Thumb Inhibitor Binding Eliminates Functionally Important Dynamics in the Hepatitis C Virus RNA Polymerase

Brittny Davis and Ian F. Thorpe*

Supporting Information

Supporting Tables

Table SI. Comparison between amino acid composition of structural motifs in NS5B and other RdRps

Domains	**PV**	**CVB3**	**HRV16**	**FMDV**	**NS5B**
Fingers	1-68, 96-190, 269-286	1-68, 96-190, 271-287	1-68, 96-190, 269-285	1-66, 97-195, 280-296	1-188, 224-308
Palm	69-95, 191-268, 287-380	69-95, 191-270, 288-381	69-95, 191-268, 286-279	67-96, 196-279, 297-392	189-223, 309-270
Thumb	381-461	382-462	380-460	393-470	371-562
Motifs					
A	229-240	229-240	229-240	236-247	216-227
B	293-312	294-313	292-311	303-322	287-306
C	322-335	323-336	321-334	332-345	312-325
D	338-362	339-363	337-361	348-373	332-353
E	363-380	364-381	362-379	374-392	354-372
F	153-178	153-178	153-178	158-183	135-162
G	113-120	113-120	113-120	114-121	95-99
Functional Regions					
I	107-112	107-112	107-112	108-113	91-94
II	184-200	184-200	184-200	189-205	168-184
III	405-420	406-421	404-419	416--430	401-414

PID to NS5B (%)	10.38	8.80	9.15	11.44	-

PV: Poliovirus, CVB3: coxsackievirus B3, HRV16: Human rhinovirus serotype 16, FMDV: Foot-and-mouth disease virus.
Sequence identity was computed with the Sequences Identities and Similarities Server (http://imed.med.ucm.es/Tools/sias.html). The Blosum 62 scoring matrix was used with a creating gap penalty of 10 and an extending gap penalty of 0.2. Percent Identity (PID) was calculated using the equation: PID=100*((Identical position) / (Length of the alignment))

Table SII. Backbone RMSD (Å) between PDB coordinates and MD average structures of NS5B, computed via aligning the palm domain

	1YUY	1YV2	2WHO	Apo
1YV2				
Palm	0.43	-	-	-
Fingers + Thumb	1.85			
Thumb	2.59			
2WHO				
Palm	0.51	0.60	-	-
Fingers + Thumb	1.23	1.56		
Thumb	1.69	2.12		
Free				
Palm	0.77	0.78	0.66	-
Fingers + Thumb	2.42	3.37	2.72	
Thumb	3.29	4.78	3.83	
Ligand Bound				
Palm	1.17	1.13	1.06	0.92
Fingers + Thumb	4.54	5.55	5.10	3.24
Thumb	6.31	7.88	7.22	4.36

Alignment was carried out by superimposing residues 181-222 and 281-366 which constitute the palm domain of the NS5B. For the purposes of the RMSD calculation, the fingers domain corresponds to residues 1-20, 41-180 and 222-280 while the thumb domain corresponds to residues 21-40 and 367-531.

Table SIII. Width of RNA duplex channel in NS5B

	Distance $\pm \sigma$ (Å)
2WHO coordinates	16.2
apo simulations	
all	14.7 $\pm$1.4
100-220 ns	14.2 $\pm$1.4
221-300 ns	15.5 $\pm$1.7

301-400 ns	14.5 ±0.9
ligand bound simulations	
all	12.3 ±0.6
100-240 ns	12.5 ±0.6
241-400 ns	12.1 ±0.6

The distance between Cα atoms of S95 and N406 in NS5B was used as a surrogate for the channel width. σ = one standard deviation.

Table SIV. Width of RNA template channel in NS5B

	Distance ± σ (Å)
2WHO coordinates	13.9
apo simulations	
all	17.2 ±1.5
100-220 ns	16.8 ±1.7
221-300 ns	17.3 ±1.2
301-400 ns	17.6 ±1.4
ligand bound simulations	
all	12.2 ±0.8
100-240 ns	12.5 ±0.9
241-400 ns	12.0 ±0.7

The distance between Cα atoms of S95 and C14 in NS5B was used as a surrogate for the channel width. σ = one standard deviation.

