

Supporting Information

Production of Proteins of SARS-CoV-2 Proteome for Drug Discovery

*Choon Kim, Kiran V. Mahasenan, Atul Bhardwaj, Olaf Wiest, Mayland Chang, and Shahriar
Mobashery**

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556,
USA

Table S1. Synthetic SARS-CoV-2 genes.

No.	Gene (peptide)	Total AA	Total length	Comment ^a
1	<i>nsp1</i> (G7-N178)	230 aa	701 bp	5'-NcoI-6xHis-2xStrep-TEV-NdeI- <i>nsp1</i> -XhoI-3'; Removed <i>N</i> -terminal disorder
2	<i>nsp2</i> (A1-G638)	639 aa	1929 bp	Full-length
3	<i>nsp3</i> (E206-N1410)	1206 aa	3630 bp	Removed <i>N</i> -terminal disorder and <i>C</i> -terminal 4 transmembrane (TM) regions
4	<i>nsp3N</i> (E206-E374)	169 aa	519 bp	ADP-ribosylhydrolase (ARH) domain of NSP3
5	<i>nsp3P</i> (E746-P1061)	317 aa	963 bp	PL ^{pro} domain of NSP3
6	<i>nsp3C</i> (D1584-G1945)	363 aa	1101 bp	<i>C</i> -terminal domain of NSP3
7	<i>nsp4N</i> (K1-D279)	247 aa	753 bp	<i>N</i> -terminal part; removed <i>N</i> - & <i>C</i> -terminal TM regions
8	<i>nsp4C</i> (S386-Q500)	116 aa	360 bp	<i>C</i> -terminal part after 4 TM regions
9	<i>nsp5</i> (S1-Q306)	307 aa	933 bp	Full-length
10	<i>nsp7</i> (S1-Q83)	84 aa	264 bp	Full-length
11	<i>nsp8</i> (A1-Q198)	199 aa	609 bp	Full-length
12	<i>nsp9</i> (N1-Q113)	114 aa	354 bp	Full-length
13	<i>nsp10</i> (A1-Q139)	140 aa	432 bp	Full-length
14	<i>nsp12</i> (S1-Q932)	933 aa	2811 bp	Full-length
15	<i>nsp12P</i> (Q117-E919)	804 aa	2424 bp	Q117-E919 in SARS-CoV crystalized
16	<i>nsp13</i> (A1-Q601)	602 aa	1818 bp	Full-length
17	<i>nsp14</i> (A1-Q527)	528 aa	1596 bp	Full-length
18	<i>nsp15</i> (S1-A346)	347 aa	1053 bp	Full-length
19	<i>nsp16</i> (S1-N298)	299 aa	909 bp	Full-length
20	<i>S</i> (Q14-K1211)	1199 aa	3609 bp	Removed signal peptide and <i>C</i> -terminal TM region
21	<i>S1</i> (Q14-R685)	673 aa	2031 bp	S1 domain; removed signal peptide
22	<i>S2</i> (S686-K1211)	527 aa	1593 bp	S2 domain; removed <i>C</i> -terminal TM region
23	<i>orf3a</i> (M125-L275)	151 aa	465 bp	Removed <i>N</i> -terminal 3 TM regions
24	<i>M</i> (R101-N203)	104 aa	324 bp	Removed <i>N</i> -terminal TM region
25	<i>orf7a</i> (E16-L96)	82 aa	258 bp	Removed signal peptide and <i>C</i> -terminal TM region
26	<i>orf8</i> (M1-I121)	121 aa	375 bp	Full-length
27	<i>N</i> (M1-A419)	419 aa	1269 bp	Full-length; highly disordered structure; phosphoprotein
28	<i>NS</i> (T49-D358)	311 aa	945 bp	Removed <i>N</i> - & <i>C</i> -terminal disorders
29	<i>orf9b</i> (M1-K97)	97 aa	303 bp	Full-length

^a Each construct except for *nsp1* contains NdeI and XhoI restriction sites at the 5'-end and the 3'-end, respectively.; a stop codon is added in each construct.

Table S2. Grouping of the compounds from virtual screening against the ARH domain of NSP3.

Group	Number	Reference ID	MW	Group	Number	Reference ID	MW
1	1	MolPort-002-285-607	326.33	3	21	MolPort-005-816-640	329.39
	2	MolPort-021-778-386	336.40		22	MolPort-023-233-580	340.39
	3	MolPort-000-123-877	291.33		23	MolPort-020-131-668	315.35
	4	MolPort-038-417-546	342.38		24	MolPort-028-804-245	284.28
	5	MolPort-035-761-911	333.35		25	MolPort-023-169-993	313.34
	6	MolPort-042-569-276	340.36		26	MolPort-009-143-528	349.36
	7	MolPort-039-034-165	341.42		27	MolPort-010-194-066	312.34
	8	MolPort-021-772-677	345.60		28	MolPort-029-915-393	344.32
	9	MolPort-002-005-643	338.81		29	MolPort-023-153-919	329.10
	10	MolPort-007-572-698	320.28		30	MolPort-009-583-655	319.36
2	11	MolPort-004-919-118	343.40	4	31	MolPort-020-046-308	344.44
	12	MolPort-009-567-089	340.39		32	MolPort-027-660-994	345.42
	13	MolPort-020-092-666	312.33		33	MolPort-004-103-984	324.38
	14	MolPort-019-642-212	315.30		34	MolPort-020-102-008	337.38
	15	MolPort-009-321-750	298.36		35	MolPort-005-716-238	303.27
	16	MolPort-023-149-743	345.29		36	MolPort-004-005-399	343.29
	17	MolPort-009-641-058	314.39		37	MolPort-028-955-822	298.31
	18	MolPort-009-020-554	321.34		38	MolPort-023-181-524	313.36
	19	MolPort-004-142-938	292.29		39	MolPort-004-273-745	304.28
	20	MolPort-023-149-032	300.32		40	MolPort-020-053-080	349.39

Table S2. Continued.

