Electronic Supplementary Information (ESI) for the Journal of Molecular Graphics and Modelling

Investigation of Structural Analogs of Hydroxychloroquine for SARS-CoV-2 Main Protease (Mpro): A Computational Drug Discovery Study

*Saima Reyaz[1#], Alvea Tasneem[1#], Gyan Prakash Rai[1#], and Hridoy R. Bairagya[2]**

[1]Department of Computer Science, Jamia Millia Islamia, New Delhi-110025, India

[2]Department of Biophysics, All India Institute of Medical Sciences, New Delhi-110029, India

**** Corresponding Author***:
Dr. Hridoy Ranjan Bairagya.
Tel: +91-8277145815
Email: hbairagya@gmail.com

[#]Authors (SR, AT, and GPR) contributed equally to this manuscript and shared first authorship.

SUPPLEMENTARY TABLES

Table S1. List of ligand-free crystal structures of Main Protease (Mpro) of SARS-CoV-2 and their different structural parameters.

S. No.	PDB ID	Classification	Method	Resolution (Å)	R-value free/ R-value work	SPACE GROUP	Entities/ Chains/ Sequence length	Refinement Methods	Unit Cell		References
									Length (Å)	Angle (∘)	
1	**6YB7** (Template)**	**Viral Protein**	**X-ray**	**1.25**	**0.192/0.171**	**C 1 2 1**	**1ab/A/306**	BUSTER	**a=112.42 b=52.79 c=44.61**	**α=90.0 β=103.0 γ=90.0**	[1]
2	6Y84	Viral Protein	X-ray	1.39	0.200/0.178	C 1 2 1	1ab/A/306	BUSTER	a=112.81 b=52.94 c=44.63	α=90.0 β=103.1 γ=90.0	[2]
3	7ALH	Viral Protein	X-ray	1.65	0.189/0.165	C 1 2 1	Mpro/A/306	PHENIX	a=115.21 b=54.18 c=44.78	α=90.0 β=101.17 γ=90.0	[3]
4	7ALI	Viral Protein	X-ray	1.65	0.217/0.189	C 1 2 1	Mpro/A/306	PHENIX	a=44.77 b=53.69 c=114.55	α=90.0 β=101.17 γ=90.0	[4]
5	7BRO	Viral Protein	X-ray	2.00	0.259/0.226	C 1 2 1	3-Clpro /A/307	PHENIX	a=113.43 b=54.25 c=44.95	α=90.0 β=100.67 γ=90.0	[5]
6	7C2Y	Viral Protein	X-ray	1.91	0.262/0.220	P 2_1 2_1 2_1	3-Clpro /A/298	PHENIX	a=67.88 b=102.34 c=103.49	α=90.0 β=90.0 γ=90.0	[6]
7	7CWB	Viral Protein	X-ray	1.90	0.257/0.218	C 1 2 1	3-Clpro /A/306	PHENIX	a=114.0 b=53.50 c=45.00	α=90.0 β=102.0 γ=90.0	[7]

8	7JR3	Hydrolase	X-ray	1.55	0.183/0.159	C 1 2 1	3-Clpro/A/306	PHENIX	a=98.99 b=80.99 c=51.74	α=90.0 β=114.7 γ=90.0	[8]
9	7JUN	Hydrolase	X-ray	2.30	0.220/0.185	C 1 2 1	3-Clpro /A/306	nCNS	a=114.92 b=54.68 c=45.21	α=90.0 β=101.46 γ=90.0	[9]
10	7JVZ	Viral Protein, Hydrolase	X-ray	2.50	0.217/0.172	C 1 2 1	3-Clpro /A/306	REFMAC	a=117.6 b=55.4 c=45.6	α=90.0 β=100.5 γ=90.0	[10]
11	7KPH	Viral Protein	X-ray	1.46	0.181/0.167	C 1 2 1	3-Clpro /A/306	PHENIX	a=112.99 b=52.84 c=44.74	α=90.0 β=102.8 γ=90.0	[11]

Table S2. Conserved water molecules (W1, W2 and W3) at catalytic region of Mpro protein and normalized B-factors with SASA values and crystallographic structural parameters (c.s.p) of those water molecules in ligand free X-ray structures.