Table SVI. Blocks of residues displaying high levels of motional correlation.

Block	Mutation Sites	Block Members and Location
a	19, 147, 155, 390	18-42, 142-159, 384-420: rear of thumb domain

b	50, 55, 138, 162	1-10, 42-61, 96-126, 133-142, 157-180, 222-232, 256-291, 341-346: rear of fingers domain
c	71, 81, 85, 177	60-93, 105-130, 164-186, 230-262, 283-290: front and left side of fingers domain
d	95, 121	79-139, 160-181, 240-290: top of fingers domain
e	338, 355	58-64, 213-222, 298-311, 320-363: bottom of palm
f	432, 441, 442, 445	405-416, 421-472, 500-527: top of thumb
g	316	187-204, 218-222, 291-296, 312-322: middle and center of palm, interface with thumb domain
h	365, 368	195-204, 217-220, 357-371, 383-387: middle and center of palm, rear of palm extending to tip
i	494	21-23, 28-33, 234-242, 482-503: rear of thumb + small section on left side of fingers

Figure Captions

Figure S1. Comparison of amino acid composition and secondary structural motifs for NS5B and group I RdRps. PV: Poliovirus, CVB3: coxsackievirus B3, HRV16: Human rhinovirus serotype 16, FMDV: Foot-and-mouth disease virus

Figure S2. Comparison of normalized B-factors for Cα atoms in the 2WHO crystallographic coordinates to those in MD simulations. The B-factors were normalized by dividing all of the values obtained for a given protein by the maximum B-factor value for that protein. Thus, the normalized B-factor scale for each protein lies between 1 and 0. The normalization procedure partially accounts for the fact that the crystal structure and MD simulations were prepared at different temperatures (100 and 300 K respectively). The plots show that the relative flexibility of specific structural elements along the peptide sequence is qualitatively consistent between 2WHO and the MD simulations. The exception to this observation is motif F located at the interface between the fingers and thumb domains that plays a role in binding RNA template. This region displays significantly enhanced flexibility relative to the remainder of the enzyme in the MD simulations. The free enzyme displays more flexibility than the ligand-bound enzyme due to the change in orientation of the thumb domain that constrains this motif in the ligand-bound enzyme.

Figure S3. Width of RNA duplex channel during NS5B molecular dynamics simulations. The distance between Cα atoms of S95 and N406 in NS5B was used as a surrogate for the channel width. The red line corresponds to the ligand-bound enzyme while the black line displays the result obtained for the free

