

Supplementary Materials:

***In silico* discovery of potent inhibitors against monkeypox's major structural proteins**

Kiran Bharat Lokhande^{1#}, Ashish Shrivastava^{1#}, Ashutosh Singh^{1#}

¹*Translational Bioinformatics and Computational Genomics Research Lab, Department of Life Sciences, Shiv Nadar Institution of Eminence, Gautam Buddha Nagar, UP-201314, India.*

Supplementary Figures:

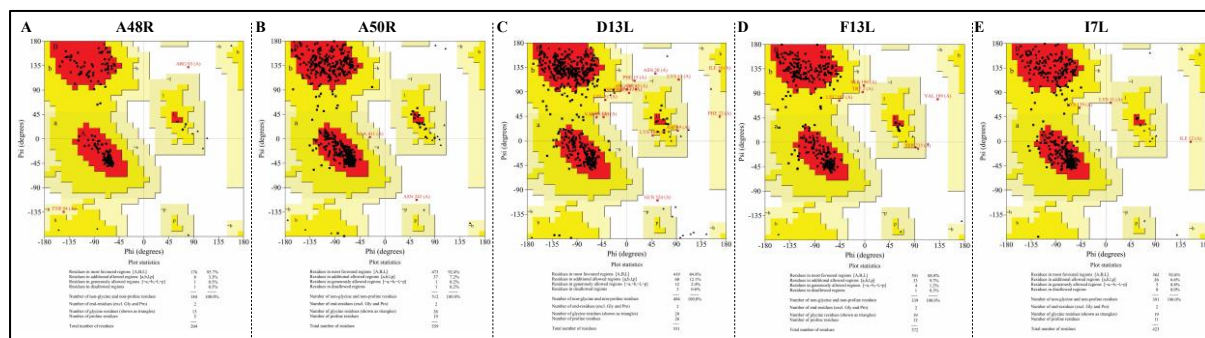


Figure S1: Stereochemical analysis of (A) Thymidylate kinase (A48R), (B) DNA ligase (A50R), (C) Scaffold protein D13 (D13L), (D) Palmytilated EEV membrane protein (F13L), and (E) Viral core cysteine proteinase (I7L), using SAVES server.

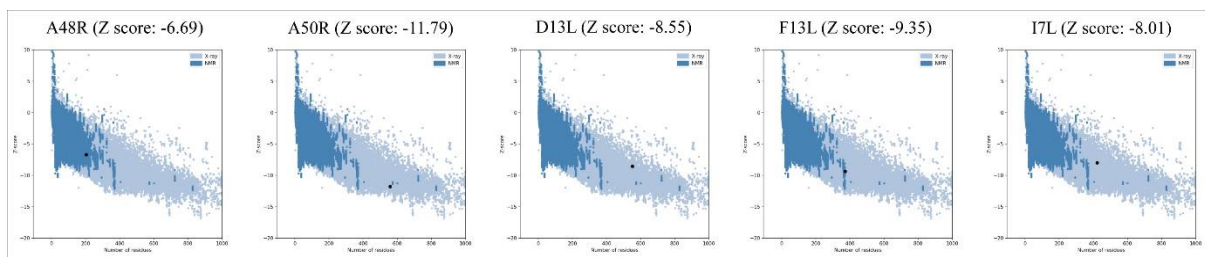


Figure S2: Stereochemical analysis of (A) Thymidylate kinase (A48R), (B) DNA ligase (A50R), (C) Scaffold protein D13 (D13L), (D) Palmytilated EEV membrane protein (F13L), and (E) Viral core cysteine proteinase (I7L), using ProSA server.

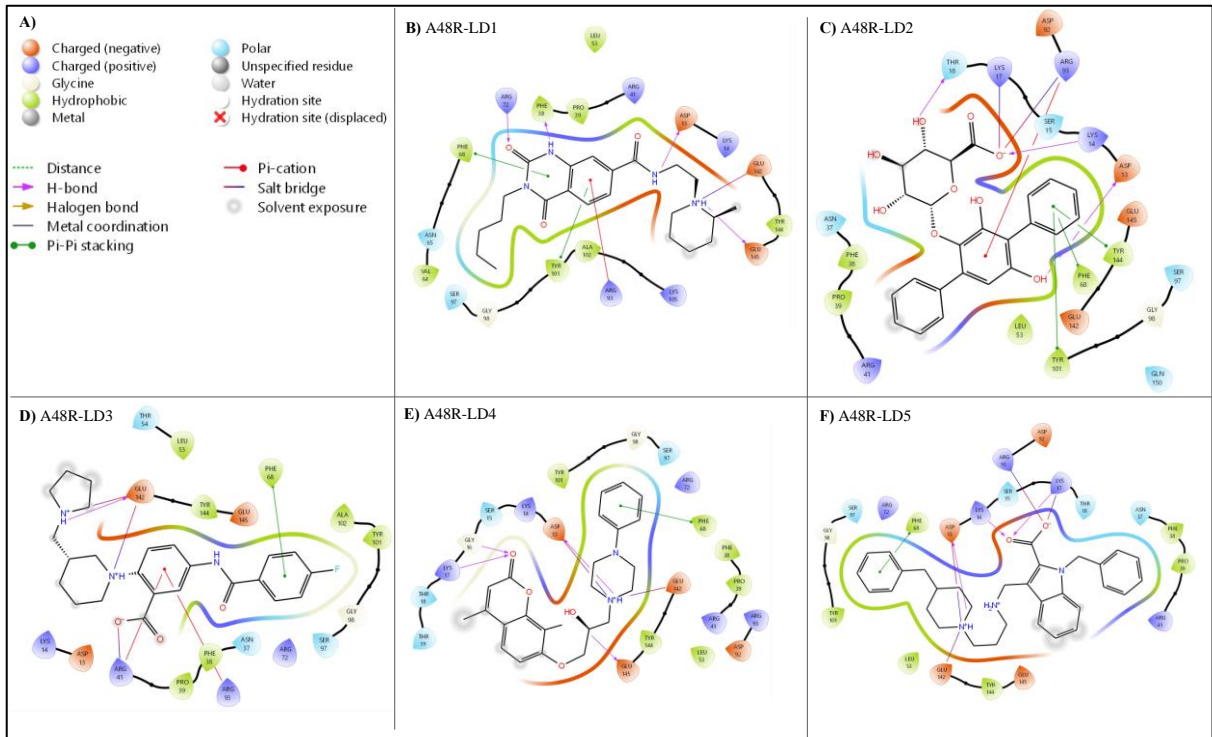


Figure S3: 2D interactions between docked ligand and binding site residues of A48R receptor. Types of interaction are depicted in figure (A). Figure (B-F), shows interactions between the A48R receptor and docked compounds.

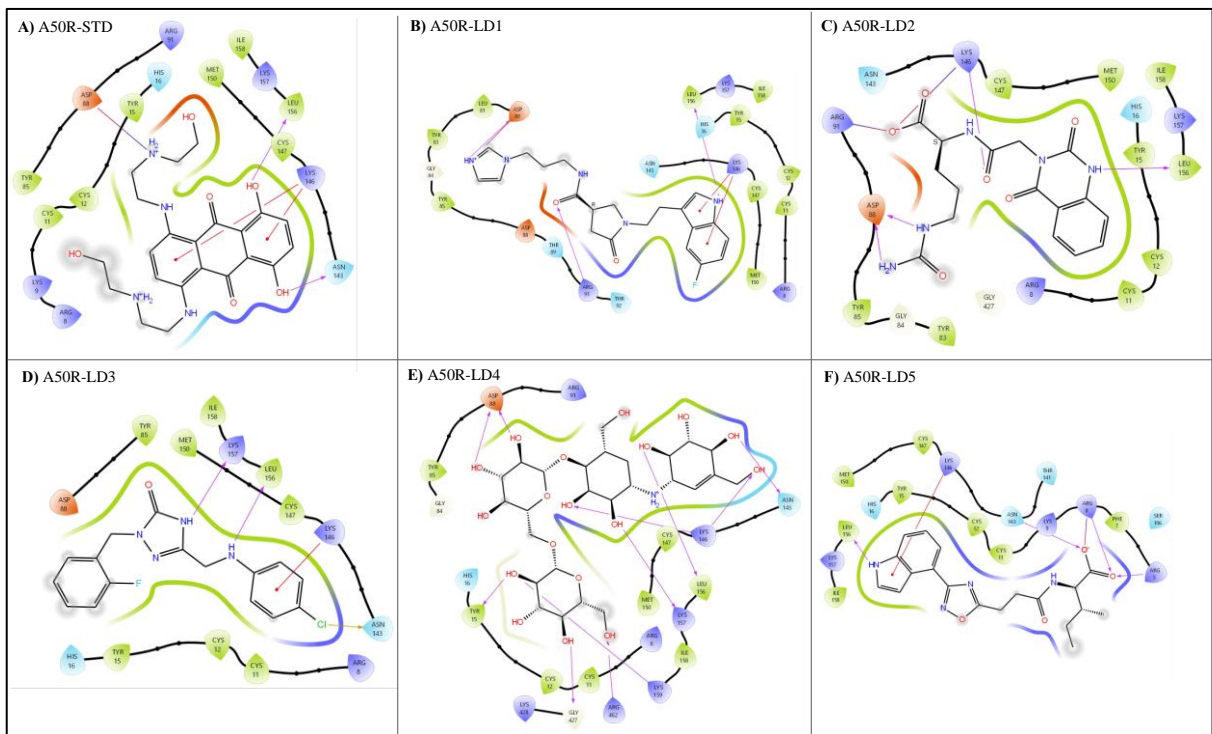


Figure S4: 2D interactions between docked ligand and binding site residues of A50R receptor. Figure (A-F), shows interactions between the A50R receptor and docked compounds.

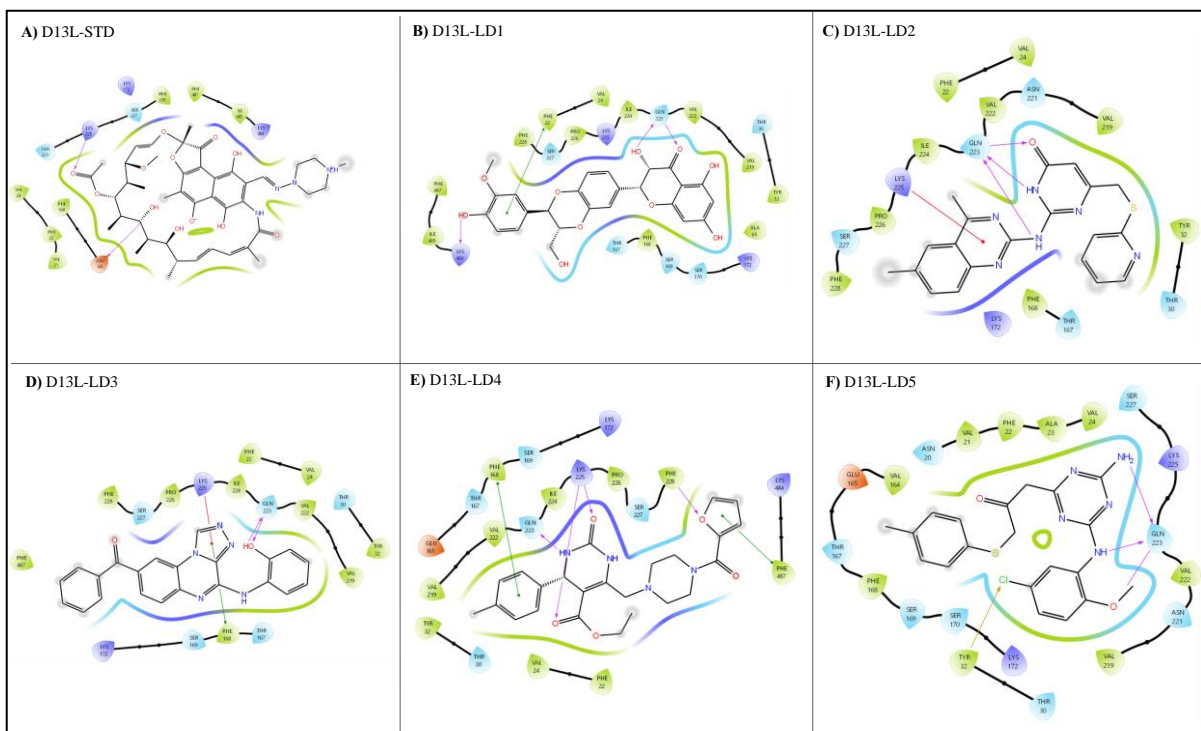


Figure S5: 2D interactions between docked ligand and binding site residues of D13L receptor. Figure (A-F), shows interactions between the D13L receptor and docked compounds.