Group	Number	Reference ID	MW	Group	Number	Reference ID	MW
5	41	MolPort-005-315-087	315.33	7	61	MolPort-020-160-315	343.31
	42	MolPort-019-645-074	315.39		62	MolPort-003-334-648	321.33
	43	MolPort-009-542-746	325.37		63	MolPort-009-545-284	274.28
	44	MolPort-019-667-806	299.29		64	MolPort-023-143-503	310.40
	45	MolPort-004-186-478	333.35		65	MolPort-009-074-353	323.36
	46	MolPort-004-217-503	347.43		66	MolPort-023-239-714	299.29
	47	MolPort-020-095-499	322.34		67	MolPort-019-683-071	336.36
	48	MolPort-009-576-607	341.39		68	MolPort-005-708-483	319.32
	49	MolPort-009-125-249	338.37		69	MolPort-027-655-284	336.35
	50	MolPort-020-097-406	309.28		70	MolPort-028-801-666	294.28
6	51	MolPort-009-539-431	312.34	8	71	MolPort-009-321-655	342.37
	52	MolPort-019-666-473	331.34		72	MolPort-023-233-484	307.35
	53	MolPort-009-448-623	283.33		73	MolPort-023-190-334	311.34
	54	MolPort-030-022-102	308.32		74	MolPort-023-188-618	305.76
	55	MolPort-005-314-867	344.39		75	MolPort-028-604-292	349.39
	56	MolPort-027-709-051	320.40		76	MolPort-044-431-079	327.34
	57	MolPort-023-139-905	323.35		77	MolPort-039-058-480	312.33
	58	MolPort-019-675-323	328.70		78	MolPort-009-650-513	325.37
	59	MolPort-023-174-763	340.32		79	MolPort-023-295-049	303.38
	60	MolPort-009-473-445	341.33		80	MolPort-023-295-351	321.34

Table S2. Continued.

Group	Number	Reference ID	MW	Group	Number	Reference ID	MW
9	81	MolPort-008-349-576	346.42	11	101	MolPort-005-973-205	341.37
	82	MolPort-042-671-819	349.39		102	MolPort-000-849-801	330.73
	83	MolPort-008-009-694	293.33		103	MolPort-003-008-282	300.33
	84	MolPort-005-980-024	343.43		104	MolPort-000-804-173	321.38
	85	MolPort-009-011-201	311.34		105	MolPort-005-988-009	344.39
	86	MolPort-006-828-377	298.32		106	MolPort-027-846-272	326.32
	87	MolPort-002-291-055	334.29		107	MolPort-001-904-777	327.38
	88	MolPort-019-952-145	299.33		108	MolPort-019-795-368	314.31
	89	MolPort-002-289-145	349.41		109	MolPort-000-835-912	333.35
	90	MolPort-003-006-209	255.23		110	MolPort-001-625-946	343.36
10	91	MolPort-000-664-093	334.11	12	111	MolPort-004-856-353	299.71
	92	MolPort-002-952-102	339.35		112	MolPort-000-813-761	339.31
	93	MolPort-019-694-241	316.32		113	MolPort-006-809-859	308.40
	94	MolPort-002-574-436	294.33		114	MolPort-028-604-154	337.40
	95	MolPort-003-000-526	329.27		115	MolPort-000-773-077	309.32
	96	MolPort-044-431-605	332.36		116	MolPort-038-429-727	335.37
	97	MolPort-002-665-215	327.31		117	MolPort-006-805-445	338.32
	98	MolPort-028-604-773	344.37		118	MolPort-006-808-970	308.34
	99	MolPort-005-919-460	300.32		119	MolPort-038-430-484	337.43
	100	MolPort-005-308-515	316.32				

Table S2. Continued.

Group	Number	Reference ID	MW	Group	Number	Reference ID	MW
13	121	MolPort-021-767-645	300.32	15	139	MolPort-044-809-313	354.41
	122	MolPort-029-903-833	367.45		140	MolPort-000-782-326	387.41
	123	MolPort-007-606-006	325.21		141	MolPort-044-809-457	329.37
	124	MolPort-007-06-035	381.44		142	MolPort-000-841-069	353.33
	125	MolPort-027-694-345	348.45		143	MolPort-044-431-816	360.80
	126	MolPort-009-564-626	297.36		144	MolPort-000-827-816	353.33
	127	MolPort-035-704-162	337.43		145	MolPort-044-809-742	376.82
	128	MolPort-035-704-312	337.43		146	MolPort-044-431-634	360.80
	129	MolPort-035-702-334	365.48		147	MolPort-000-838-047	351.79
14	130	MolPort-029-882-857	309.37	16	148	MolPort-044-809-270	325.37
	131	MolPort-029-882-697	351.35		149	MolPort-0000-814-999	271.28
	132	MolPort-035-703-846	337.43		150	MolPort-044-431-643	352.44
	133	MolPort-029-884-837	295.35		151	MolPort-044-809-282	405.25
	134	MolPort-029-883-430	337.43		152	MolPort-044-809-847	353.43
	135	MolPort-029-883-647	323.40		153	MolPort-000-820-216	328.33
	136	MolPort-029-883-157	309.37		154	MolPort-044-809-526	369.38
	137	MolPort-039-059-996	357.42		155	MolPort-044-809-269	325.37
	138	MolPort-002-510-495	355.31		156	MolPort-000-827-950	353.33

Table S2. Continued.