Water ID	**PDB ID**	**Water molecule**	**Normalized B-factor (B_{Norm}) (Å^2)**	**SASA (Å^2)**	**Crystallographic Structural parameters (c.s.p.)**			
					Mean B-factor of Protein <Bi>	***Matthews coefficient***	***Solvent content***	***Ratio (N_{PROT} / N_{HOH})***
W1	6YB7	W639	5.07	2.6	16.87	1.91	35.50	5.34
	6Y84	W615	2.70	0.0	18.98	1.92	35.89	6.00
	7ALH	W573	0.95	20.4	33.22	2.03	39.32	19.43
	7ALI	W563	0.42	19.5	28.58	2.00	38.41	17.08
	7BRO	W499	0.63	17.6	31.12	2.01	38.70	19.04
	7JR3	W484	1.52	2.1	27.41	2.78	55.82	7.10
	7KPH	W629	1.12	1.1	24.40	1.93	36.11	14.71
W2	6YB7	W864	3.77	24.9	16.87	1.91	35.50	5.34
	6Y84	W800	3.89	19.3	18.98	1.92	35.89	6.00
	7KPH	W691	2.24	27.4	24.40	1.93	36.11	14.71
W3	7C2Y	W346	-0.78	0.0	33.56	2.83	56.48	20.93
	7CWB	W441	-1.20	0.0	42.62	1.98	38.01	21.37
	7JUN	W409	-1.54	0.0	39.97	2.06	40.25	35.87
	7JVZ	W429	-0.78	0.0	70.83	2.16	43.03	51.45

Abbreviation, SASA, Solvent accessible surface area, N_{PROT}, Number of atoms in protein, N_{HOH}, Number of atoms in solvent.

Table S3. Interactions of ligand HCQ and its structural analogs with the residues of Mpro protein and their binding energies in molecular docking studies.

S.no.	Ligand ID (Name)	H-bonding interaction (distances within 3.50 Å)	Short contacts (distances between 3.50 to 4.50 Å)	Binding Free energy (kcal/mol)
1	DB01611 (Hydroxychloroquine)	^{OB}T26, **^{ND1}H41,** ^{OG}S46	**^{SG}C145,** ^{NB}E166, ^{OE1}Q189	-5.80
2	DB00608 (Chloroquine)	^{OG1}T25, ^{OB}T26, ^{NB}E166	**^{ND1}H41,** ^{OB}C44, ^{OB}H164	-5.30
3	DB00353 (Methylergometrine)	^{OB}T24, ^{OG1}T25, ^{OB}C44, ^{NB}S46	^{OB}T45, **^{ND1}H41,** ^{OE1}Q189	-8.00
4	DB05039 (Indacaterol)	^{OG1}T25, **^{ND1}H41**, ^{NE2}H41, ^{NB}G143, ^{OB}H164	^{OB}T24, ^{NB}T25, ^{OB}T26, ^{SD}M49, **^{SG}C145**	-8.60
5	CHEMBL42527	**^{ND1}H41,** ^{OB}H164	^{NE2}H41, ^{OB}C44, ^{OG}S46, ^{OD1}N142, **^{SG}C145**, ^{NB}M165, ^{OE1}E166	-6.70
6	ZINC44136213	^{OG1}T25, ^{OB}F140, ^{OB}L141, ^{OB}G143, ^{OG}S144, **^{SG}C145,** ^{NE2}H163	^{OB}C44, **^{ND1}H41,** ^{OD1}N142, ^{OB}H164, ^{NB}M165, ^{NB}E166	-6.50
7	ZINC77376915	^{OB}E166, ^{OB}H164, **^{ND1}H41, ^{NE2}H41**	^{NB}P168, ^{OD1}A142, ^{SD}M49, **^{SG}C145**	-6.90
8	ZINC77520192	^{OD1}A142, ^{OB}E166	^{OG1}T25, **^{ND1}H41,** ^{OB}C44, ^{OB}T45, ^{NB}S46	-7.10
9	ZINC40444366	^{OB}C44, ^{OB}T45, ^{OB}H164	^{OG1}T25, **^{ND1}H41,** ^{SD}M49, **^{SG}C145,** ^{NB}M165, ^{NB}E166	-6.40
10	ZINC95367069	**^{ND1}H41**	^{NB}V42, ^{OB}E166, ^{OB}R188, ^{OE1}Q189, ^{NB}T190, ^{NE2}Q192	-7.40
11	ZINC28706440	^{OB}C44, ^{NB}E166, ^{OG}S144, **^{SG}C145**	^{OG1}T25, ^{OB}T45, **^{ND1}H41,** ^{NB}C44, ^{OG}S46, ^{SD}M49, ^{OB}L141, ^{NE2}H163, ^{OB}E166	-5.60
12	ZINC44136210	^{OB}T26, ^{OB}E166, ^{NE2}Q192, ^{OB}T190, ^{OB}R188	^{OG1}T25, **^{ND1}H41,** ^{OB}C44, **^{SG}C145,** ^{NB}G143, ^{NB}P168	-6.70
13	ZINC01706244	^{OB}C44, ^{OB}T45, ^{SD}M49, **^{ND1}H41,** ^{OB}M164	^{OG1}T25, ^{NB}T45, ^{NB}S46, ^{NE2}H41, ^{SG}C145, ^{OD1}N142, ^{OE1}E166	-6.20
14	ZINC95367068	^{OB}C44, ^{OB}T45, **^{SG}C145,** ^{OB}E166,	^{OG1}T25, ^{NB}S46, ^{SD}M49, **^{ND1}H41,** ^{NE2}H41, ^{NB}G143, ^{OE1}E189	-7.30
15	ZINC40412048	^{OB}T26, ^{NB}T26, ^{NE2}H41, **^{ND1}H41,**	^{OB}T24, ^{SD}M49, **^{SG}C145,** ^{OB}H164	-6.40