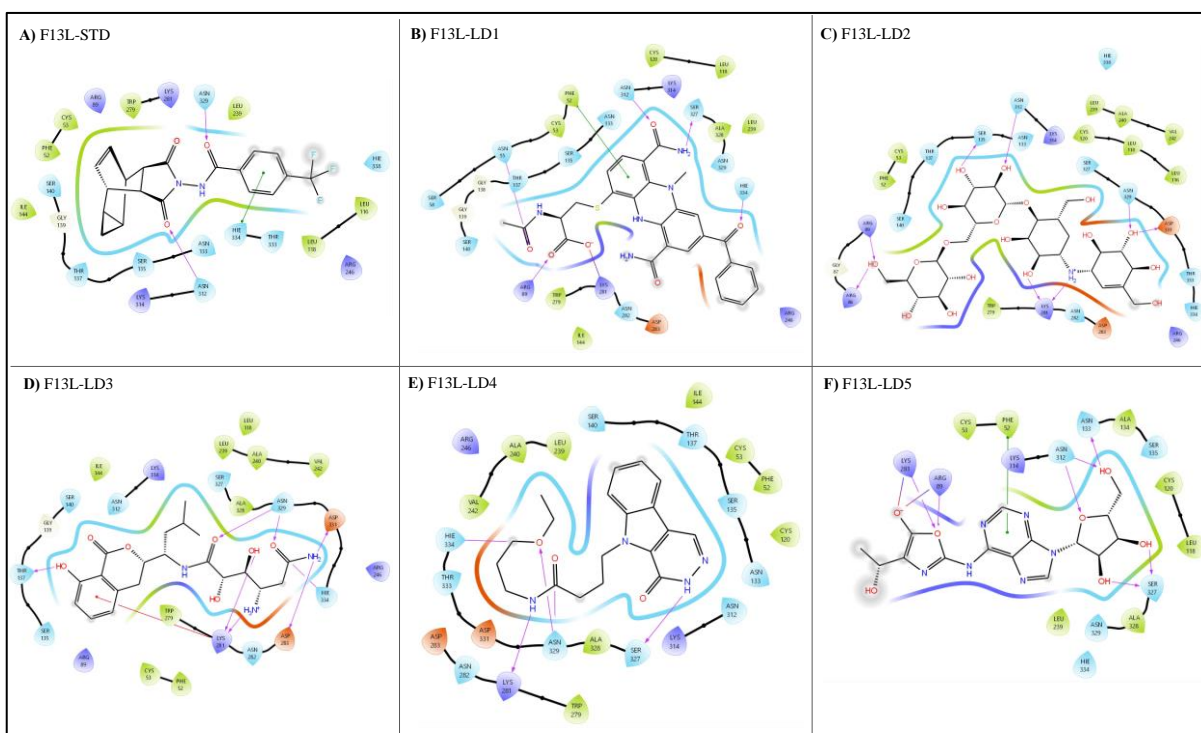


Figure S6: 2D interactions between docked ligand and binding site residues of F13L receptor. Figure (A-F), shows interactions between the F13L receptor and docked compounds.

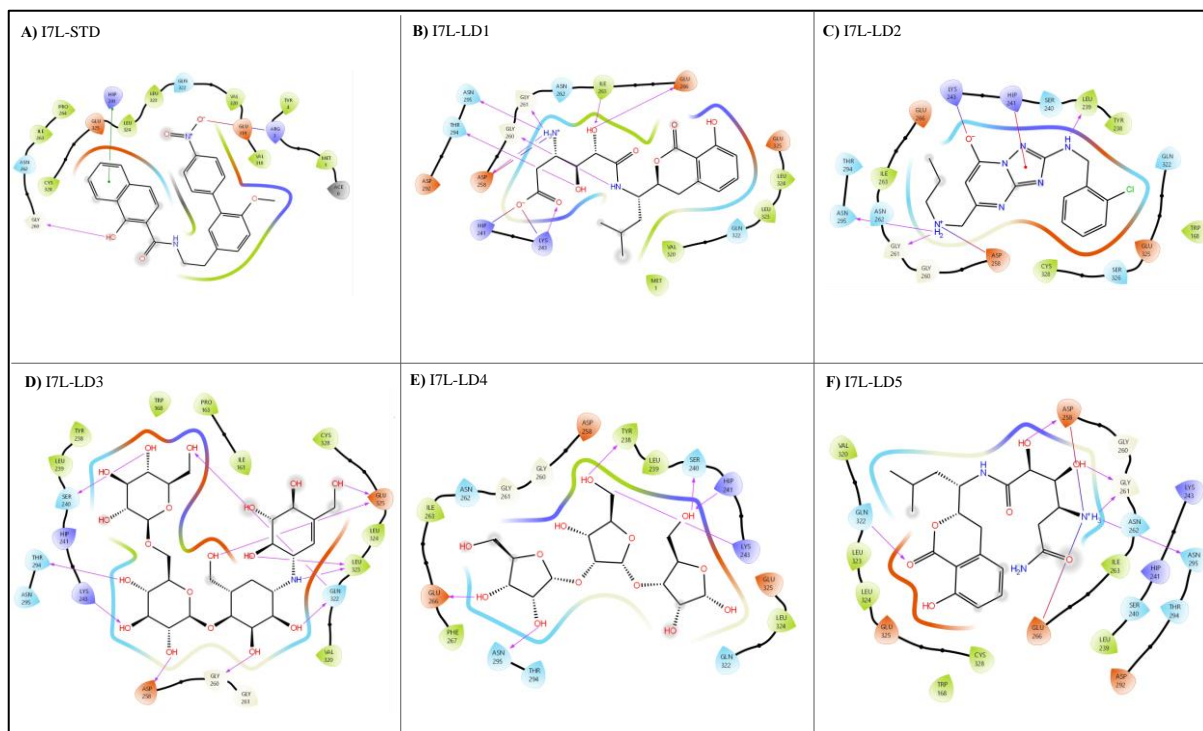


Figure S7: 2D interactions between docked ligand and binding site residues of I7L receptor. Figure (A-F), shows interactions between the I7L receptor and docked compounds.

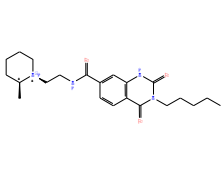
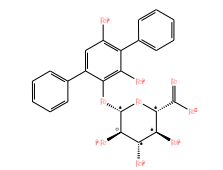
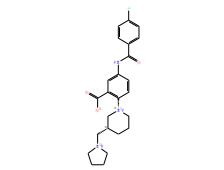
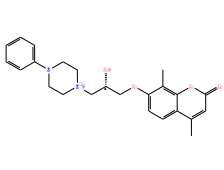
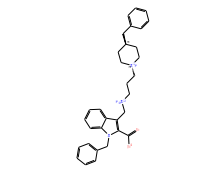
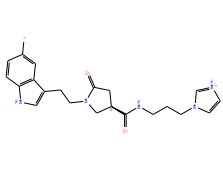
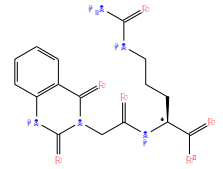
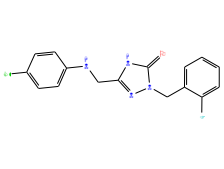
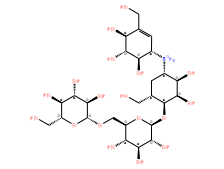
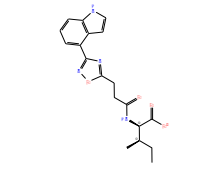
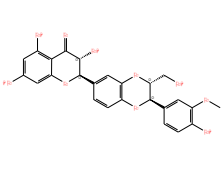
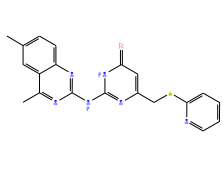
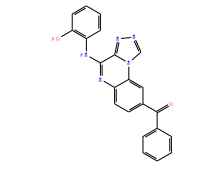
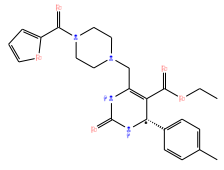
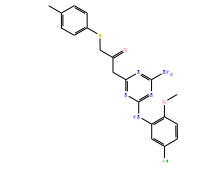
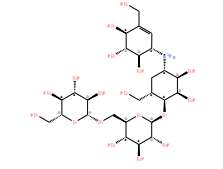
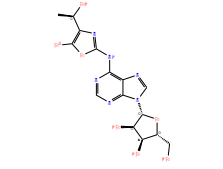
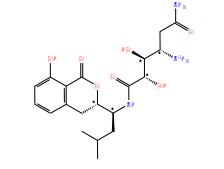
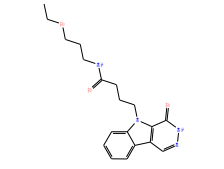
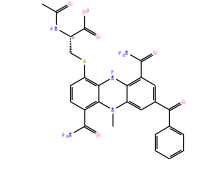
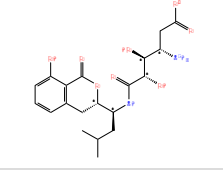
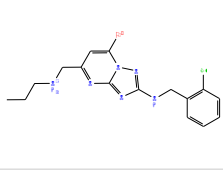
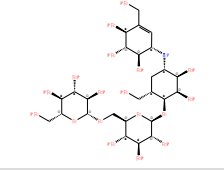
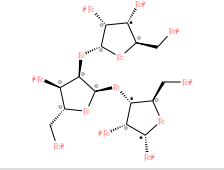
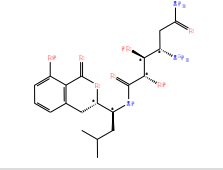
				
A48R-LD-1 -10.902	A48R-LD-2 -9.921	A48R-LD-3 -9.713	A48R-LD-4 -9.596	A48R-LD-5 -9.404
				
A50R-LD-1 -8.89	A50R-LD-2 -8.614	A50R-LD-3 -8.526	A50R-LD-4 -8.469	A50R-LD-5 -8.275
				
D13L-LD-1 -10.958	D13L-LD-2 -9.984	D13L-LD-3 -9.448	D13L-LD-4 -9.347	D13L-LD-5 -9.259
				
F13L-LD-1 -10.764	F13L-LD-2 -9.442	F13L-LD-3 -9.305	F13L-LD-4 -8.913	F13L-LD-5 -8.245
				
I7L-LD-1 -11.273	I7L-LD-2 -9.538	I7L-LD-3 -9.34	I7L-LD-4 -8.411	I7L-LD-5 -7.831

Figure S8: 2D structure of selected lead compounds for their respective targets with Glide docking score (kcal/mol) obtained from docking and virtual screening.