Group	Number	Reference ID	MW	Group	Number	Reference ID	MW
17	157	MolPort-044-809-976	426.41	19	175	MolPort-000-835-972	331.38
	158	MolPort-000-814-712	331.33		176	MolPort-005-977-370	346.35
	159	MolPort-000-810-604	427.45		177	MolPort-000-817-103	311.34
	160	MolPort-044-809-279	354.41		178	MolPort-000-825-694	362.35
	161	MolPort-044-431-637	358.83		179	MolPort-044-431-633	358.83
	162	MolPort-044-809-460	357.43		180	MolPort-044-809-850	374.39
	163	MolPort-000-843-327	347.37		181	MolPort-044-431-632	368.44
	164	MolPort-000-791-060	443.90		182	MolPort-000-816-837	313.36
	165	MolPort-000-829-044	347.37		183	MolPort-000-841-359	317.35
18	166	MolPort-000-835-912	333.35	20	184	MolPort-044-431-630	370.41
	167	MolPort-044-809-461	358.37		185	MolPort-044-809-856	349.36
	168	MolPort-000-831-303	363.37		186	MolPort-000-851-044	271.24
	169	MolPort-000-817-289	307.71		187	MolPort-000-832-070	353.33
	170	MolPort-044-809-977	406.41		188	MolPort-044-809-263	358.83
	171	MolPort-000-838-658	333.35		189	MolPort-044-431-639	403.28
	172	MolPort-044-431-625	368.44		190	MolPort-044-809-845	436.44
	173	MolPort-044-431-627	417.31		191	MolPort-044-809-257	368.44
	174	MolPort-000-836-247	377.40		192	MolPort-044-809-957	354.41

Table S2. Continued.

Group	Number	Reference ID	MW
21	193	MolPort-000-845-338	343.34
	194	MolPort-000-823-841	335.34
	195	MolPort-000-823-182	299.33
	196	MolPort-044-431-646	356.40
	197	MolPort-044-809-452	331.37
	198	MolPort-044-431-644	338.41
	199	MolPort-002-600-674	327.34
	200	MolPort-002-985-610	285.31
	201	MolPort-002-143-318	357.33
	202	MolPort-002-025-637	361.36

Table S3. Binding check of the compounds in groups 10 and 12 to the ARH domain of NSP3 by MST.

Number	Reference ID	S/N ratio ^a	Number	Reference ID	S/N ratio ^a
91	MolPort-000-664-093	0.8	111	MolPort-004-856-353	4.4
92	MolPort-002-952-102	3.4	112	MolPort-000-813-761	7.5
93	MolPort-019-694-241	3.6	113	MolPort-006-809-859	1.0
94	MolPort-002-574-436	1.3	114	MolPort-028-604-154	5.4
95	MolPort-003-000-526	12.7	115	MolPort-000-773-077	1.3
96	MolPort-044-431-605	14.8	116	MolPort-038-429-727	5.9
97	MolPort-002-665-215	1.0	117	MolPort-006-805-445	1.9
98	MolPort-028-604-773	7.8	118	MolPort-006-808-970	1.7
99	MolPort-005-919-460	1.1	119	MolPort-038-430-484	13.5
100	MolPort-005-308-515	4.7			

^a The signal-to-Noise (S/N) ratio higher than 5 in MST indicates that a ligand binds to a target.

The compounds showing the binding are in bold. The compounds exhibiting the high S/N ratio are indicated in gray shades.

Table S4. Thirteen out of 82 compounds in the second set (groups from 13 to 21 in Table S2) exhibiting the interaction to the ARH domain of NSP3.

Number	Reference ID	S/N ratio ^a	Number	Reference ID	S/N ratio ^a
121	MolPort-021-767-645	6.9	138	MolPort-002-510-495	9.6
122	MolPort-029-903-833	5.4	147	MolPort-000-838-047	13.6
123	MolPort-007-606-006	18.8	173	MolPort-044-431-627	41.9
127	MolPort-035-704-162	7.8	175	MolPort-000-835-972	13.3
128	MolPort-035-704-312	13.8	190	MolPort-044-809-845	6.9
129	MolPort-035-702-334	7.0	200	MolPort-002-985-610	7.1
130	MolPort-029-882-857	13.3			

^a The signal-to-Noise (S/N) ratio higher than 5 in MST indicates that a ligand binds to a target.

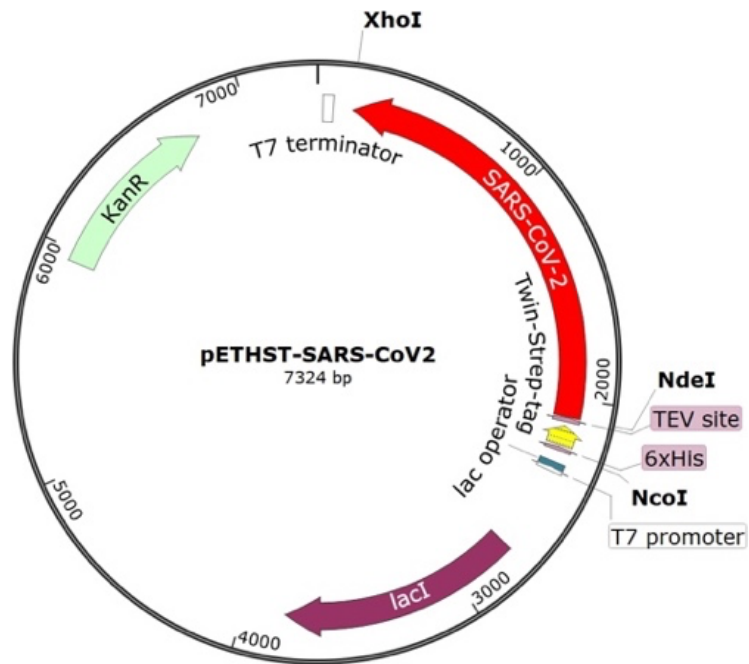


Figure S1. Schematic of the pETHST-SARS-CoV-2 plasmid. The thrombin cleavage site and the T7 tag in the vector pET28a (+) were replaced to the Twin-Strep tag and the TEV cleavage site to generate the expression vector pETHST.