Table S4. Physicochemical properties of ligand HCQ and its structural analogs obtained from OSIRIS Property Explorer and Swiss ADME program.

S.No.	LIGAND ID	Molecular Formula	OSIRIS Property Explorer						Swiss ADME		
			cLogP	Solubility	Molweight (gm/mol)	TPSA	Drug likeness	Drug-score	Molar Refractivity	Lipinski	Bioavailability Score	Synthetic Accessibility
1	**DB01611 (HCQ)**	**$C_{18}H_{26}ClN_3O$**	**3.08**	**-3.55**	**335**	**48.39**	**6.54**	**0.48**	**98.57**	**YES**	**0.55**	**2.82**
2	DB00608 (Chloroquine)	$C_{18}H_{26}ClN_3$	4.01	-4.06	319	28.16	7.39	0.25	97.41	YES	0.55	2.76
3	DB00353 (Methylergometrine)	$C_{20}H_{25}N_3O_2$	1.33	-2.51	339	68.36	8.03	0.32	103.37	YES	0.55	4.36
4	DB05039 (Indacaterol)	$C_{24}H_{28}N_2O_3$	2.81	-4.42	392	81.59	2.12	0.66	116.83	YES	0.55	3.68
5	CHEMBL42527	$C_{24}H_{25}Cl_2N_5$	4.63	-6.18	453	61.87	1.73	0.22	132.03	YES	0.55	2.70
6	ZINC44136213	$C_{16}H_{20}ClN_3O_2$	2.06	-3.60	321	74.25	3.20	0.49	89.06	YES	0.55	2.53
7	ZINC77376915	$C_{19}H_{24}ClN_3O$	2.95	-3.40	345	28.60	4.69	0.48	106.02	YES	0.55	2.89
8	ZINC77520192	$C_{19}H_{24}ClN_3O$	3.06	-3.84	345	39.60	6.48	0.46	106.10	YES	0.55	2.53
9	ZINC40444366	$C_{15}H_{17}ClN_2O$	2.76	-3.70	276	36.36	1.79	0.46	82.17	YES	0.55	2.31
10	ZINC95367069	$C_{19}H_{24}ClN_3O$	3.06	-3.84	345	39.60	5.90	0.46	106.10	YES	0.55	3.31
11	ZINC28706440	$C_{14}H_{18}ClN_3O$	1.60	-2.97	279	57.18	1.74	0.50	79.25	YES	0.55	2.01
12	ZINC44136210	$C_{16}H_{20}ClN_3O_2$	2.06	-3.60	321	74.25	3.20	0.49	89.06	YES	0.55	2.53
13	ZINC01706244	$C_{16}H_{22}ClN_3O$	2.34	-3.45	307	57.18	1.75	0.47	88.90	YES	0.55	2.24
14	ZINC95367068	$C_{19}H_{24}ClN_3O$	3.06	-3.84	345	39.60	5.90	0.46	106.10	YES	0.55	3.31
15	ZINC40412048	$C_{14}H_{15}ClN_2O$	2.56	-3.65	262	36.36	5.02	0.50	77.36	YES	0.55	2.22

Abbreviation, cLogP, calculated Log P; Molweight, Molecular weight; TPSA, topological polar surface area

Table S5. Toxicity properties of ligand HCQ and its structural analogs obtained from OSIRIS Property Explorer.