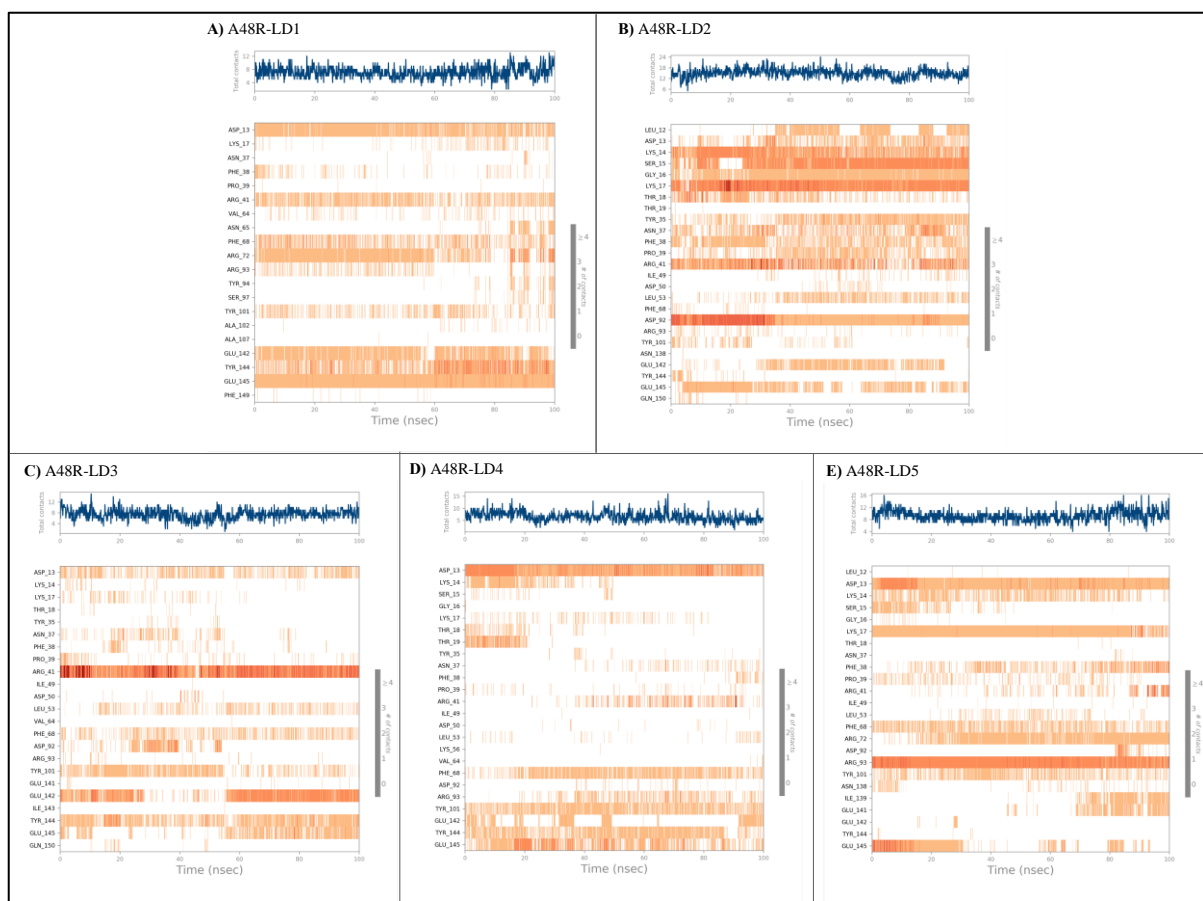


Figure S9: Protein-ligand contact plot: showing a timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). The top panel depicts the total number of individual interactions made by the protein with the ligand during the simulation time. In each trajectory frame, the bottom panel displays which residues interact with the ligand. Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange.

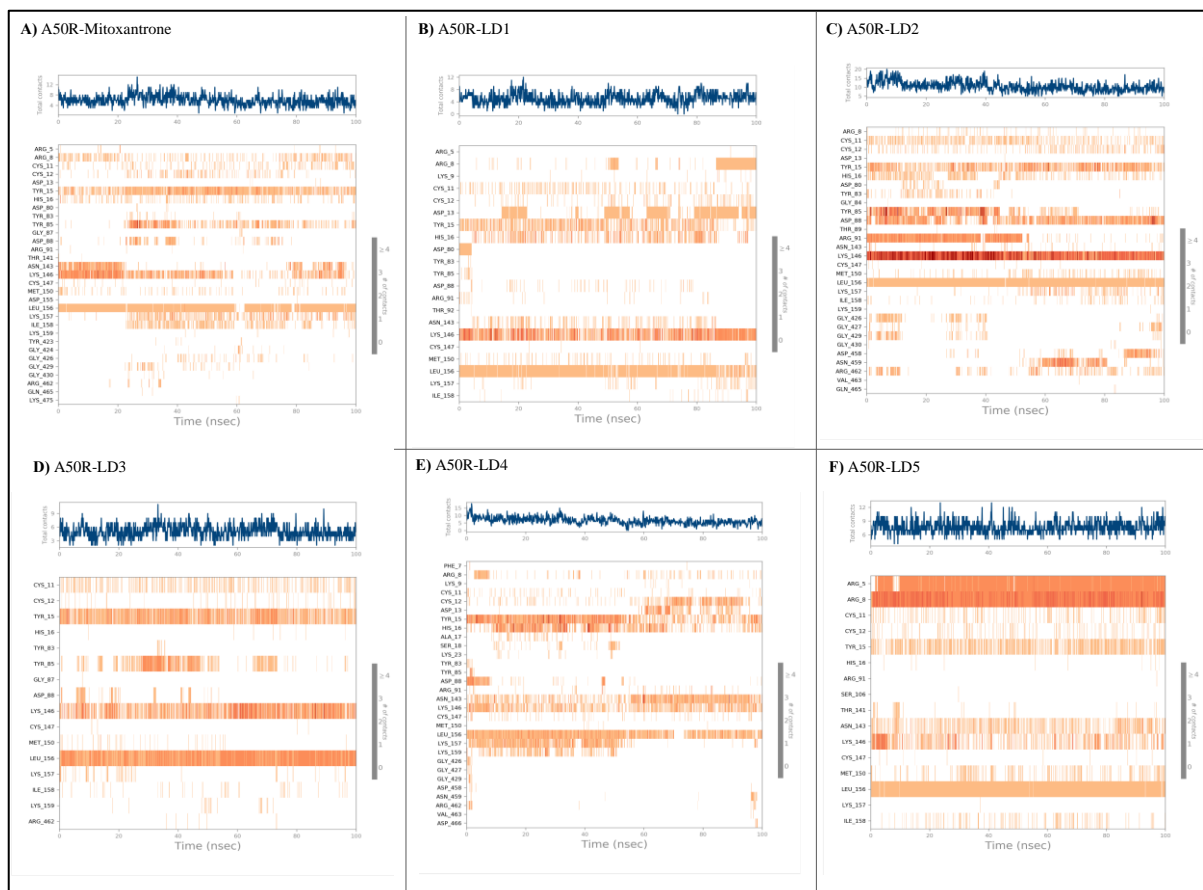


Figure S10: Protein-ligand contact plot: showing a timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). The top panel depicts the total number of individual interactions made by the protein with the ligand during the simulation time. In each trajectory frame, the bottom panel displays which residues interact with the ligand. Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange.

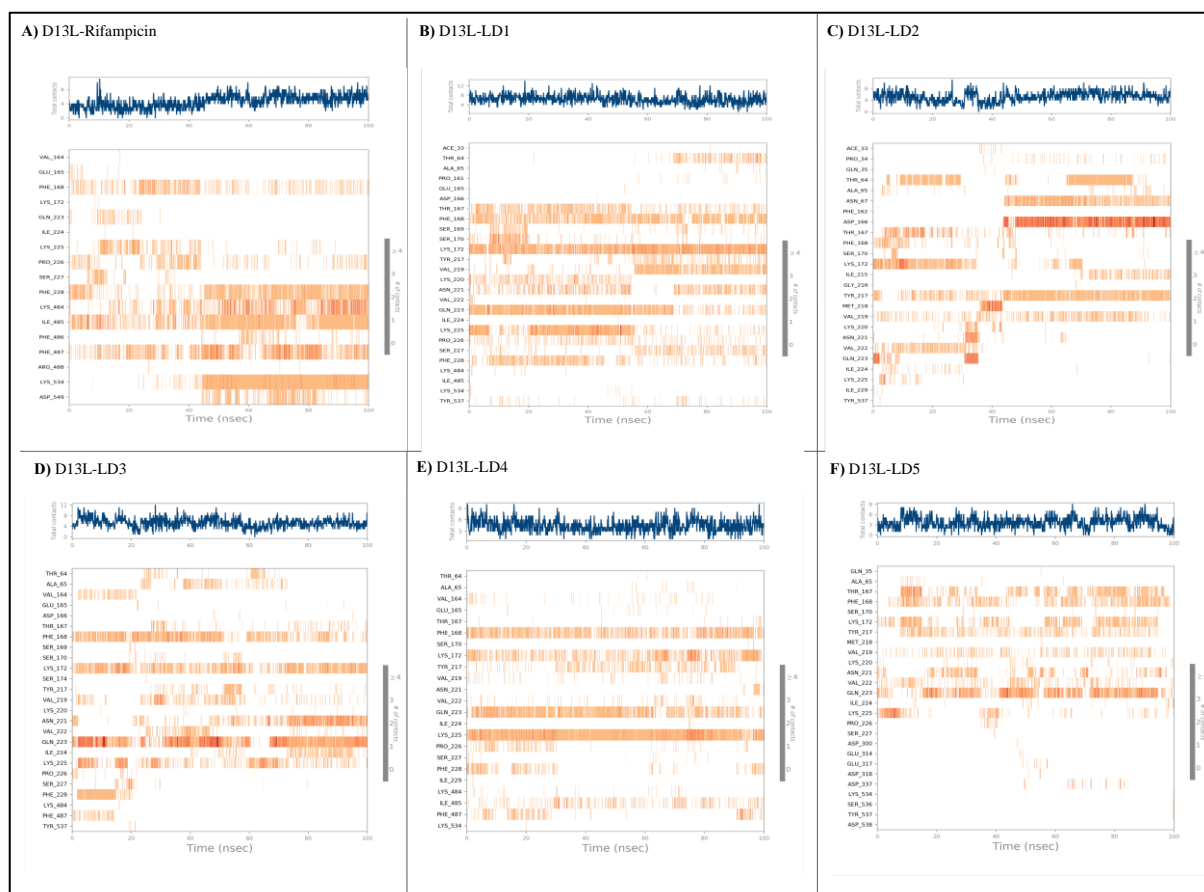


Figure S11: Protein-ligand contact plot: showing a timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). The top panel depicts the total number of individual interactions made by the protein with the ligand during the simulation time. In each trajectory frame, the bottom panel displays which residues interact with the ligand. Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange.

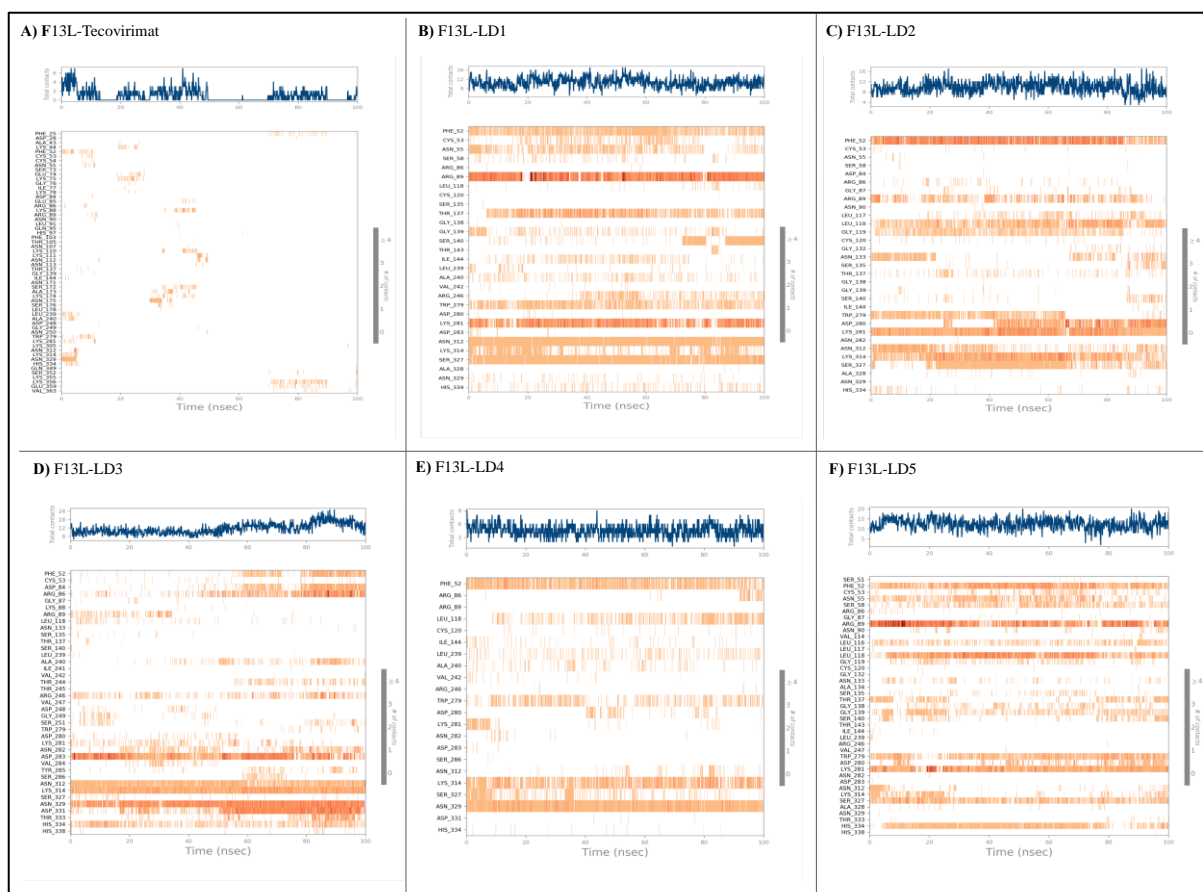


Figure S12: Protein-ligand contact plot: showing a timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). The top panel depicts the total number of individual interactions made by the protein with the ligand during the simulation time. In each trajectory frame, the bottom panel displays which residues interact with the ligand. Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange.

