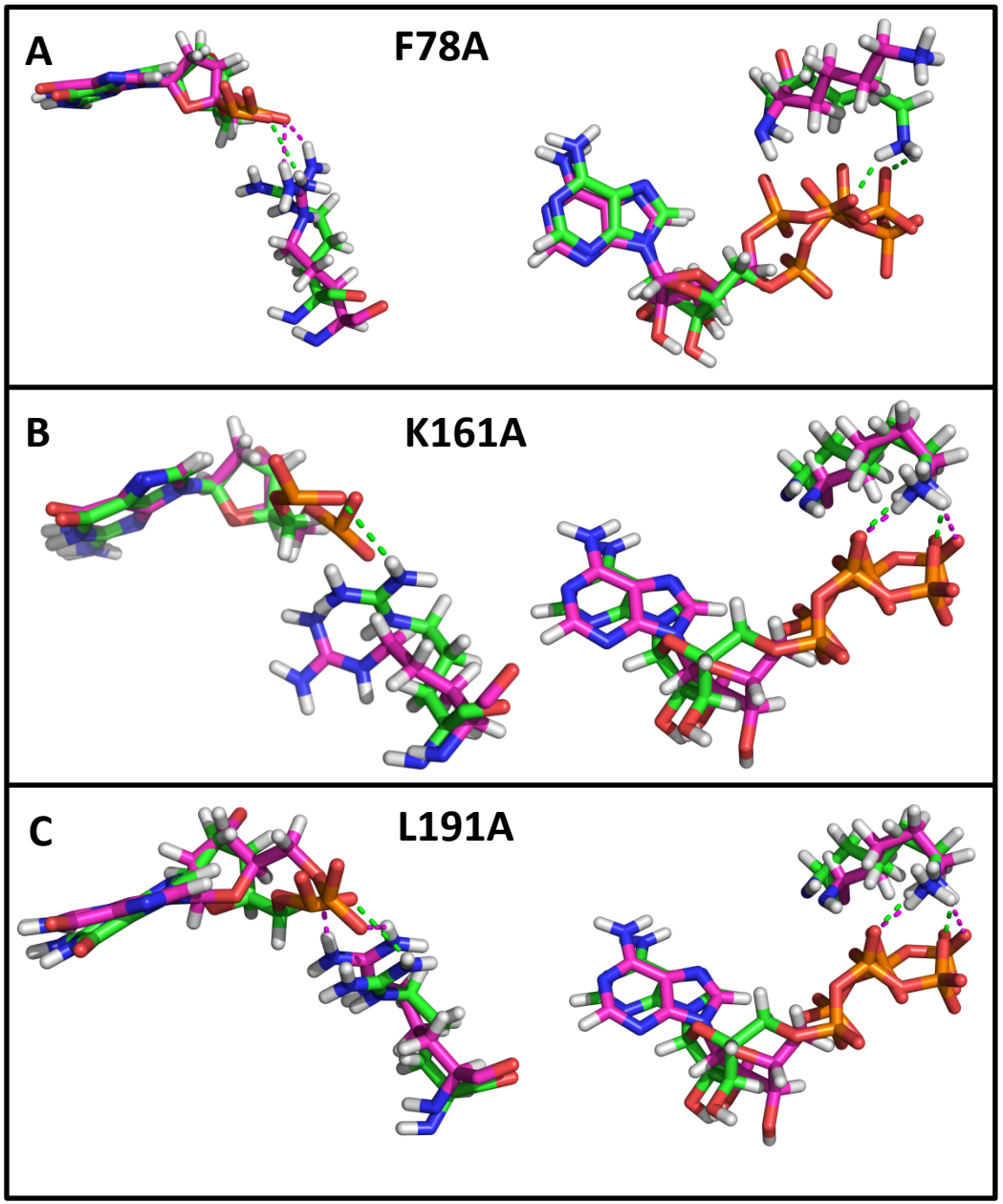


Figure S6