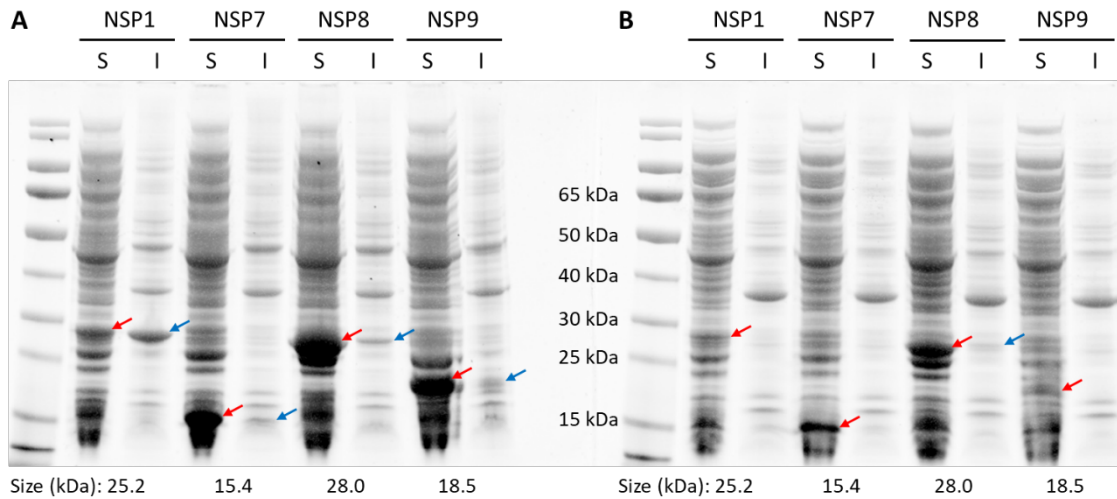


Figure S2-1. SDS-PAGE for expression of SARS-CoV-2 NSP1, NSP7, NSP8 and NSP9. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (A) and with 0.4 mM of IPTG for 18 hours at 16 °C (B). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

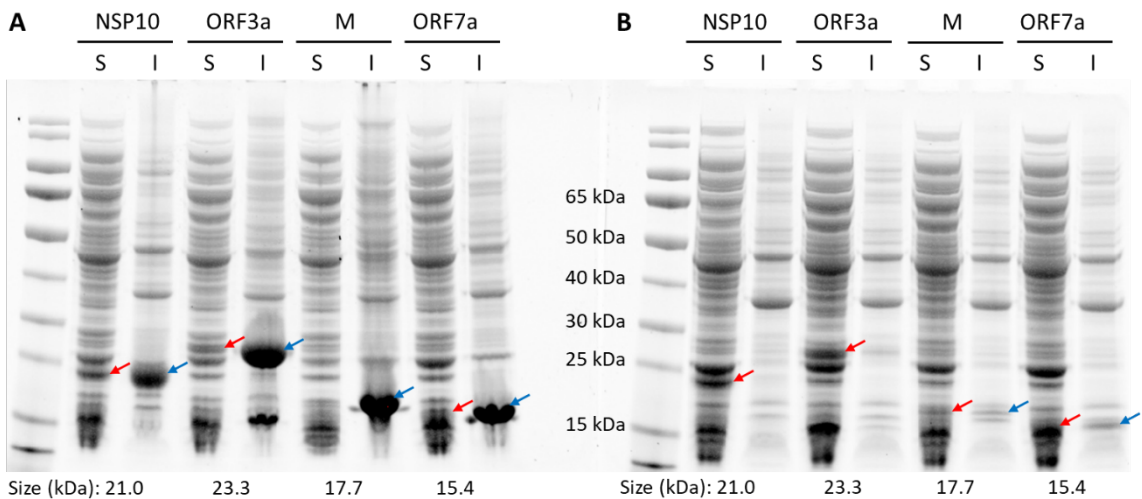


Figure S2-2. SDS-PAGE for expression of SARS-CoV-2 NSP10, ORF3a, M and ORF7a. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (A) and with 0.4 mM of IPTG for 18 hours at 16 °C (B). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

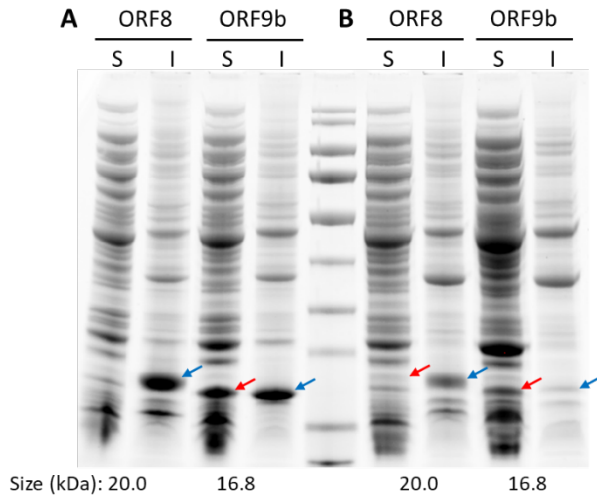


Figure S2-3. SDS-PAGE for expression of SARS-CoV-2 NSP10, ORF8 and ORF9b. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (**A**) and with 0.4 mM of IPTG for 18 hours at 16 °C (**B**). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