S.No.	Ligand Id	Toxicity Risks			
		Mutagenic	**Tumorigenic**	**Irritant**	**Reproductive effective**
1	**DB01611** (Hydroxychloroquine)	**High**	Low	Low	Low
2	DB00608 (Chloroquine)	**High**	Low	**High**	Low
3	DB00353 (Methylergometrine)	Low	Low	**High**	**High**
4	DB05039 (Indacaterol)	Low	Low	Low	Low
5	CHEMBL42527	**High**	Low	Low	Low
6	ZINC44136213	**High**	Low	Low	Low
7	ZINC77376915	**High**	Low	Low	Low
8	ZINC77520192	**High**	Low	Low	Low
9	ZINC40444366	**High**	Low	Low	Low
10	ZINC95367069	**High**	Low	Low	Low
11	ZINC28706440	**High**	Low	Low	Low
12	ZINC44136210	**High**	Low	Low	Low
13	ZINC01706244	**High**	Low	Low	Low
14	ZINC95367068	**High**	Low	Low	Low
15	ZINC40412048	**High**	Low	Low	Low

Table S6: The short contacts (within 4.50 Å) of ligand HCQ, indacterol, and ZINC28706440 with Mpro protein in MD structures

S.no	Ligands	H-bond Interaction (within 4.5 Å)
1	HCQ	^{NE2}H41, ^{SG}C145, ^{OB}H164, ^{OB}V186,^{NB}R188, ^{OE1}Q189
2	Indacterol	^{ND1}H41, ^{OG1}S46, ^{OE1}E47, ^{ND1}N142, ^{NE2}Q189
3	ZINC28706440	^{NE2}H41, ^{NE2}H163, ^{NB}H164, ^{OE1}E166, ^{OB}R188,^{NE2}Q189

Table S7: Ligands from molecular docking and MD-Simulated structures not occupied the W1, W2, and W3 water molecules in the 6YB7 and 7CWB crystal structures. Distances (Å) between the atom of each ligand to W1, W2, and W3 water molecules of 6YB7 and 7CWB crystal structures are mentioned.

S.No.	Ligand	Structures	Distances (Å) between an atom of ligand to W1 and W2 position of 6YB7 and W3 position of 7CWB crystal structure		
			W1 (W639)	**W2** (W864)	**W3** (W441)
1	HCQ	MD Simulation	*-	$C8^{HCQ}$ (1.23)	-
		Docking	-	$C11^{HCQ}$ (1.44)	*NA
2	Indacaterol	MD Simulation	-	-	NA
		Docking	-	$C22^{Indacaterol}$ (0.82)	NA
3	ZINC28706440	MD Simulation	-	-	NA
		Docking	$O1^{ZINC28706440}$ (0.72)	$C2^{ZINC28706440}$ (0.97)	NA