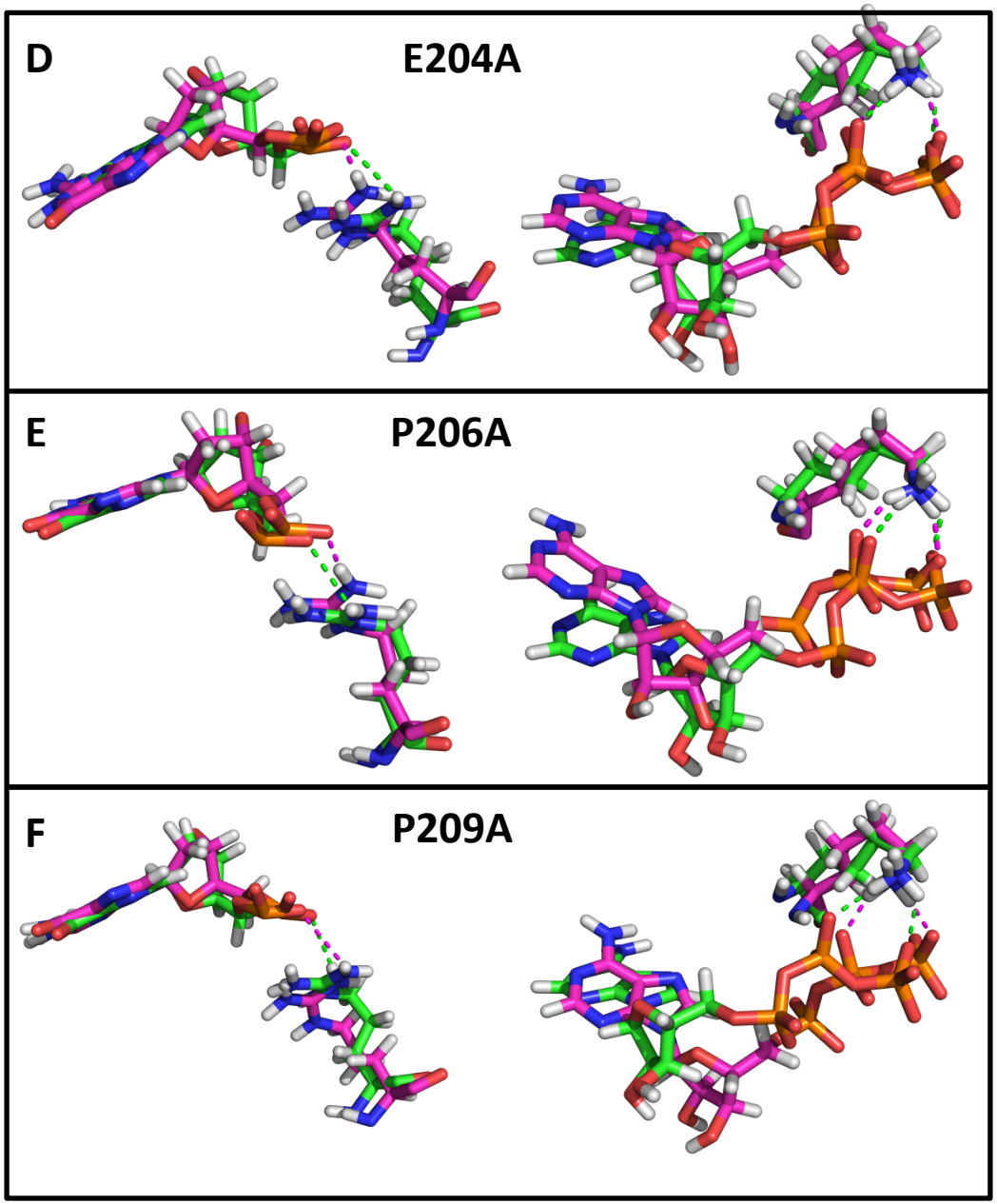
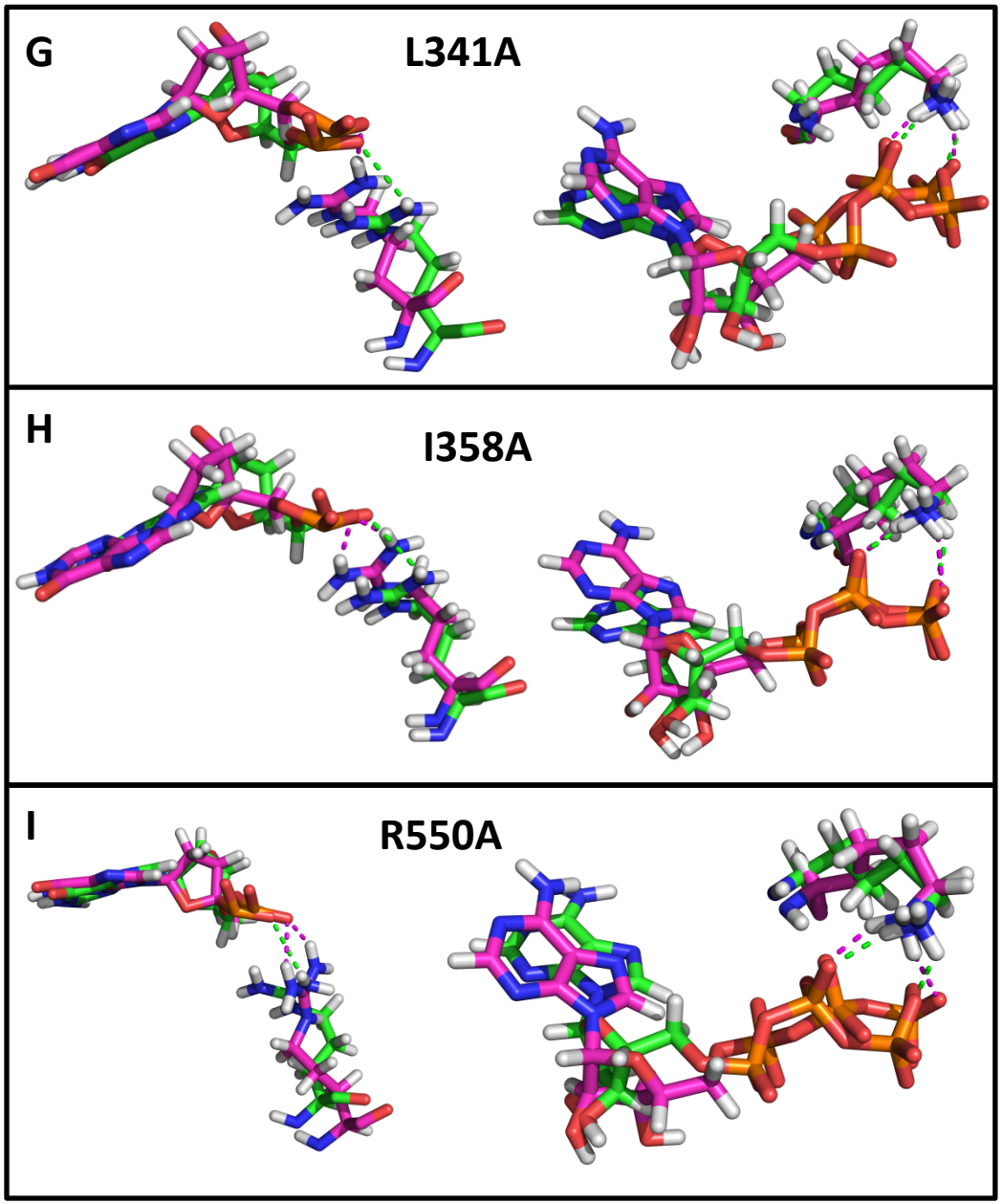