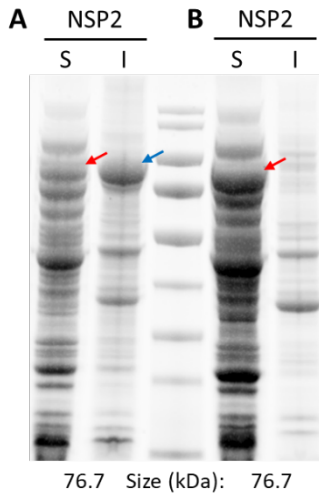


Figure S2-4. SDS-PAGE for expression of SARS-CoV-2 NSP2. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (**A**) and with 0.4 mM of IPTG for 18 hours at 16 °C (**B**). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

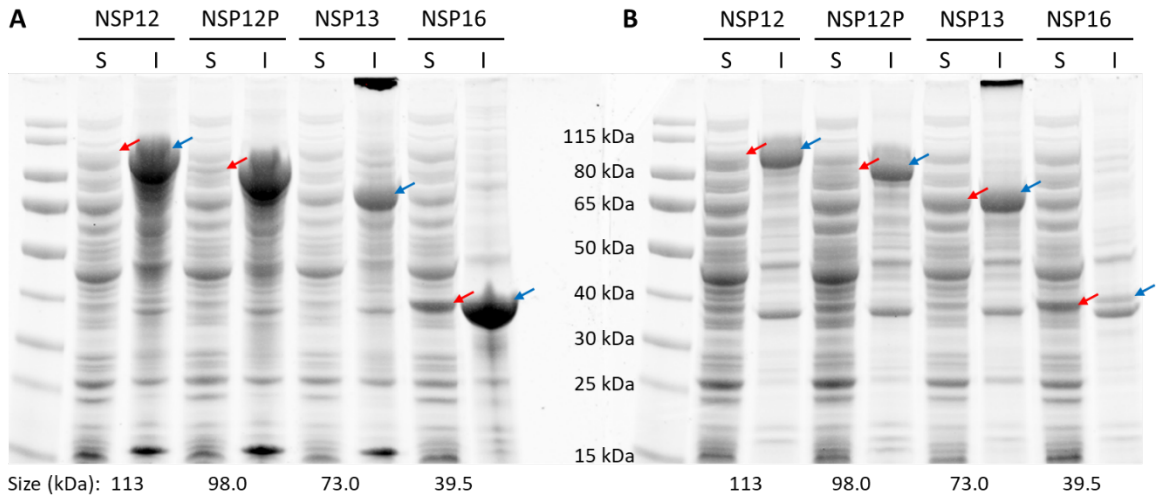


Figure S2-5. SDS-PAGE for expression of SARS-CoV-2 NSP12, NSP12P, NSP13 and NSP16. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (**A**) and with 0.4 mM of IPTG for 18 hours at 16 °C (**B**). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

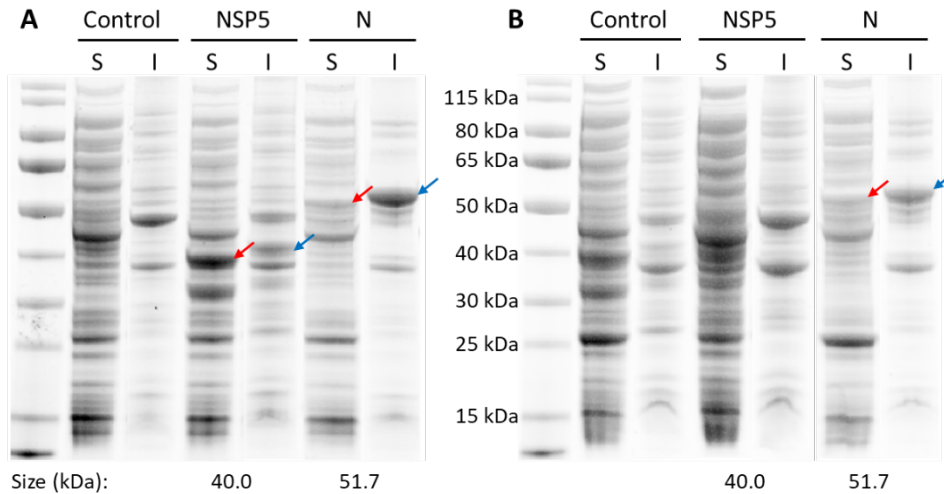


Figure S2-6. SDS-PAGE for expression of SARS-CoV-2 NSP5 and N. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C (**A**) and with 0.4 mM of IPTG for 18 hours at 16 °C (**B**). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

