Title - Whole Genome Sequence Analysis and Homology Modelling of Main Protease and Non-Structural Protein 3 of the SARS-CoV-2 reveals an Aza-Peptide and a Lead Inhibitor with Possible Antiviral Properties

Authors - Arun K. Shanker, Divya Bhanu, Anjani Alluri and Samriddhi Gupta

Journal – New Journal of Chemistry

Supplementary Material

Supplementary Fig.1 Multiple alignment of the Polyproteins of Severe acute respiratory syndrome coronavirus 2 isolate virus isolates

29

30

31

Supplementary Table 1 List of Severe Acute Respiratory Syndrome coronavirus 2 isolate sequences taken for bioinformatic analysis

Genbank Accession Number	Title	Description
MN988713.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-IL1/2020	Complete genome
MN938384.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-002a_2020	Complete genome
MN975262.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-005b_2020	Complete genome
MN985325.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-WA1/2020	Complete genome
NC_045512.2	Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1	Complete genome
MN908947.3	Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1	Complete genome
MN938385.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-001_2020 ORF1ab polyprotein, RdRp region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN938386.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-004_2020 ORF1ab polyprotein, RdRp region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN975263.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007a_2020 ORF1ab polyprotein, RdRp region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN975264.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007b_2020 ORF1ab polyprotein, RdRp region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN975265.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007c_2020 ORF1ab polyprotein, RdRp region, (orf1ab) gene, partial cds	Polyprotein, RdRp region

MN970003.1	Severe acute respiratory syndrome coronavirus 2 isolate SI200040-SP orf1ab polyprotein, RdRP region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN970004.1	Severe acute respiratory syndrome coronavirus 2 isolate SI200121-SP orf1ab polyprotein, RdRP region, (orf1ab) gene, partial cds	Polyprotein, RdRp region
MN938387.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-001_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN938388.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-002b_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN938389.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-004_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN938390.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-005_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN975266.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007a_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN975267.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007b_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein
MN975268.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-007c_2020 surface glycoprotein (S) gene, partial cds	Glycoprotein

Supplementary Table 2 Comparison of binding properties of Novel Coronavirus protein from region 3268-3573 (2a5i _SARS-CoV-2) and 2a5i template to ligand AZP

Hydrophobic Interactions

Index	Residue		AA		Distance		Ligand Atom		Protein Atom	
	2a5i SARS-CoV-2	2a5i	2a5i SARS-CoV-2	2a5i	2a5i SARS-CoV-2	2a5i	2a5i SARS-CoV-2	2a5i	2a5i SARS-CoV-2	2a5i
1	3308A	41A	HIS	HIS	3.64	3.70	2393	2461	307	308
2	3316A	49A	MET	MET	3.81	3.86	2395	2463	368	368
3	3435A	168A	PRO	PRO	3.42	3.73	2376	2443	1303	1347
4	3456A	189A	GLN	GLN	3.84	3.93	2396	2464	1462	1507

Hydrogen Bonds

Index	Residue		AA		Distance H-A		Distance D-A		Donor Angle		Protein donor?		Sidechain		Donor Atom		Acceptor Atom	
	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i
1	3407A	140A	PHE	PHE	2.61	2.46	3.47	3.33	146.60	147.28	✗	✗	✗	✗	2404 [Nam]	2472 [Nam]	1081 [O2]	1112 [O2]
2	3409A	142A	ASN	ASN	2.51	2.52	2.87	2.87	102.18	100.93	✗	✗	✓	✓	2410 [O3]	2478 [O3]	1103 [O2]	1134 [O2]

3	3410A	143A	GLY	GLY	1.94	1.83	2.78	2.73	142.36	150.57	✓	✓	✗	✗	1105 [Nam]	1136 [Nam]	2407 [O2]	2475 [O2]
4	3411A	144A	SER	SER	3.37	3.44	3.76	3.80	106.27	104.79	✓	✓	✓	✓	1114 [O3]	1145 [O3]	2405 [O2]	2473 [O2]
5	3411A	144A	SER	SER	2.60	2.61	3.24	3.20	122.58	118.23	✓	✓	✗	✗	1109 [Nam]	1140 [Nam]	2407 [O2]	2475 [O2]
6	3412A	145A	CYS	CYS	2.50	2.57	3.39	3.37	150.10	137.96	✓	✓	✗	✗	1115 [Nam]	1146 [Nam]	2407 [O2]	2475 [O2]
7	3431A	164A	HIS	HIS	1.70	1.73	2.63	2.67	156.72	157.31	✗	✗	✗	✗	2399 [Nam]	2467 [Nam]	1266 [O2]	1307 [O2]
8	3433A	166A	GLU	GLU	2.08	2.01	3.04	2.97	165.52	163.22	✓	✓	✗	✗	1281 [Nam]	1325 [Nam]	2387 [O2]	2455 [O2]
9	3433A	166A	GLU	GLU	2.06	2.13	2.88	2.92	140.23	135.67	✗	✗	✗	✗	2370 [Nam]	2438 [Nam]	1284 [O2]	1328 [O2]
10	3456A	189A	GLN	GLN	1.90	1.82	2.84	2.77	158.82	161.93	✗	✗	✓	✓	2388 [Nam]	2456 [Nam]	1464 [O2]	1509 [O2]