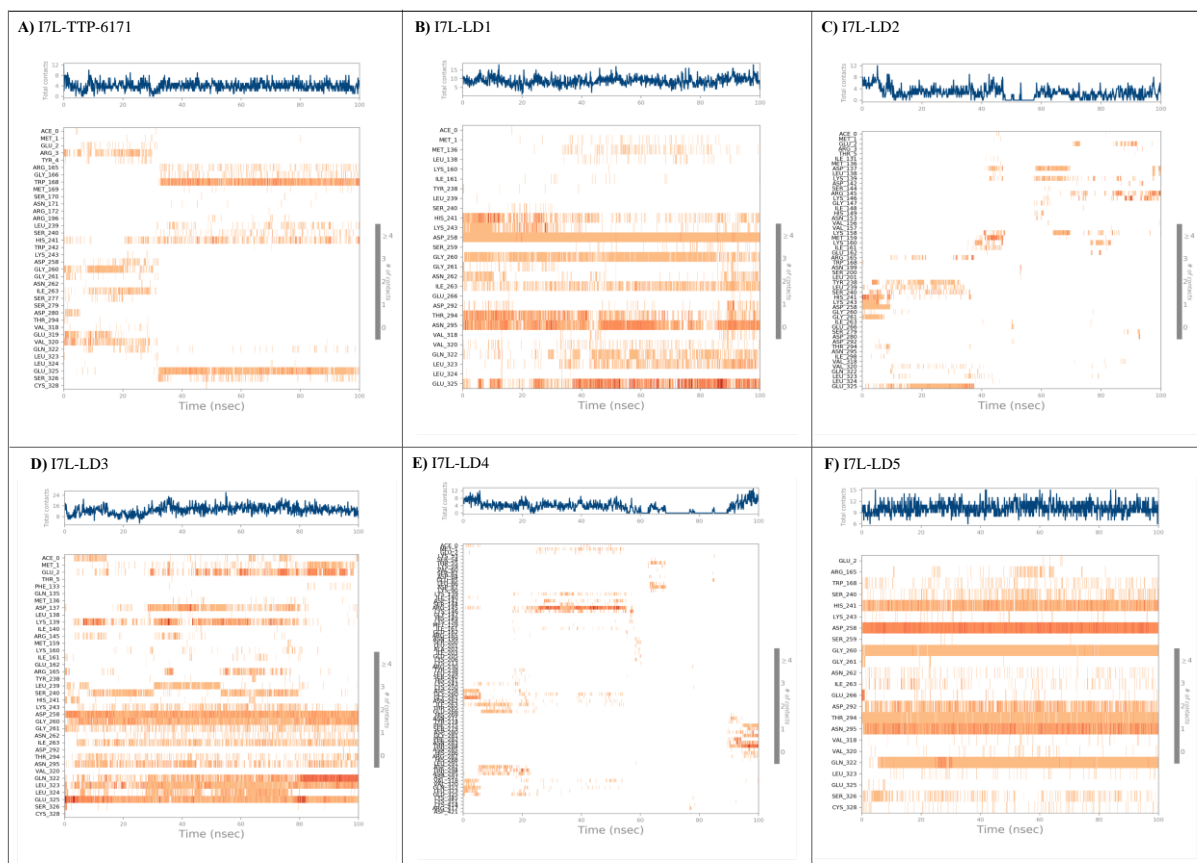


Figure S13: Protein-ligand contact plot: showing a timeline representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water bridges). The top panel depicts the total number of individual interactions made by the protein with the ligand during the simulation time. In each trajectory frame, the bottom panel displays which residues interact with the ligand. Some residues make more than one specific contact with the ligand, which is represented by a darker shade of orange.

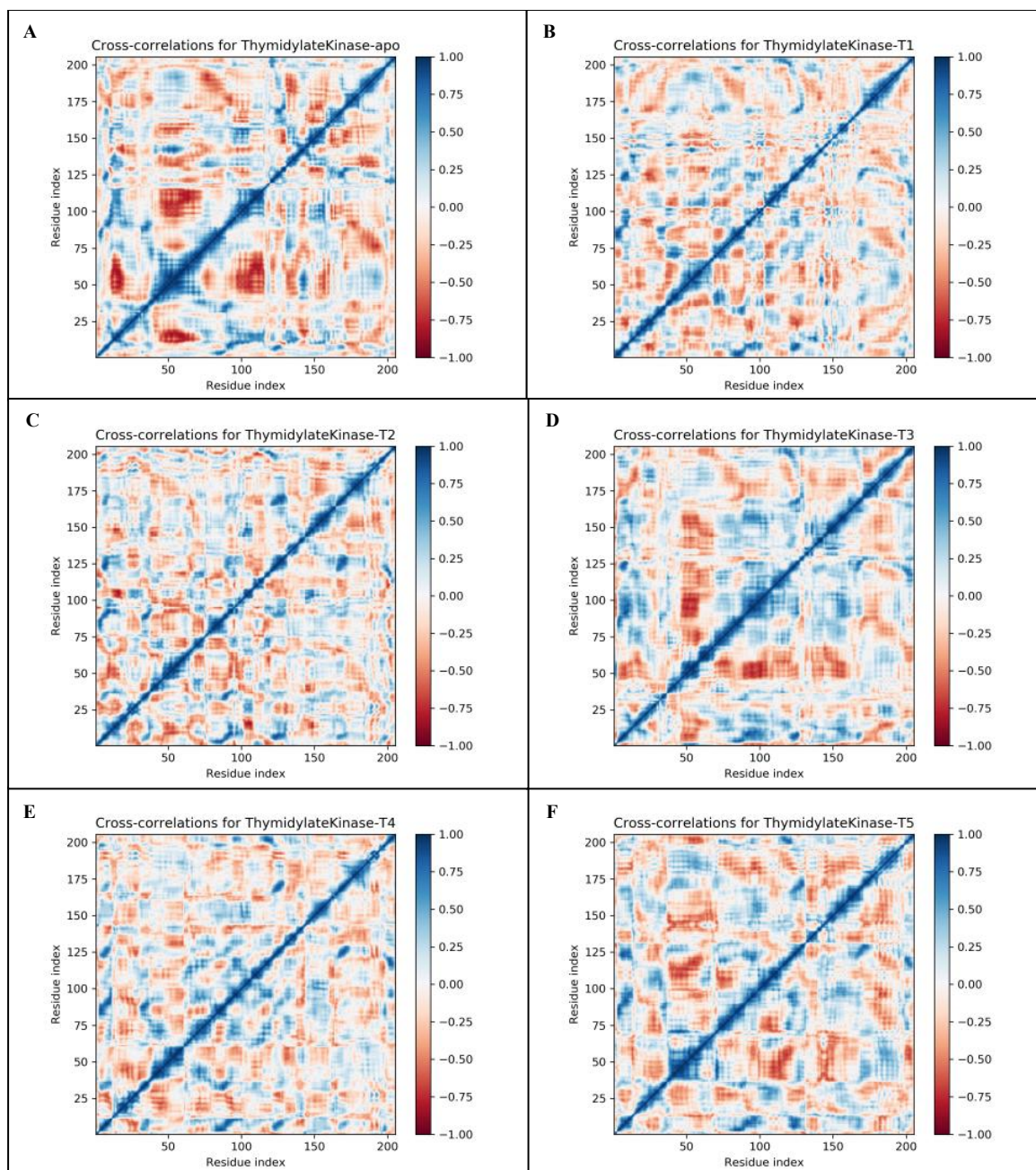


Figure S14: Cross correlations matrix obtained from essential dynamics of MD simulation trajectories for an individual amino acid. The color gradient represents correlated motion, where blue indicates highly correlated motion, and red indicates negatively correlated motion.

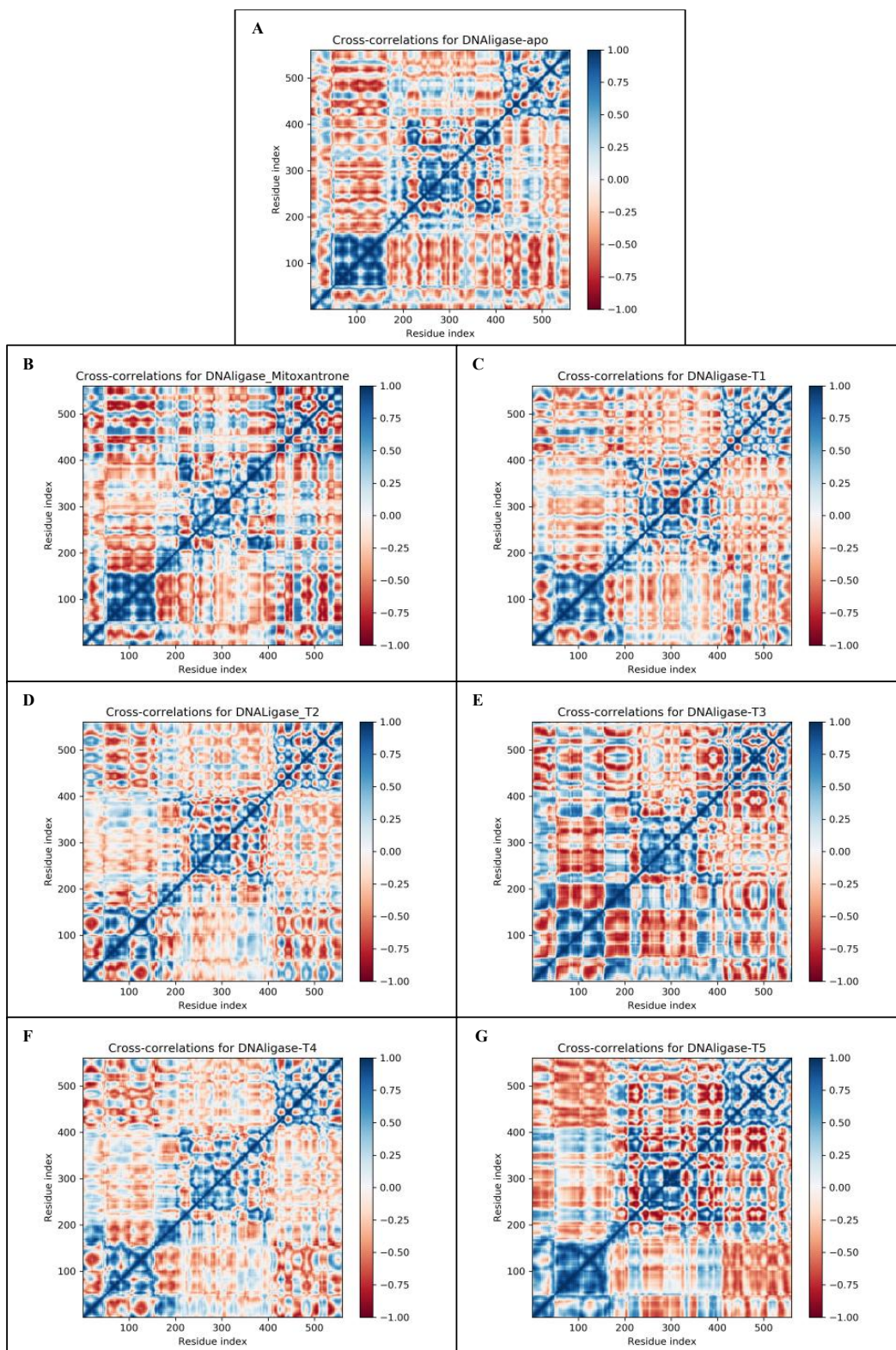


Figure S15: Cross correlations matrix obtained from essential dynamics of MD simulation trajectories for an individual amino acid. The color gradient represents correlated motion, where blue indicates highly correlated motion, and red indicates negatively correlated motion.