*'-', Not Occupied, NA, Not Available

SUPPLEMENTARY FIGURES

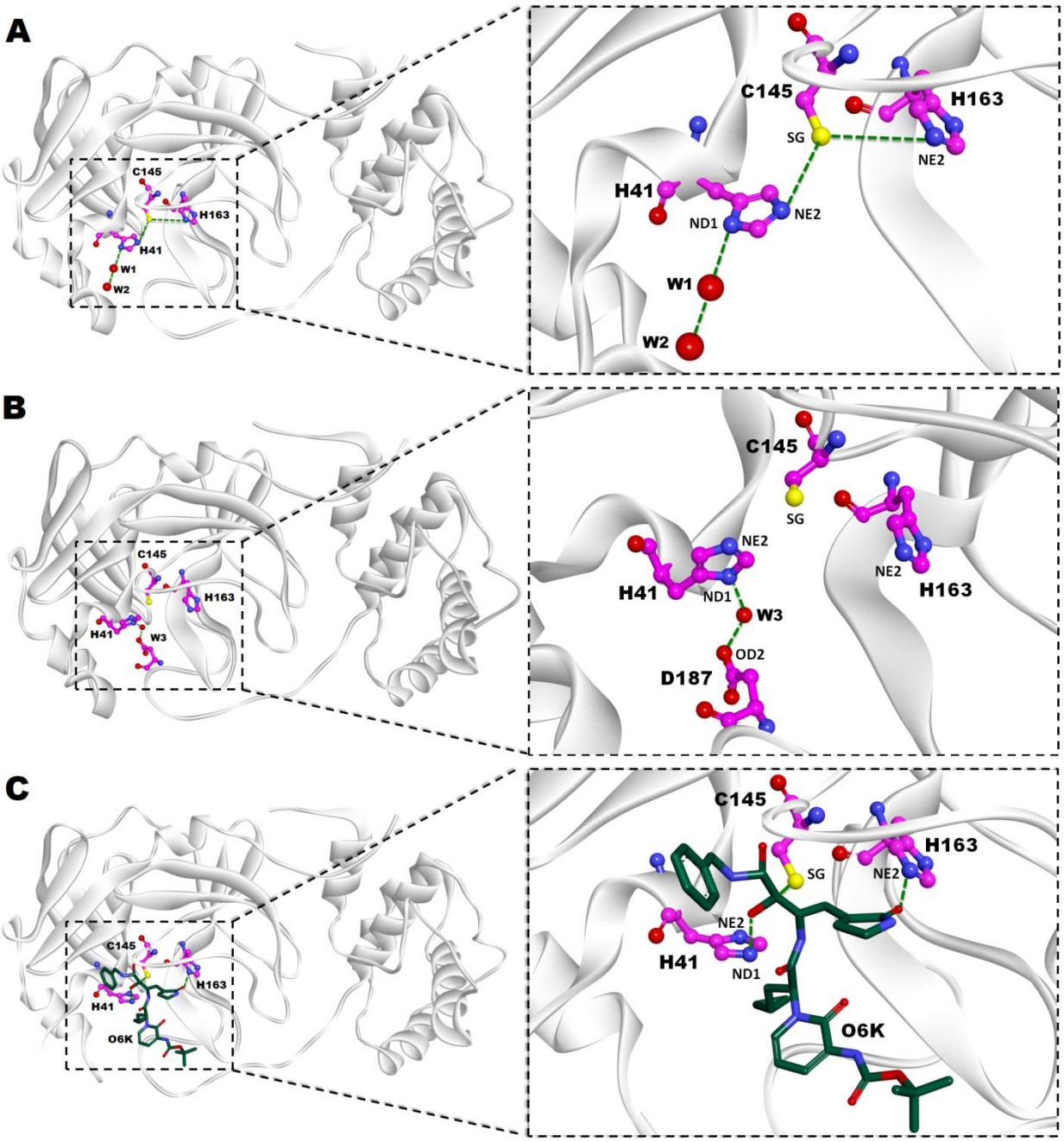

Fig. S1: H-bonding interactions between catalytic residues (His41, Cys145, and His163) of Mpro and water molecules (W1, W2, and W3) are shown as dotted lines in **(A)** native conformation of PDB ID: 6YB7, **(B)** native conformation of PDB ID: 7JUN, and **(C)** ligand-bound conformation of PDB ID: 6Y2F. Upper right panel indicates the recognition of W2---W1---His41---Cys145---His163. Middle panel indicates the recognition of D187---W3---His41. Lower right panel represents the interactions of His41---ligand (PDB name: O6K) ---Cys145 and His41---ligand (PDB name: O6K)---His163 at ligand-binding pocket in crystal structure of 6Y2F.