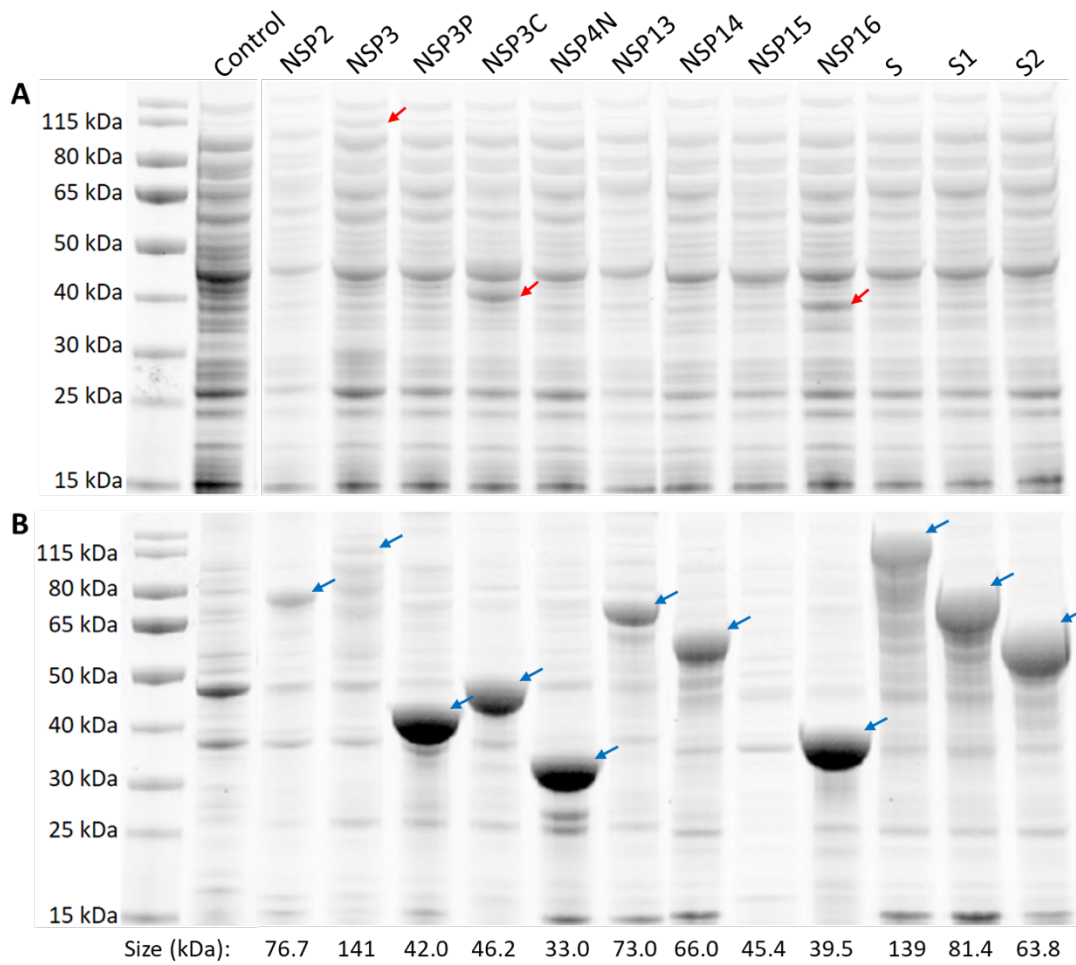


Figure S2-7. SDS-PAGE for expression of SARS-CoV-2 NSP5 and N. The protein expression was induced with 0.2 mM of IPTG for 2 hours at 37 °C. The red and blue arrows indicate the target proteins in soluble (**A**) and insoluble (**B**) fractions, respectively.

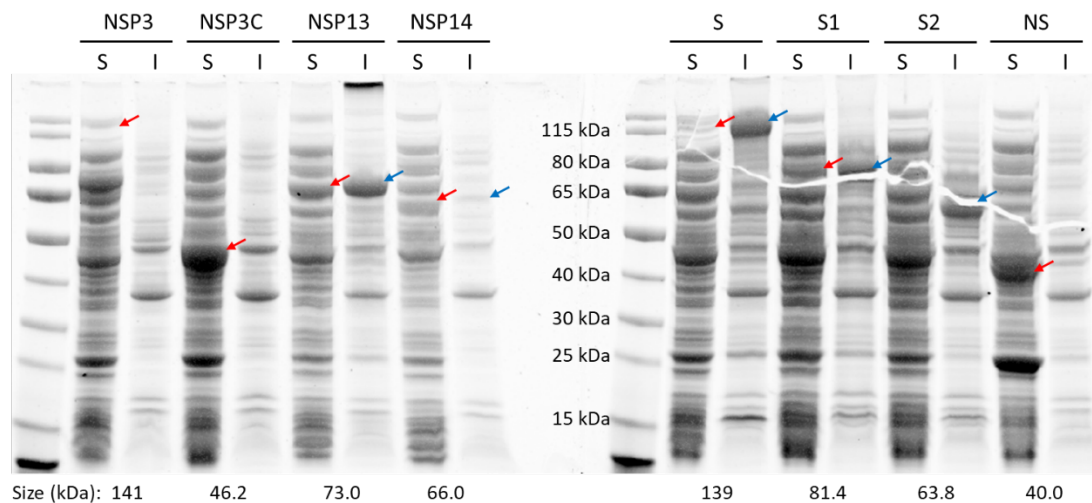


Figure S2-8. SDS-PAGE for expression of SARS-CoV-2 NSP3, NSP3C, NSP13, NSP14, S, S1, S2 and NS. The protein expression was induced with 0.4 mM of IPTG for 18 hours at 16 °C. The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.