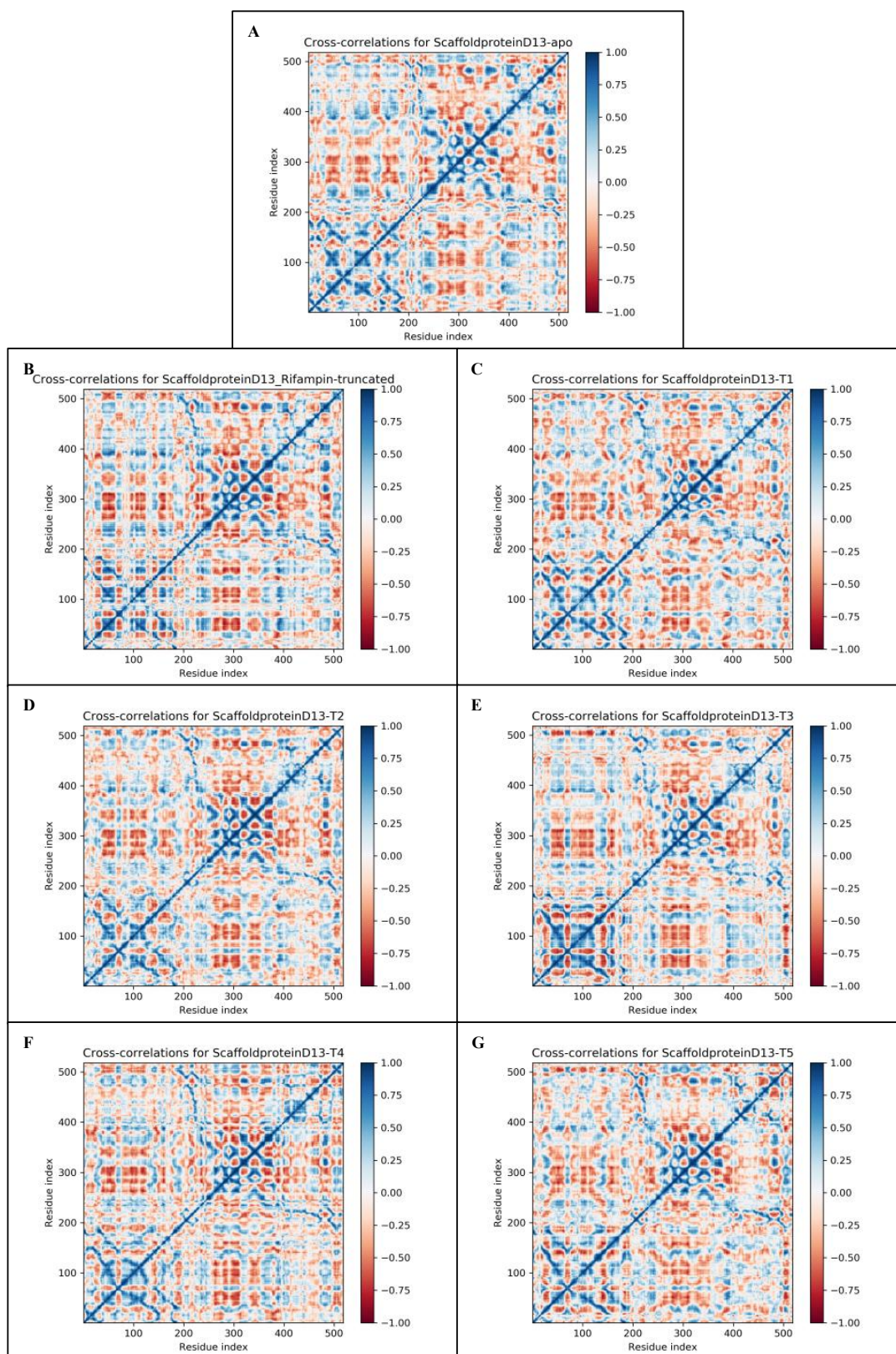


Figure S16: Cross correlations matrix obtained from essential dynamics of MD simulation trajectories for an individual amino acid. The color gradient represents correlated motion, where blue indicates highly correlated motion, and red indicates negatively correlated motion.

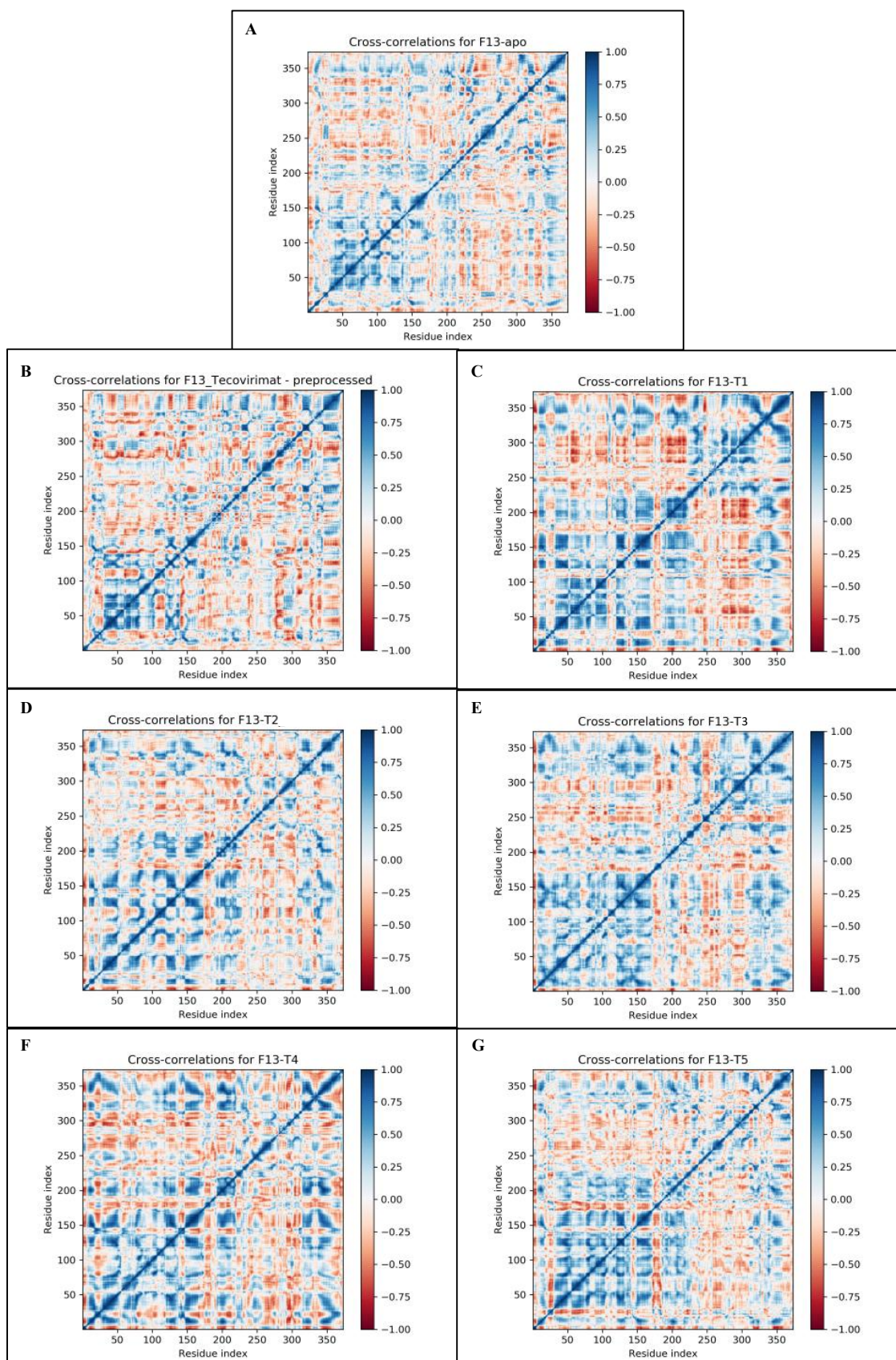


Figure S17: Cross correlations matrix obtained from essential dynamics of MD simulation trajectories for an individual amino acid. The color gradient represents correlated motion, where blue indicates highly correlated motion, and red indicates negatively correlated motion.

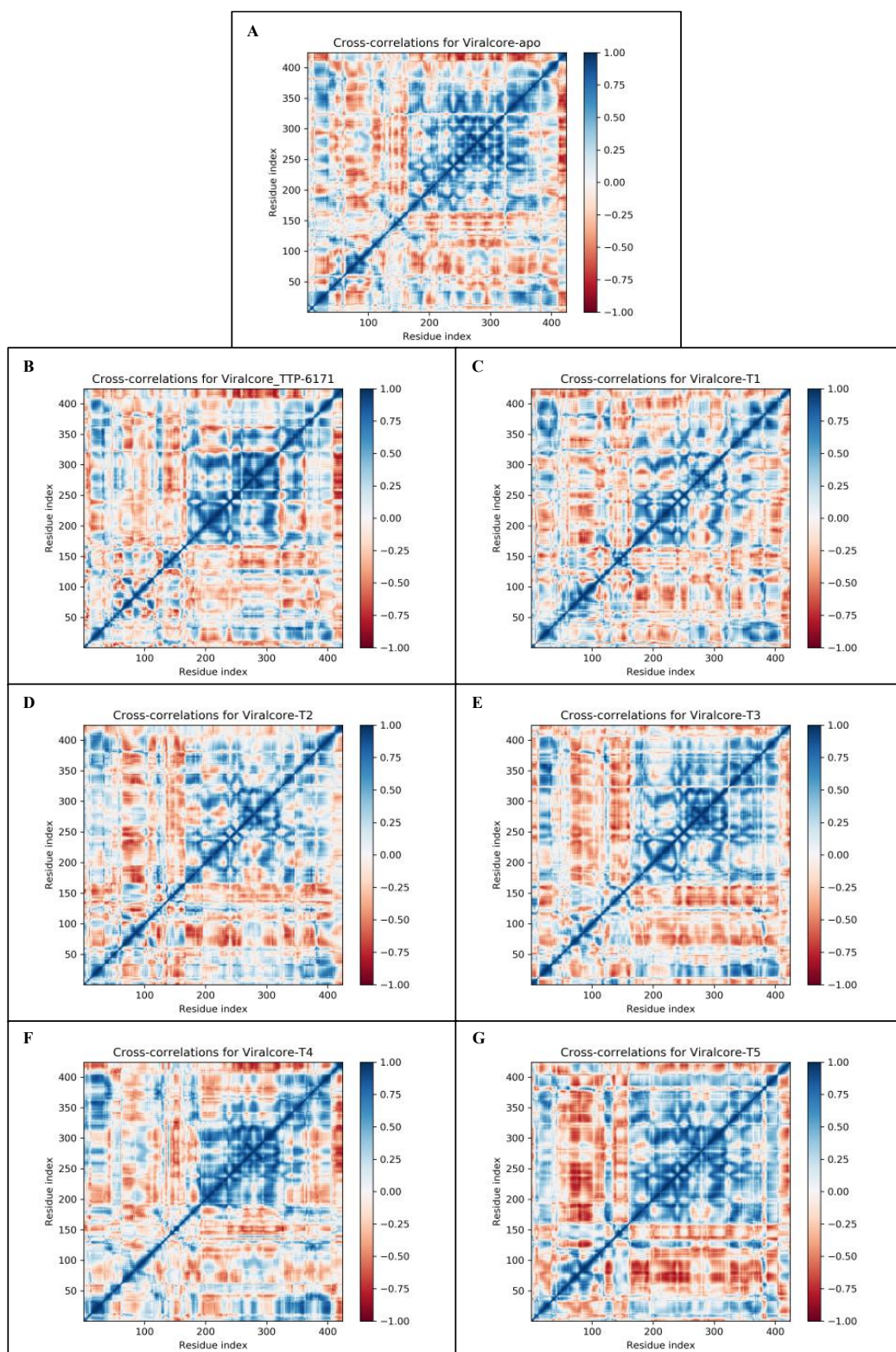


Figure S18: Cross correlations matrix obtained from essential dynamics of MD simulation trajectories for an individual amino acid. The color gradient represents correlated motion, where blue indicates highly correlated motion, and red indicates negatively correlated motion.