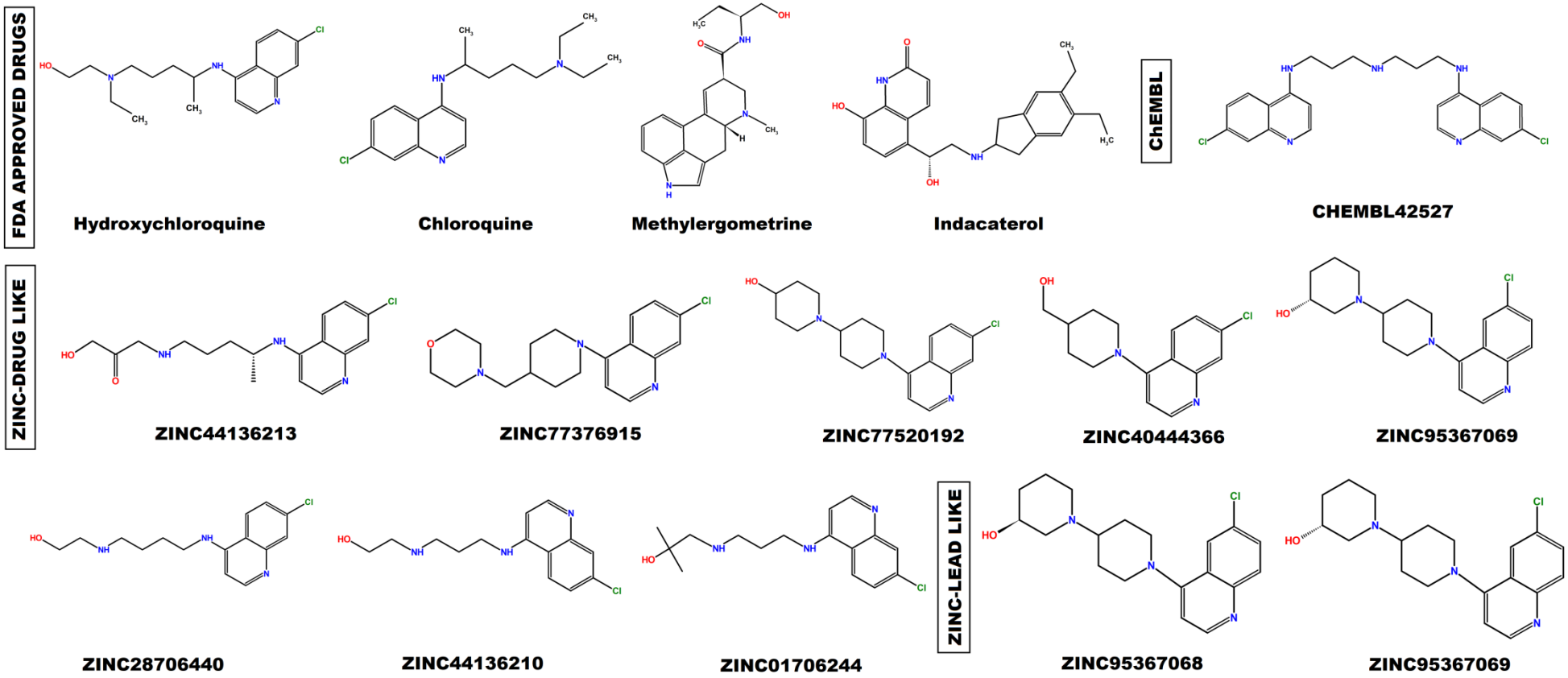

Fig. S2: Schematic presentations of the chemical structures of Hydroxychloroquine and its analogs that were obtained from FDA approved drugs, ChEMBL, Zinc-Drug like, and Zinc-Lead like databases.