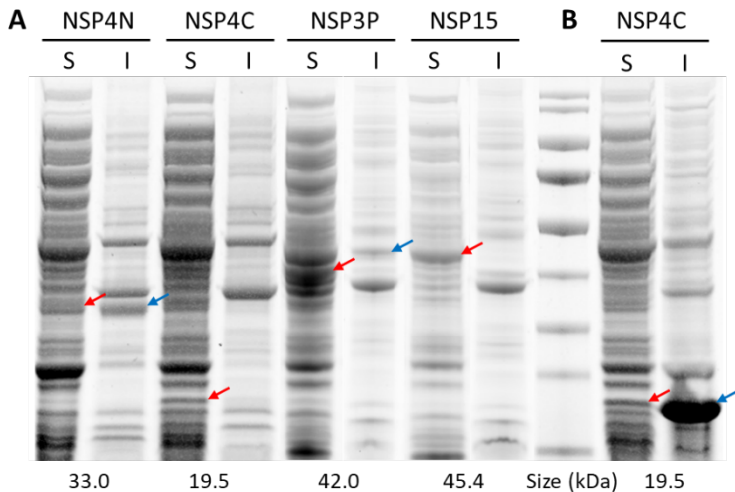
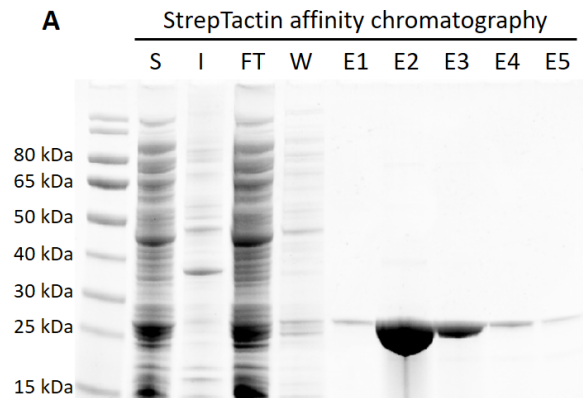


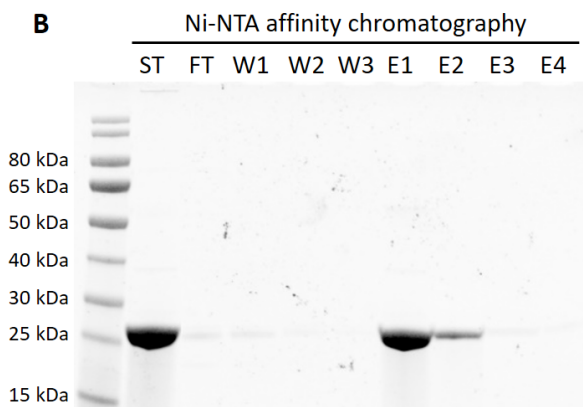
Figure S2-9. SDS-PAGE for expression of SARS-CoV-2 NSP4N, NSP4C, NSP3P and NSP15. The protein expression was induced with 0.4 mM of IPTG for 18 hours at 16 °C (A) and with 0.2 mM of IPTG for 2 hours at 37 °C (B). The red and blue arrows indicate the target proteins in soluble (S) and insoluble (I) fractions, respectively.



S, soluble; I, insoluble; FT, flow-through; W, washing; E, elution

Washing buffer I: 50 mM Tris-Cl, pH 8.0, 150 mM NaCl

Elution buffer: 50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin



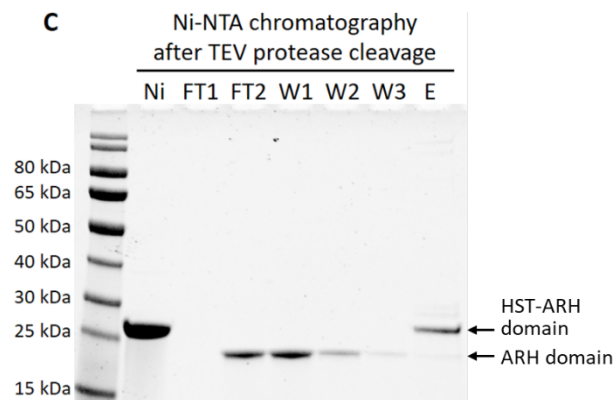
ST: Eluted fraction from **A**; W1, W2 & W3: washing buffer II (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 20 mM Imidazole)

E1: Washing buffer II + 200 mM Imidazole

E2: Washing buffer II + 400 mM Imidazole

E3: Washing buffer II + 800 mM Imidazole

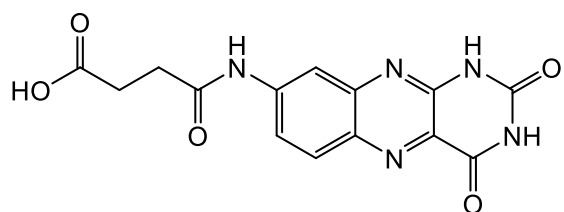
E4: Washing buffer II + 1 M Imidazole



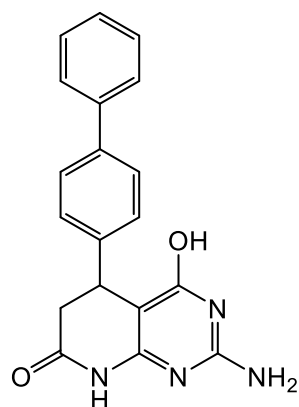
Ni: Eluted fraction from **B**; W1, W2 & W3: washing buffer II (50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 20 mM Imidazole)

E: Washing buffer II + 800 mM Imidazole

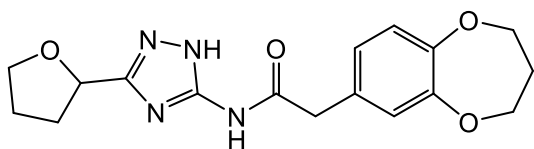
Figure S3. SDS-PAGE gels for the purification of SARS-CoV-2 NSP3 (ARH domain).