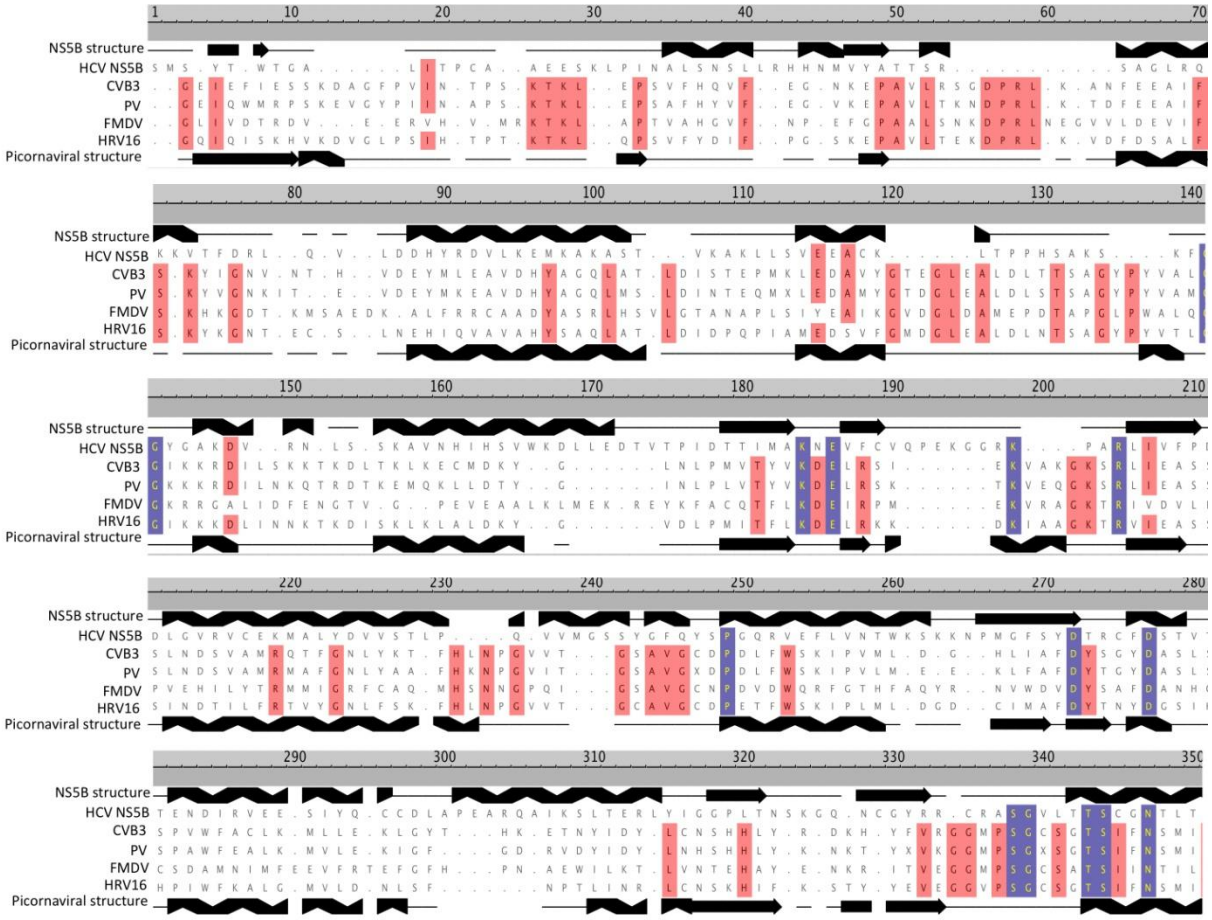
enzyme. The gray box indicates the equilibration period of the trajectory. The circled region corresponds to the portion of the free trajectory that displays elevated SASA for residues lining the RNA template binding channel. Comparison with Figure 4 in the main text reveals that the widths of the RNA template and RNA duplex channels increase in concert with one another.