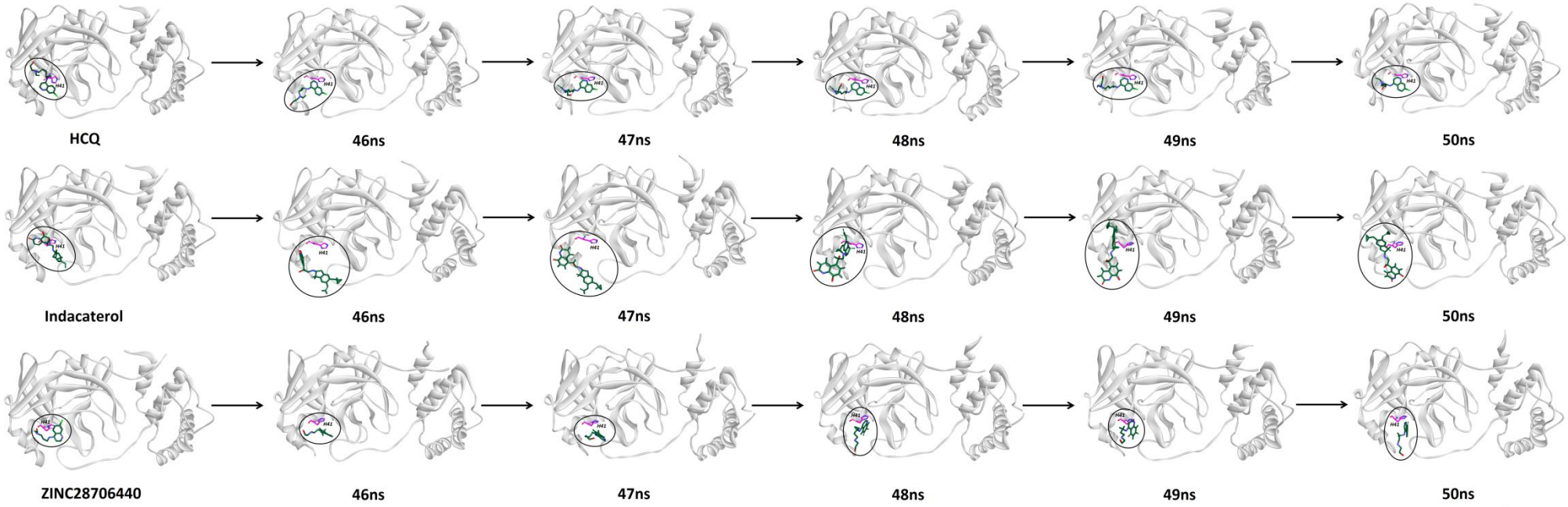

Fig. S3: Conformational change and binding position of each ligand with respect to His41 during MD simulation of HCQ (upper panel), Indacaterol (middle), and ZINC28706440 (lower panel) bound-MD structures in 6YB7.

References:

1. Owen C.D., Lukacik P., Strain-Damerell C.M., Douangamath A., Powell A.J., Fearon D., et al., SARS-CoV-2 main protease with unliganded active site (2019-nCoV, coronavirus disease 2019, COVID-19), Worldwide Protein Data Bank, (2020), http://dx.doi.org/10.2210/pdb6yb7/pdb.
2. Owen C.D., Lukacik P., Strain-Damerell C.M., Douangamath A., Powell A.J., Fearon D., et al., SARS-CoV-2 main protease with unliganded active site (2019-nCoV, coronavirus disease 2019, COVID-19), Worldwide Protein Data Bank, (2020), http://dx.doi.org/10.2210/pdb6y84/pdb.
3. Costanzi E., Demitri N., Giabbai B., Heroux A., Storici P., Crystal structure of the main protease (3CLpro/Mpro) of SARS-CoV-2 at 1.65A resolution (spacegroup C2), Worldwide Protein Data Bank, (2020), http://dx.doi.org/10.2210/pdb7alh/pdb.
4. Costanzi E., Demitri N., Giabbai B., Heroux A., Storici P., Crystal structure of the main protease (3CLpro/Mpro) of SARS-CoV-2 at 1.65A resolution (spacegroup P2(1)), Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7ali/pdb.
5. Fu L.F., Crystal structure of the 2019-nCoV main protease, Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7bro/pdb.
6. Zhou X.L., Zhong F.L., Lin C., Zhou H., Hu X.H., Wang Q.S., et al., The crystal structure of COVID-2019 main protease in the apo state, Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7c2y/pdb.
7. DeMirci H., Ambient-Temperature Serial Femtosecond X-ray Crystal structure of SARS-CoV-2 Main Protease at 1.9 A Resolution (C121), Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7cwb/pdb.
8. De Oliveira R.R., Nascimento A.F.Z., Zeri A.C.M., Trivella D.B.B, SARS-CoV-2 3CL protease crystallized under reducing conditions, Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7jr3/pdb.
9. Kneller D.W., Phillips G., Weiss K.L., Pant S., Zhang Q., O'Neill H.M., Coates L., Kovalevsky A., Unusual zwitterionic catalytic site of SARS–CoV-2 main protease revealed by neutron crystallography, Journal of Biological Chemistry, 295(2020),17365-73, https://doi.org/10.1074/jbc.AC120.016154.
10. Worldwide Protein Data Bank, http://dx.doi.org/10.2210/pdb7jvz/pdb.
11. Noske G.D., Nakamura A.M., Gawriljuk V.O., Lima G.M.A., Zeri A.C.M., Nascimento A.F.Z., et al., SARS-CoV-2 Main Protease in mature form, Worldwide Protein Data Bank, 2020, http://dx.doi.org/10.2210/pdb7kph/pdb.