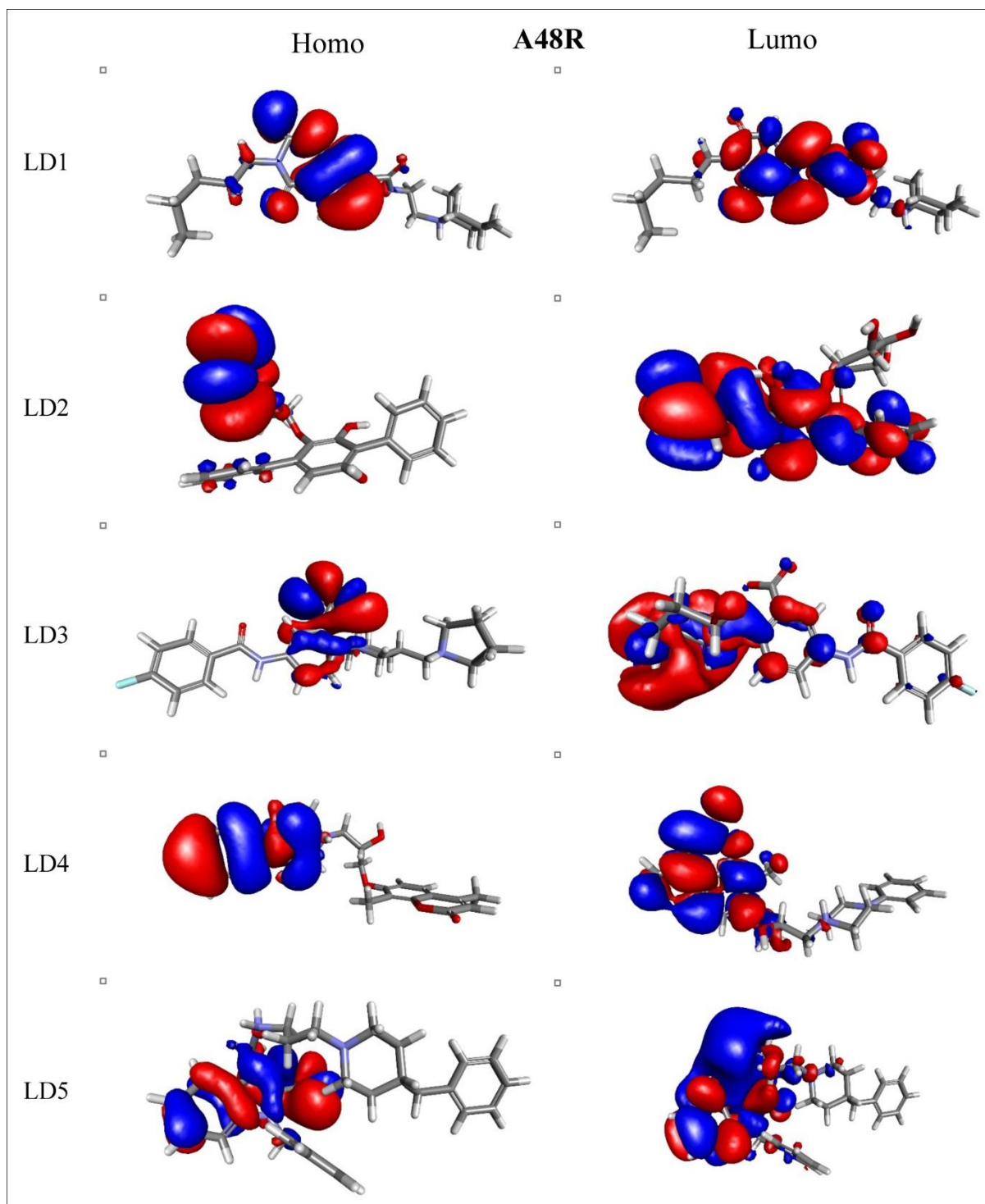


Figure S19: HOMO and LUMO shown for top 5 lead compounds against Thymidylate kinase (A48R).

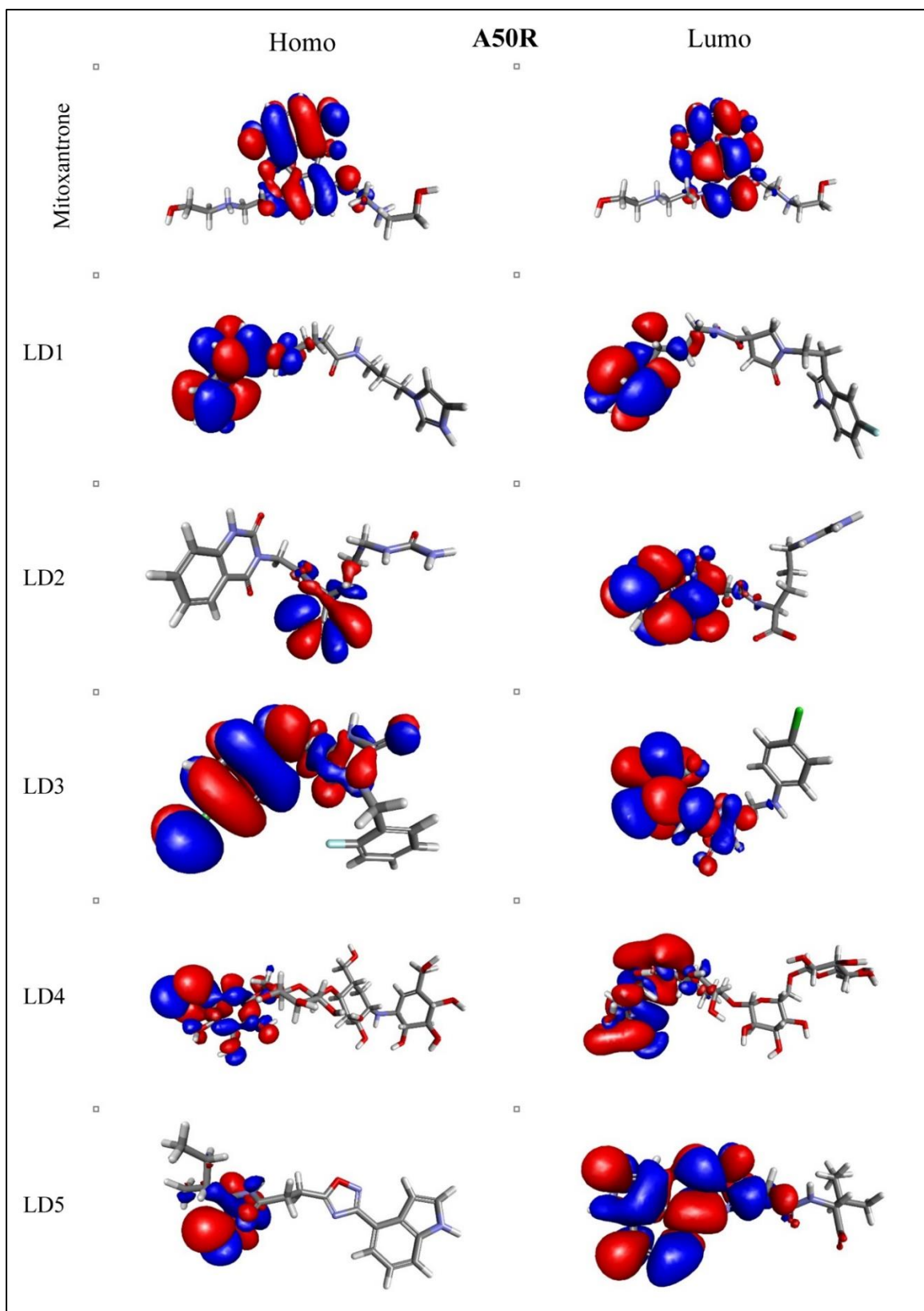


Figure S20: HOMO and LUMO shown for Mitoxantrone and top 5 lead compounds against DNA Ligase (A50R).

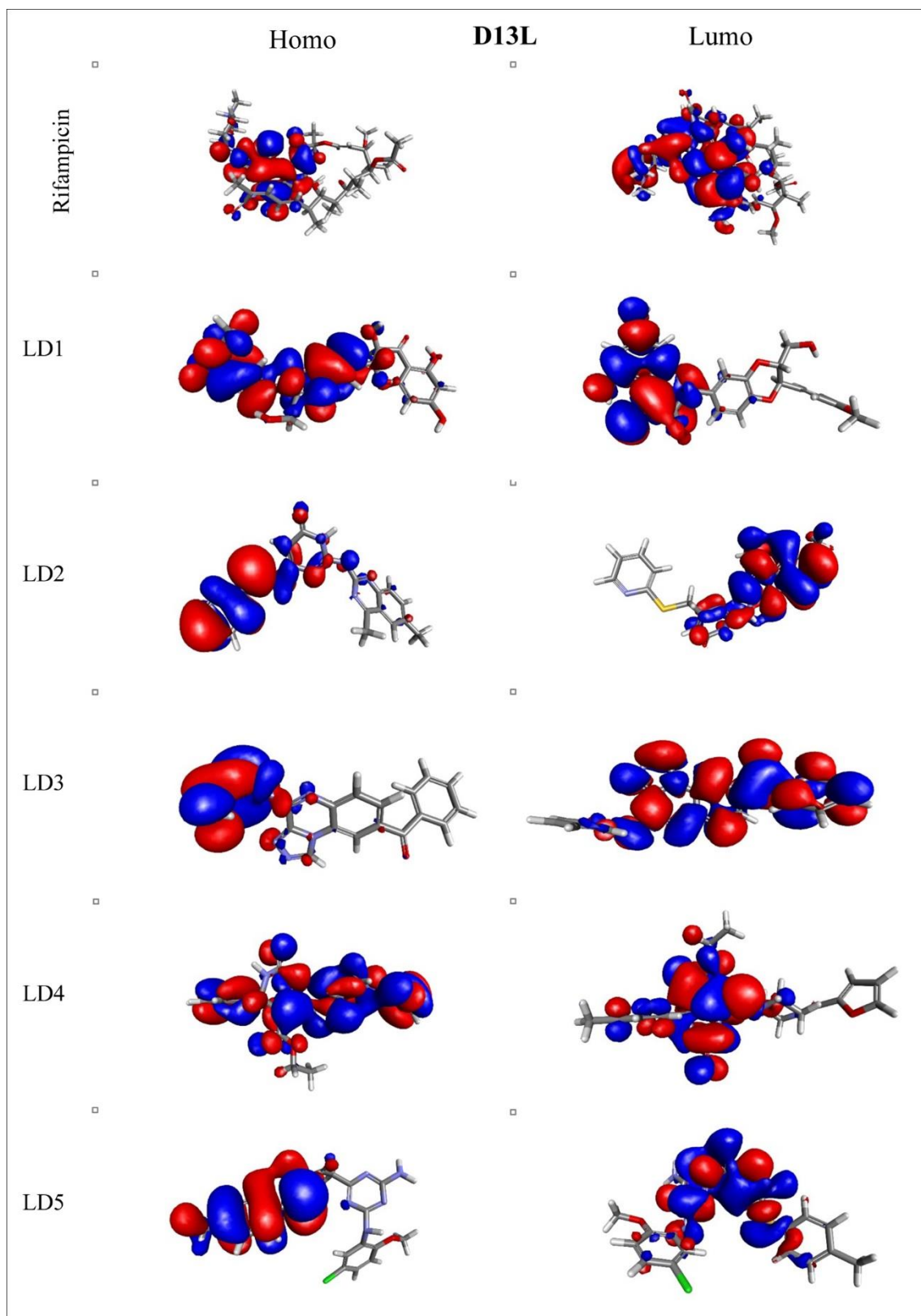


Figure S21: HOMO and LUMO shown for Rifampicin and top 5 lead compounds against Scaffold Protein (D13L).