Figure S6



	Position of mutation in MutS	H-bond to DNA	H-bond to ATP or ADP	Domain destabilization in S1, S2 subunits
ATP_{S1}-ATP_{S2}	Wild type	73%	87%	NA
	R172A	11%	0%	S1-IV, S2-IV
	I553A	15%	0%	S1-I, S2-I
	Y167A	55%	65%	—
ADP_{S1}-ATP_{S2}	Wild type	79%	68%	NA
	R172A	14%	0%	S2-I, IV
	I553A	18%	0%	S1-I, S2-I, IV
	Y167A	61%	78%	—
ATP_{S1}-ADP_{S2}	Wild type	67%	72%	NA
	R172A	17%	0%	S1-I, IV, S2-IV
	I553A	12%	0%	S1-IV, S2-IV
	Y167A	38%	49%	S1-I, S2-IV
ADP_{S1}-ADP_{S2}	Wild type	56%	98%	NA
	R172A	3%	91%	S1-I, S2-IV
	I553A	7%	88%	S1-I, S2 IV
	Y167A	69%	84%	—

TABLE S3. Effects of sector (R172A, I553A) and non-sector (Y167A) mutations on ADP-bound forms of MutS. Percent hydrogen bond retention in MD trajectory snapshots between reporter R76 $N\eta^2$ and guanine 1546 backbone oxygen in DNA, and reporter K589 N_ζ and ATP/ADP P^β, in the S1 subunit, as well as the domains destabilized by mutations in S1 and S2 subunits are listed.