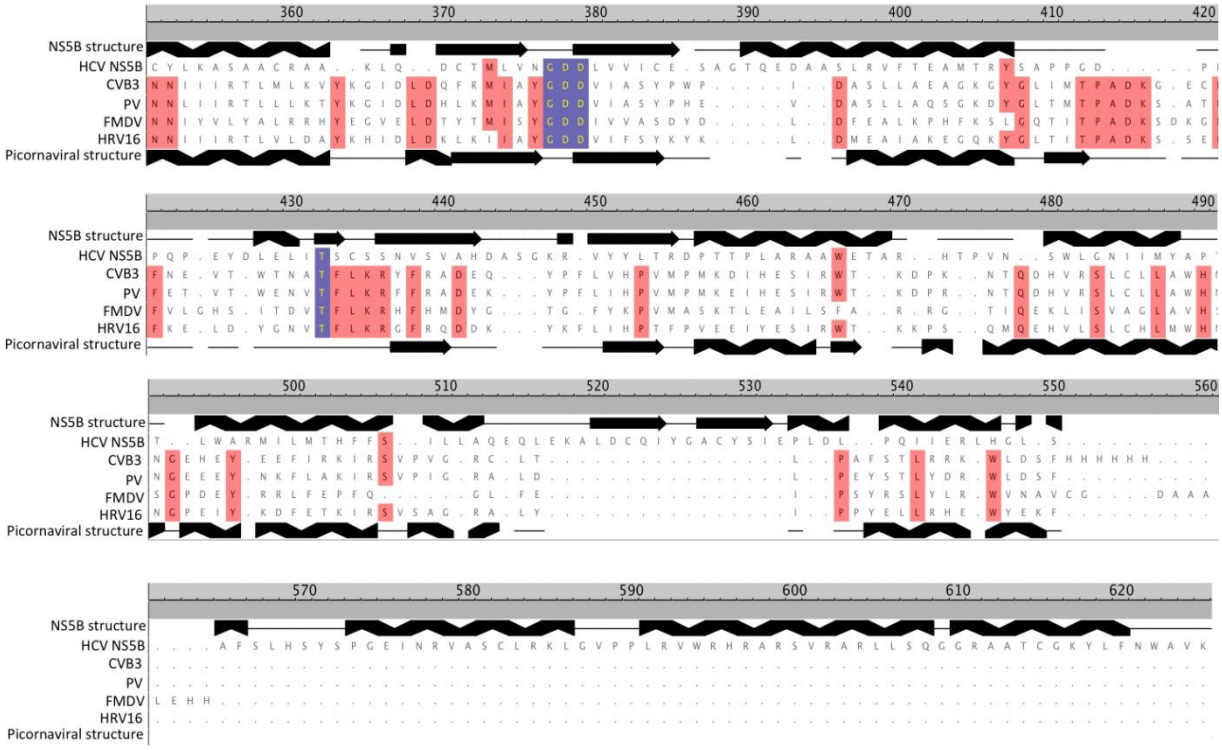
Figure S4. Covariance matrices constructed from discriminating modes identified from the free enzyme simulations. a) and b) were generated by filtering out only PC mode 1 and 3 from the free enzyme simulations and reconstructing the covariance matrix as described in Methods. Notation in the figure is as shown in Figure 6 of the main text.