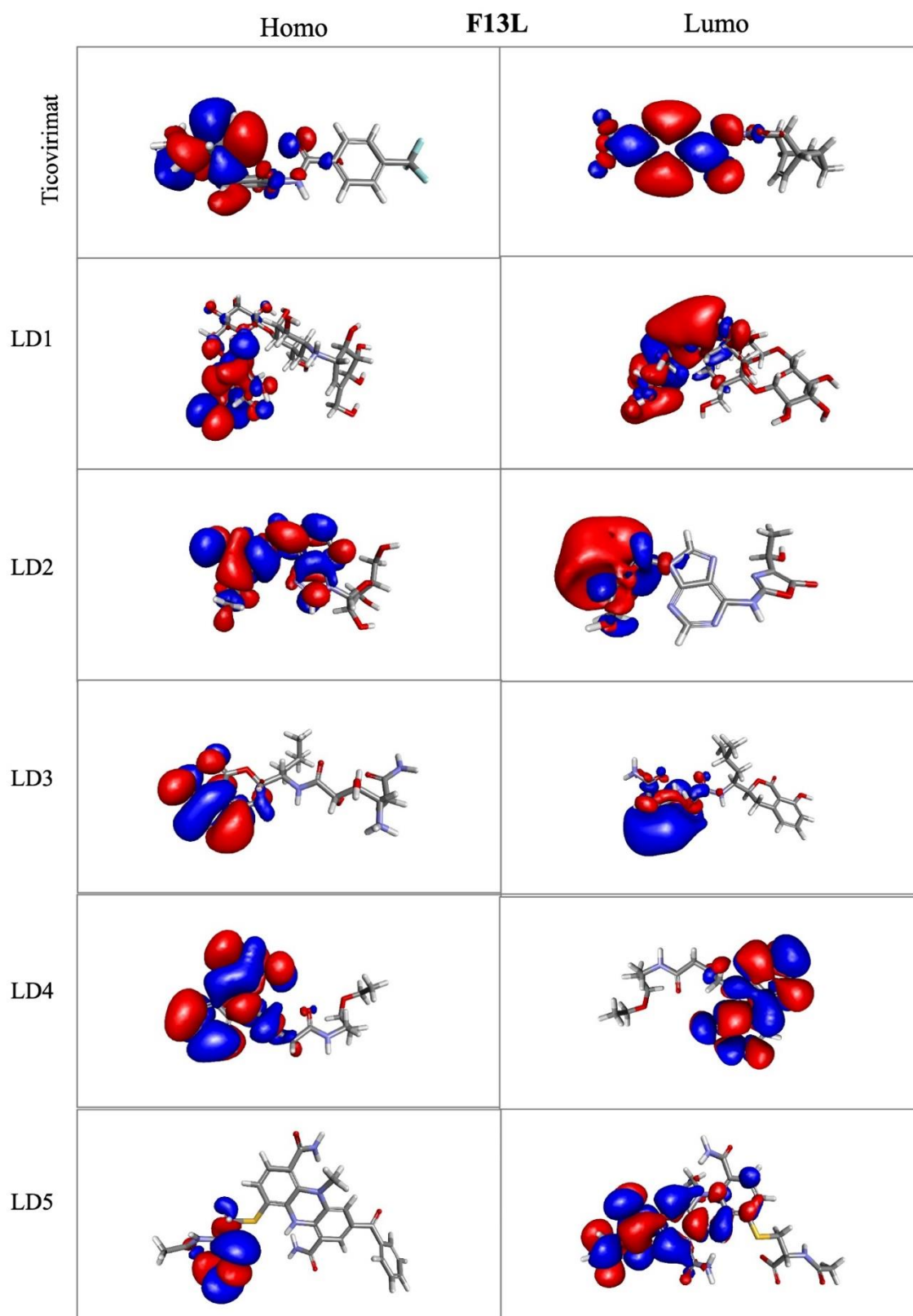


Figure S22: HOMO and LUMO shown for Ticovirimat and top 5 lead compounds against Palmytilated EEV membrane protein (F13L).

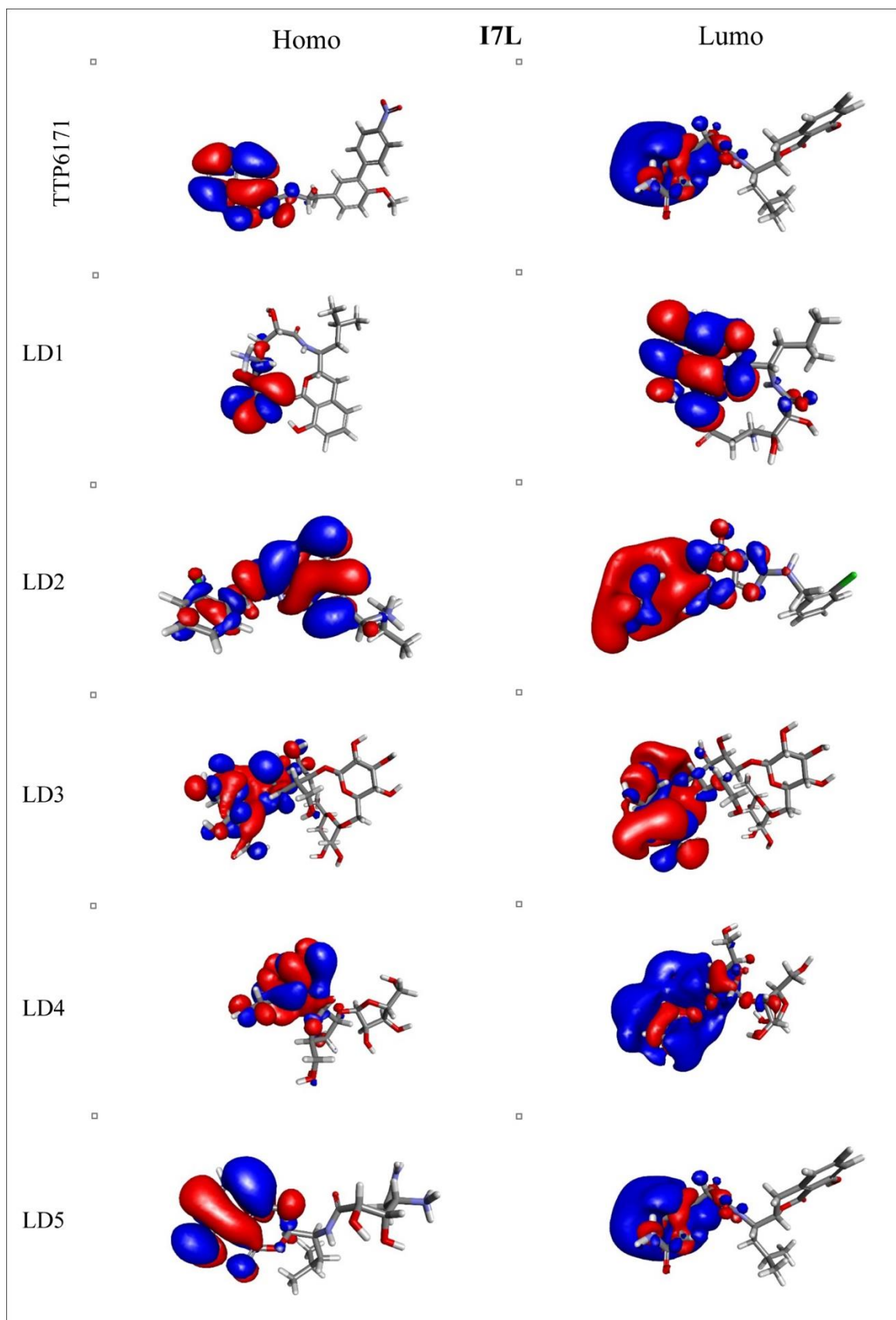


Figure S23: HOMO and LUMO shown for TTP6171 and top 5 lead compounds against Viral core cysteine proteinase (I7L).

Supplementary Tables:

Table S1: Details of the intermolecular interactions between Thymidylate kinase and top lead compounds from the screening studies.

Receptor	Compound Name	Docking Score (kcal/mol)	Interacting Amino Acid Residues	Type of Interaction	Bond Distance (Å)
Thymidylate kinase (A48R)	A48R-LD-1	-10.902	Asp13	H-Bond	1.94
			Phe38	H-Bond	2.38
			Phe68	Pi-pi	4.18
			Arg72	H-Bond	2.11
			Arg93	Pi-cation	4.81
			Tyr101	Pi-pi	5.18
			Glu142	Salt-Bridge	4.34
			Glu145	H-Bond	2.05
	A48R-LD-2	-9.921	Asp13	H-Bond	1.92
			Lys14	H-Bond	2.27
			Lys17	Salt-Bridge	3.12
			Thr18	H-Bond	1.69
			Phe68	Pi-pi	4.89
			Arg93	Pi-cation	5.01
			Arg93	Salt-Bridge	4.30
			Tyr101	Pi-cation	3.69
	A48R-LD-3	-9.713	Arg41	Salt-Bridge	3.77
			Arg41	Pi-cation	4.23
			Phe68	Pi-pi	3.40
			Arg93	Pi-cation	5.06
			Glu142	2 H-Bond	2.08, 4.41
			Glu142	Salt-Bridge	4.06
	A48R-LD-4	-9.596	Asp13	H-Bond	2.13
			Asp13	Salt-Bridge	3.91
			Gly16	H-Bond	2.10
Lys17			H-Bond	2.21	
Phe68			Pi-pi	3.71	
Glu142			Salt-Bridge	4.15	
Glu145			H-Bond	2.66	
A48R-LD-5	-9.404	Asp13	H-Bond	2.18	
		Asp13	Salt-Bridge	4.00	
		Lys14	H-Bond	2.25	
		Lys17	H-Bond	2.33	
		Lys17	Salt-Bridge	3.82	
		Phe68	Pi-pi	3.62	
		Arg93	Salt-Bridge	3.40	
		Glu142	Salt-Bridge	3.65	

Table S2: Details of the intermolecular interactions between DNA ligase, *in vitro* proven drug and top lead compounds from the screening studies.

Receptor	Compound Name	Docking Score (kcal/mol)	Interacting Amino Acid Residues	Type of Interaction	Bond Distance (Å)
DNA ligase (A50R)	A50R-LD-1	-8.890	Asp80	H-Bond	1.99
			Asp80	Salt-Bridge	3.71
			Arg91	H-Bond	2.72
			Lys146	2 Pi-cation	4.93, 4.93
	A50R-LD-2	-8.614	Leu156	2 H-Bond	1.76, 1.90
			Asp88	2 H-Bond	2.13, 2.15
			Arg91	Salt-Bridge	3.29
			Lys146	2 Salt-Bridge	4.21, 1.90
	A50R-LD-3	-8.526	Asn143	Halogen Bond	3.11
			Lys146	Pi-cation	4.46
			Leu156	H-Bond	1.78
			Lys157	H-Bond	1.74
	A50R-LD-4	-8.469	Tyr15	H-Bond	1.72
			Asp88	2 H-Bond	1.73, 2.01
			Asn143	2 H-Bond	1.85, 2.32
			Lys146	2 H-Bond	2.27, 2.11
			Leu156	H-Bond	2.33
			Lys157	H-Bond	2.73
			Lys159	H-Bond	2.34
			Gly427	H-Bond	2.02
	A50R-LD-5	-8.275	Arg462	H-Bond	2.27
			Arg5	H-Bond	2.53
			Arg8	H-Bond	2.26
			Arg8	Salt-Bridge	3.68
			Asn143	H-Bond	2.19
			Lys146	Pi-cation	4.86
	Mitoxantrone	-4.776	Leu156	H-Bond	1.78
Asp88			Salt-Bridge	4.62	
Asn143			H-Bond	2.68	
Lys146			2 Pi-cation	4.26, 4.65	
			Leu156	H-Bond	1.72

Table S3: Details of the intermolecular interactions between Scaffold protein D13, *in vitro* proven drug and top lead compounds from the screening studies.