Salt Bridges

Index	Residue		AA		Distance		Protein positive?		Ligand Group		Ligand Atoms	
	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i
1	3308A	41A	HIS	HIS	5.10	5.08	✔	✔	Carboxylate	Carboxylate	2412, 2413	2481, 2480

Water Bridges

Index	Residue		AA		Dist. A-W		Dist. D-W		Donor Angle		Water Angle		Protein donor?		Donor Atom		Acceptor Atom		Water Atom	
	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i	2a5i _SARS-CoV-2	2a5i
1	189A	-	GLN	-	4.07	-	3.93	-	115.60	-	91.13	-	✔	-	1510 [Nam]	-	2456 [Nam]	-	2543	-

Supplementary Table 3 Comparison of binding properties of Novel Coronavirus protein from region 1568-1882 (3e9s_SARS-CoV-2) and 3e9s template to ligand Small molecule Noncovalent Lead Inhibitor

Hydrophobic Interactions

Index	Residue		AA		Distance		Ligand Atom		Protein Atom	
	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s
1	1731A	165A	ASP	ASP	3.82	3.83	2502	2504	1308	1320
2	1814A	248A	PRO	PRO	3.79	3.80	2501	2503	1955	1952
3	1815A	249A	PRO	PRO	3.52	3.75	2503	2505	1963	1960
4	1831A	265A	TYR	TYR	3.50	3.57	2503	2505	2090	2089
5	1831A	265A	TYR	TYR	3.63	3.67	2504	2506	2091	2090
6	1835A	-	TYR	-	3.62	-	2502	-	2121	-
7	1836A	270A	GLN	GLN	3.60	3.59	2502	2504	2130	2129
8	1836A	270A	GLN	GLN	3.58	3.62	2507	2509	2130	2129
9	1868A	302A	THR	THR	3.37	3.54	2514	2516	2381	2385

Hydrogen Bonds

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Sidechain	Donor Atom	Acceptor Atom

	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s
1	1731A	165A	ASP	ASP	2.12	2.19	2.98	3.03	151.51	149.71	✔	✔	✔	✔	1311 [O3]	1323 [O3]	2512 [Nam]	2514 [Nam]
2	1731A	165A	ASP	ASP	2.00	2.05	2.98	3.03	173.11	175.80	✖	✖	✔	✔	2512 [Nam]	2514 [Nam]	1311 [O3]	1323 [O3]
3	1835A	269A	TYR	TYR	2.78	2.81	3.61	3.64	143.08	142.33	✖	✖	✔	✔	2511 [Npl]	2513 [Npl]	2124 [O3]	2123 [O3]
4	1836A	270A	GLN	GLN	1.87	1.77	2.83	2.75	164.21	174.65	✔	✔	✖	✖	2125 [Nam]	2124 [Nam]	2509 [O2]	2511 [O2]

π-Stacking

Index	Residue		AA		Distance		Angle		Offset		Type		Ligand Atoms	
	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s	3e9s_SARS-CoV-2	3e9s

1	1835A	269A	TYR	TYR	5.30	5.31	84.17	85.46	1.70	1.70	T	T	2497, 2500, 2503, 2506, 2508, 2516	2499, 2502, 2505, 2508, 2510, 2518
2	1835A	269A	TYR	TYR	5.09	5.10	84.10	85.40	0.83	0.85	T	T	2497, 2498, 2500, 2501, 2515, 2517	2499, 2500, 2502, 2503, 2517, 2519
3	-	269A	-	TYR	-	5.18	-	73.97	-	1.96	-	T	-	2498, 2501, 2504, 2507, 2509, 2512