Figure S5. Blocks of residues displaying high levels of motional correlation in ligand-free NS5B simulations. These dynamic units are roughly consistent with the known structural motifs of NS5B. However, the simulations allow these structural motifs to be subdivided into more fine grained motional units, offering more detailed information about how the motion of different regions of NS5B is connected. (a) Space filling representations of regions a, b, g and h given in Table SVI, with each region drawn in a separate color. (b) Space filling representations of regions c, d, e, f and i given in Table SVI. For each row the left panel displays the same view given in the left panel of Figure 2 in the main text while the right panel displays the view given in the upper panel of Figure 5 in the main text. Multiple colors for the same region indicate residues participate in more than one dynamic domain.

Movies of principal components that discriminate between free and ligand-bound trajectories are available online.

Figure S1.


NS5B structure
HCV NS5B
CVB3
PV
FMDV
HRV16
Picornaviral structure



This alignment was performed using the ClustalW server (http://www.ch.embnet.org/software/ClustalW.html). The Blosum Scoring matrix was employed with an open gap penalty of 10, end gap penalty of 10, extending gap penalty of 0.05 and a separation gap penalty of 0.05.

Figure S2.

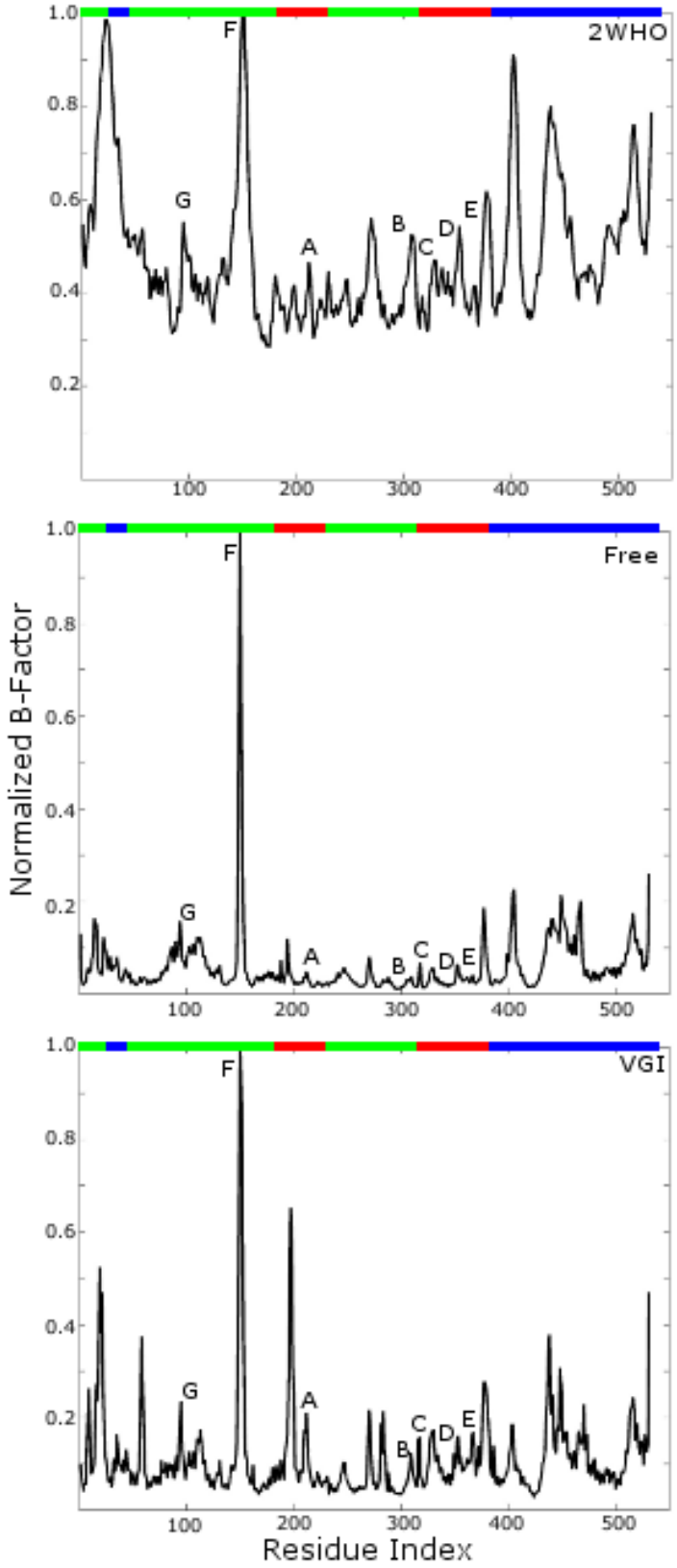

2WHO
Free
VGI
F
G
A
B
C
D
E
1.0
0.8
0.6
0.4
0.2
100
200
300
400
500
Normalized B-Factor
Residue Index

Figure S3.

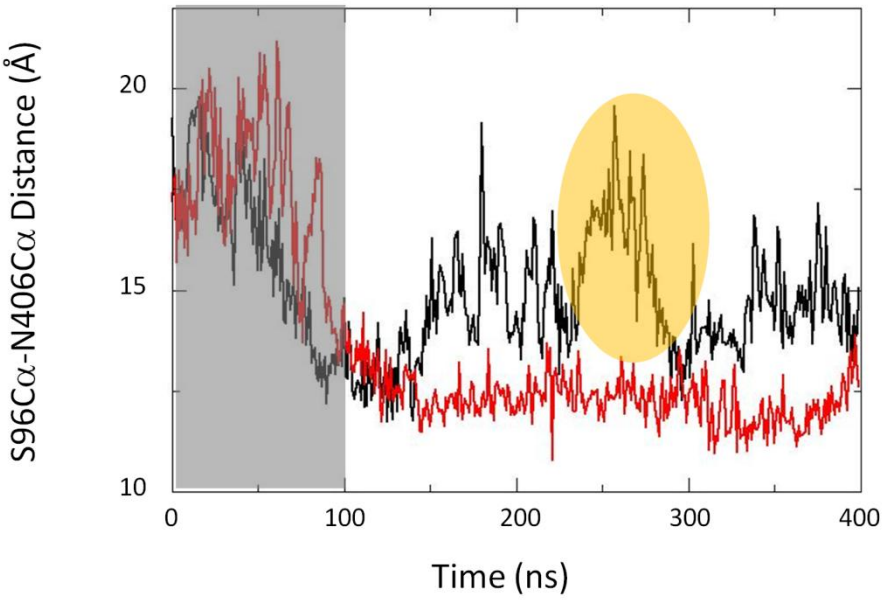


Figure S4.