Receptor	Compound Name	Docking Score (kcal/mol)	Interacting Amino Acid Residues	Type of Interaction	Bond Distance (Å)
Scaffold protein D13 (D13L)	D13L-LD-1	-10.958	Phe22	Pi-pi	5.22
			Gln223	2 H-Bond	1.83, 2.32
			Lys484	H-Bond	1.71
	D13L-LD-2	-9.984	Gln223	3 H-Bond	1.73, 2.10, 2.35
			Lys225	Pi-cation	5.37
	D13L-LD-3	-9.448	Phe168	Pi-pi	4.82
			Lys225	Pi-cation	4.16
	D13L-LD-4	-9.347	Phe168	Pi-pi	5.30
			Gln223	H-Bond	1.66
			Lys225	2 H-Bond	1.83, 2.75
			Phe228	H-Bond	2.53
			Phe487	Pi-pi	4.96
	D13L-LD-5	-9.259	Gln223	3 H-Bond	1.90, 2.01, 2.25
Rifampicin	-4.631	Glu165	H-Bond	1.81	
		Lys225	H-Bond	2.23	

Table S4: Details of the intermolecular interactions between Palmytilated EEV membrane protein, *in vitro* proven drug and top lead compounds from the screening studies.

Receptor	Compound Name	Docking Score (kcal/mol)	Interacting Amino Acid Residues	Type of Interaction	Bond Distance (Å)
Palmytilated EEV membrane protein (F13L)	F13L-LD-1	-10.764	Phe52	Pi-pi	5.10
			Asn55	H-Bond	2.70
			Arg89	H-Bond	2.27
			Lys281	Salt-Bridge	3.35
			Asn312	H-Bond	2.13
			Ser327	H-Bond	1.85
			His334	H-Bond	1.91
	F13L-LD-2	-9.442	Arg86	H-Bond	2.78
			Arg89	H-Bond	2.14
			Lys281	2 H-Bond	1.82, 2.00
			Ser135	H-Bond	1.80
			Asn312	H-Bond	1.87
			Asn329	H-Bond	2.11
	F13L-LD-3	-9.305	Asp331	H-Bond	2.01
			Thr137	H-Bond	1.91
			Asp283	H-Bond	2.11
			Asn329	2 H-Bond	1.92, 2.21
			Asp331	H-Bond	2.37
	F13L-LD-4	-8.913	His334	H-Bond	1.90
			Ser327	H-Bond	1.86
			Asn329	2 H-Bond	2.30, 2.57
F13L-LD-5	-8.245	His334	H-Bond	2.15	
		Phe52	Pi-pi	5.10	
		Asn55	H-Bond	2.70	
		Asn312	H-Bond	2.13	
		Ser327	H-Bond	1.85	
Tecovirimat	-4.528	His334	H-Bond	1.91	
		Asn312	H-Bond	2.33	
		Asn329	H-Bond	2.42	
			His334	Pi-pi	4.43

Table S5: Details of the intermolecular interactions between Viral core cysteine proteinase, *in vitro* proven drug and top lead compounds from the screening studies.

Receptor	Compound Name	Docking Score (kcal/mol)	Interacting Amino Acid Residues	Type of Interaction	Bond Distance (Å)
Viral core cysteine proteinase (I7L)	I7L-LD-1	-11.274	His241	Salt-Bridge	4.36
			Lys243	Salt-Bridge	4.93
			Lys243	H-Bond	2.70
			Asp258	Salt-Bridge	2.87
			Asp258	H-Bond	2.04
			Gly260	H-Bond	2.28
			Gly261	H-Bond	1.99
			Ile263	H-Bond	2.69
			Glu266	H-Bond	1.93
			Thr294	H-Bond	1.76
			Asn295	H-Bond	1.96
	I7L-LD-2	-9.538	Leu239	H-Bond	2.22
			His241	Pi-cation	4.34
			Lys243	Salt-Bridge	3.16
			Asp258	Salt-Bridge	3.60
			Gly261	H-Bond	1.91
	I7L-LD-3	-9.340	Ans295	H-Bond	3.60
			Ser240	H-Bond	2.08
			Lys243	H-Bond	2.40
			Asp258	H-Bond	2.70
			Gly260	H-Bond	2.56
			Thr294	H-Bond	2.47
			Gln322	2 H-Bond	2.02, 2.39
			Leu323	2 H-Bond	1.74, 1.87
	I7L-LD-4	-8.411	Glu325	2 H-Bond	2.07, 2.55
Tyr238			H-Bond	2.14	
Ser240			H-Bond	2.73	
Lys243			H-Bond	2.43	
Glu266			H-Bond	1.80	
I7L-LD-5	-7.831	Asn295	H-Bond	1.86	
		Asp258	H-Bond	2.51	
		Asp258	Salt-Bridge	4.44	
		Gly261	2 H-Bond	1.68, 2.11	
		Glu266	Salt-Bridge	3.97	
		Asn295	H-Bond	2.02	
TTP-6171	-4.918	Gln322	H-Bond	2.67	
		Arg3	Salt-Bridge	2.76	
		His241	Pi-pi	5.06	
			Gly260	H-Bond	2.15

Table S6: Selected lead compounds with their SMILES codes.

Name of the Lead Compounds	SMILES
A48R-LD-1	<chem>C[C@H]1CCCC[N@H+]1CCNC(=O)c(cc2)cc(c23)[nH]c(=O)n(c3=O)CCCC</chem>
A48R-LD-2	<chem>[O-]C(=O)[C@@H]1[C@@H](O)[C@H](O)[C@@H](O)[C@@H](O1)Oc(c(cc2O)-c3cccc3)c(O)c2-c4cccc4</chem>
A48R-LD-3	<chem>c1cc(F)ccc1C(=O)Nc(c2)ccc(c2C([O-])=O)[N@@H+](C3)CCC[C@@H]3C[NH+]4CCCC4</chem>
A48R-LD-4	<chem>Cc1cc(=O)oc(c12)c(C)c(cc2)OC[C@@H](O)C[NH+](CC3)CCN3c4cccc4</chem>
A48R-LD-5	<chem>c1cccc1Cn(c(c23)cccc3)c(C([O-])=O)c2C[NH2+]CCC[N@H+](CC4)CC[C@@H]4Cc5cccc5</chem>
A50R-LD-1	<chem>c1[nH+]ccn1CCNC(=O)[C@H](CC2=O)CN2CCc3c[nH]c(c34)ccc(F)c4</chem>
A50R-LD-2	<chem>NC(=O)NCCC[C@@H](C([O-])=O)NC(=O)Cn(c1=O)c(=O)[nH]c(c12)cccc2</chem>
A50R-LD-3	<chem>c1cc(Cl)ccc1NCc2nn(c(=O)[nH]2)Cc3c(F)cccc3</chem>
A50R-LD-4	<chem>OCC1=C[C@@H]([C@@H](O)[C@H](O)[C@H]1O)[NH2+][C@@H](C2)[C@@H](O)[C@@H](O)[C@H]([C@@H]2CO)O[C@H](O3)[C@H](O)[C@@H](O)[C@H](O)[C@H]3CO[C@H](O4)[C@H](O)[C@@H](O)[C@H](O)[C@H]4CO</chem>
A50R-LD-5	<chem>CC[C@@H](C)[C@H](C([O-])=O)NC(=O)CCc1nc(no1)-c2cccc(c23)[nH]cc3</chem>
D13L-LD-1	<chem>COc1c(O)ccc(c1)[C@H]([C@H](O2)CO)Oc(c23)ccc(c3)[C@H]([C@H](C4=O)O)Oc(c45)cc(O)cc5O</chem>
D13L-LD-2	<chem>n1cccc1SCc2cc(=O)[nH]c(n2)Nc(nc3C)nc(c34)ccc(C)c4</chem>
D13L-LD-3	<chem>c1cccc1C(=O)c(c2)ccc(c2n(c34)cnn4)nc3Nc5c(O)cccc5</chem>
D13L-LD-4	<chem>o1cccc1C(=O)N2CCN(CC2)CC(=C3C(=O)OCC)NC(=O)N[C@H]3c(cc4)ccc4C</chem>
D13L-LD-5	<chem>Cc1ccc(cc1)SCC(=O)Cc2nc(nc(n2)N)Nc3c(OC)ccc(Cl)c3</chem>
F13L-LD-1	<chem>OCC1=C[C@@H]([C@@H](O)[C@H](O)[C@H]1O)[NH2+][C@@H](C2)[C@@H](O)[C@@H](O)[C@H]([C@@H]2CO)O[C@H](O3)[C@H](O)[C@@H](O)[C@H](O)[C@H]3CO[C@H](O4)[C@H](O)[C@@H](O)[C@H](O)[C@H]4CO</chem>
F13L-LD-2	<chem>C[C@@H](O)c1c([O-])oc(n1)Nc2ncnc(c23)n(cn3)[C@H](O4)[C@H](O)[C@H](O)[C@H]4CO</chem>
F13L-LD-3	<chem>NC(=O)C[C@H]([NH3+])[C@H](O)[C@H](O)C(=O)N[C@@H](CC(C)C)[C@@H](OC1=O)Cc(c12)cccc2O</chem>
F13L-LD-4	<chem>CCOCCNC(=O)CCN(c(c1c23)c(=O)[nH]nc1)c2cccc3</chem>
F13L-LD-5	<chem>c1cccc1C(=O)c(cc2C(=O)N)cc(N3C)c2Nc(c34)c(ccc4C(=O)N)SC[C@@H](C([O-])=O)NC(=O)C</chem>
I7L-LD-1	<chem>[O-]C(=O)C[C@H]([NH3+])[C@H](O)[C@H](O)C(=O)N[C@@H](CC(C)C)[C@@H](OC1=O)Cc(c12)cccc2O</chem>
I7L-LD-2	<chem>CCC[NH2+]Cc(n1)cc([O-])n(c12)nc(n2)NCc3c(Cl)cccc3</chem>
I7L-LD-3	<chem>OCC1=C[C@@H]([C@@H](O)[C@H](O)[C@H]1O)N[C@@H](C2)[C@@H](O)[C@@H](O)[C@H]([C@@H]2CO)O[C@H](O3)[C@H](O)[C@@H](O)[C@H](O)[C@H]3CO[C@H](O4)[C@H](O)[C@@H](O)[C@H](O)[C@H]4CO</chem>
I7L-LD-4	<chem>OC[C@@H]1[C@H]([C@@H](O)[C@H](O1)O)O[C@@H](O2)[C@@H]([C@H](O)[C@H]2CO)O[C@@H](O3)[C@H](O)[C@H](O)[C@H]3CO</chem>
I7L-LD-5	<chem>NC(=O)C[C@H]([NH3+])[C@H](O)[C@H](O)C(=O)N[C@@H](CC(C)C)[C@@H](OC1=O)Cc(c12)cccc2O</chem>

Data and software availability:

- The predicted structure from AlphaFold2 server, lead compound structures and structure of the MPXV target's known inhibitors are provided with GitHub (https://github.com/jivkiran007/MPXV_TBCGRL.git);
- Monkeypox virus's selected target protein sequences were retrieved from NCBI database (<https://www.ncbi.nlm.nih.gov/>);
- The 3D protein structure were modeled using DeepMind's AlphaFold2 server (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>);
- Ramachandran plot analysis done using SAVES v6.0 server (<https://saves.mbi.ucla.edu/>);
- Small molecule compounds were retrieved from the databases viz. ChemDiv database (<https://www.chemdiv.com/catalog/screening-libraries/>);
- DrugBank database (<https://go.drugbank.com/>), and
- PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>);
- Molecule visualization and DFT calculation done using BIOVIA Discovery Studio (Licensed software) (<https://www.3ds.com/products-services/biovia/>);
- Docking, HTVS and binding free energy were calculated using Schrodinger's software package (version 2022-3) (Licensed software) (<https://www.schrodinger.com/>);
- Molecular dynamics simulation performed using Desmond D.E. Shaw (Version 2020-1) (https://www.deshawresearch.com/downloads/download_desmond.cgi/);
- Essential dynamics analysis was performed using Schrodinger's script "trj_essential_dynamics.py" (<https://www.schrodinger.com/scriptcenter/>);
- Porcupine plots were generated using PyMol and "modevectors.py" script (<https://pymolwiki.org/index.php/Modevectors>);
- PCA plots were generated using OriginPro 2023 (Learning edition) (<https://www.originlab.com/index.aspx?go=Purchase/LicensingOptions&pid=923>).