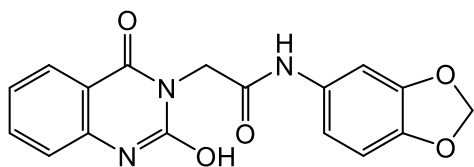
95



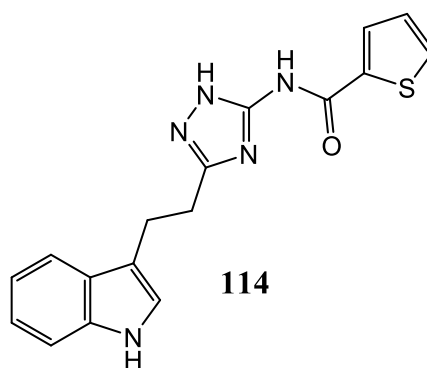
96



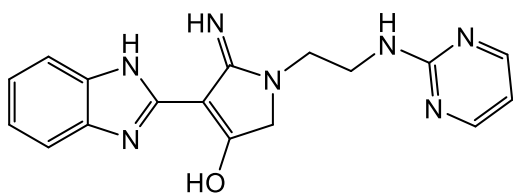
98



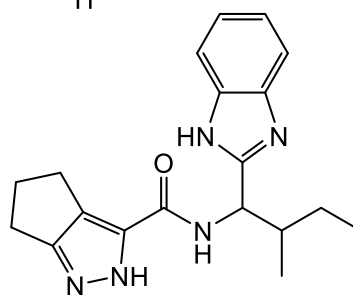
112



114



116



119

Figure S4. The chemical structures of seven compounds that bind to the ARH domain in the first set. Compounds **96**, **98** and **119** are racemates.

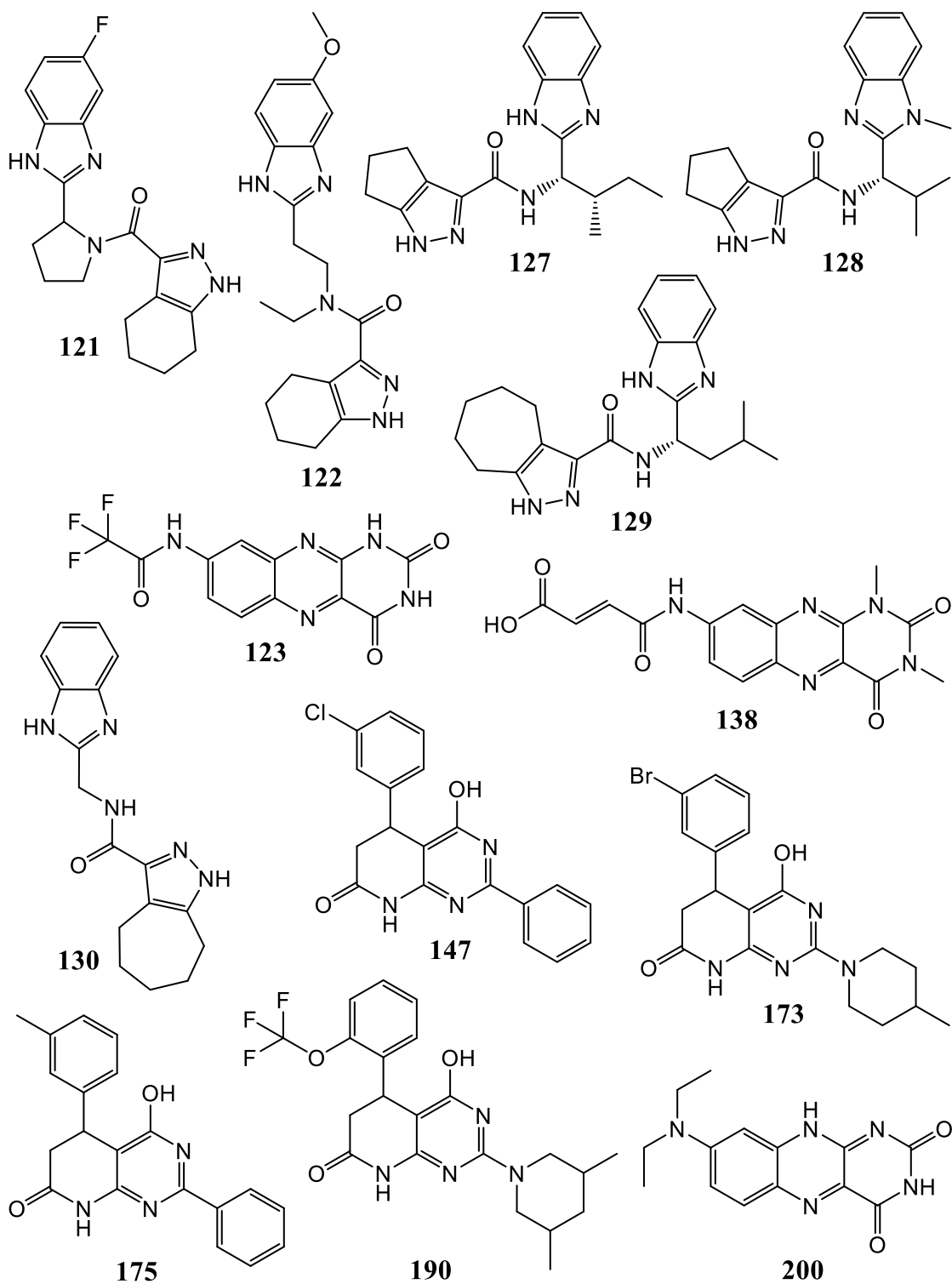


Figure S5. The chemical structures of 13 compounds that bind to the ARH domain in the second set. Compounds **121**, **147**, **174**, **175** and **190** are racemates.