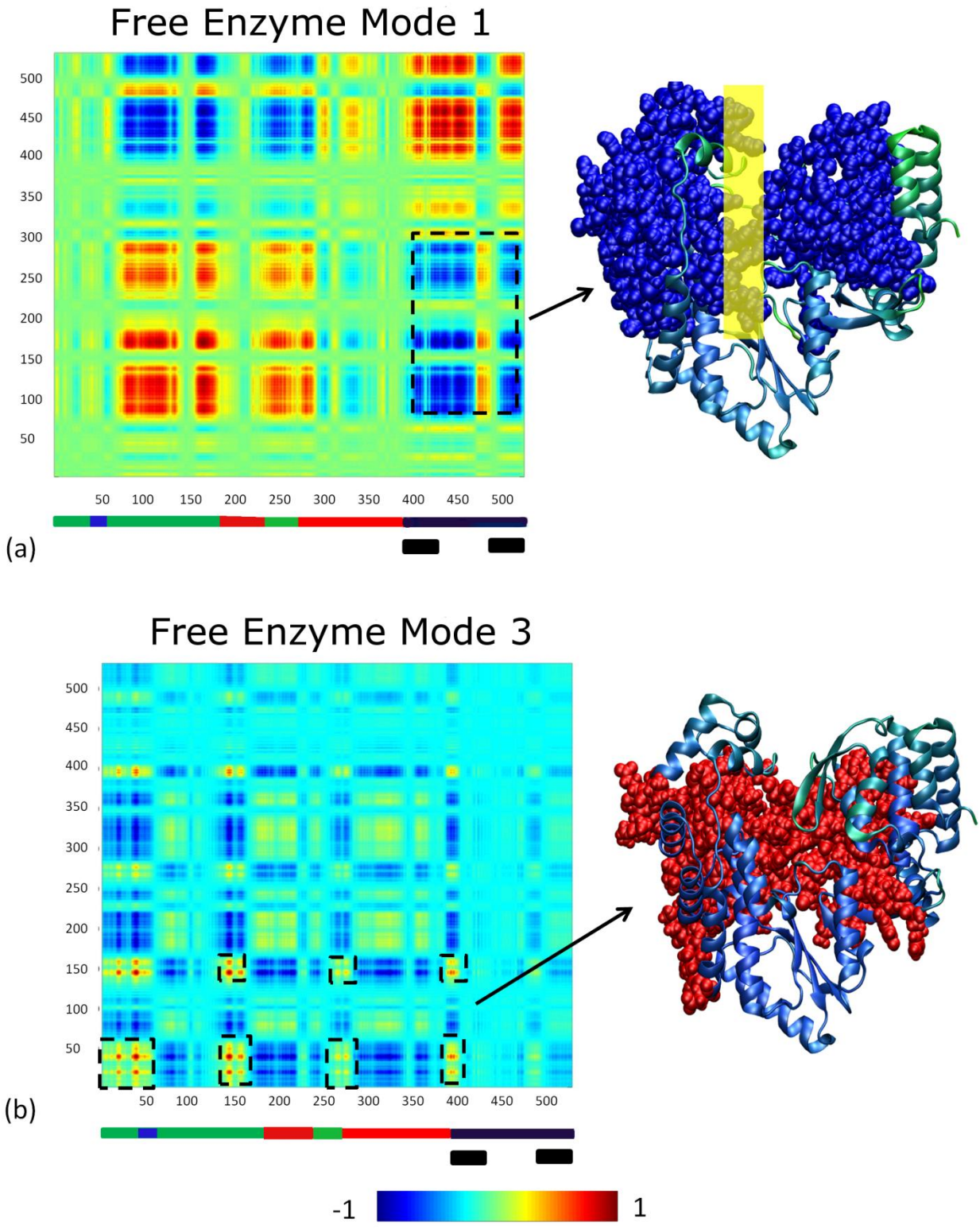

Free Enzyme Mode 1
500
450
400
350
300
250
200
150
100
50
50 100 150 200 250 300 350 400 450 500
(a)
Free Enzyme Mode 3
500
450
400
350
300
250
200
150
100
50
50 100 150 200 250 300 350 400 450 500
(b)
-1
1

Figure S5.