Figure S1

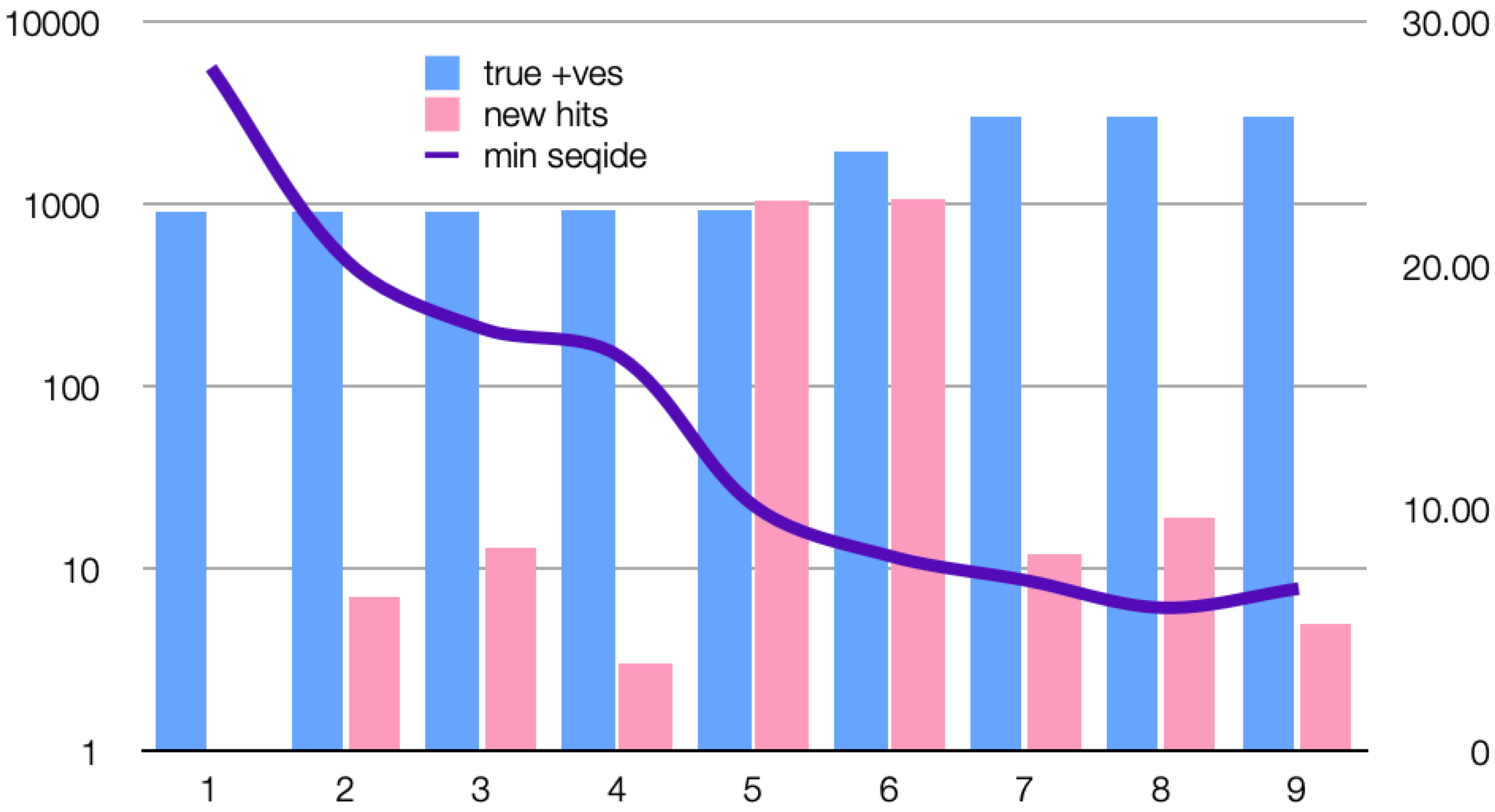
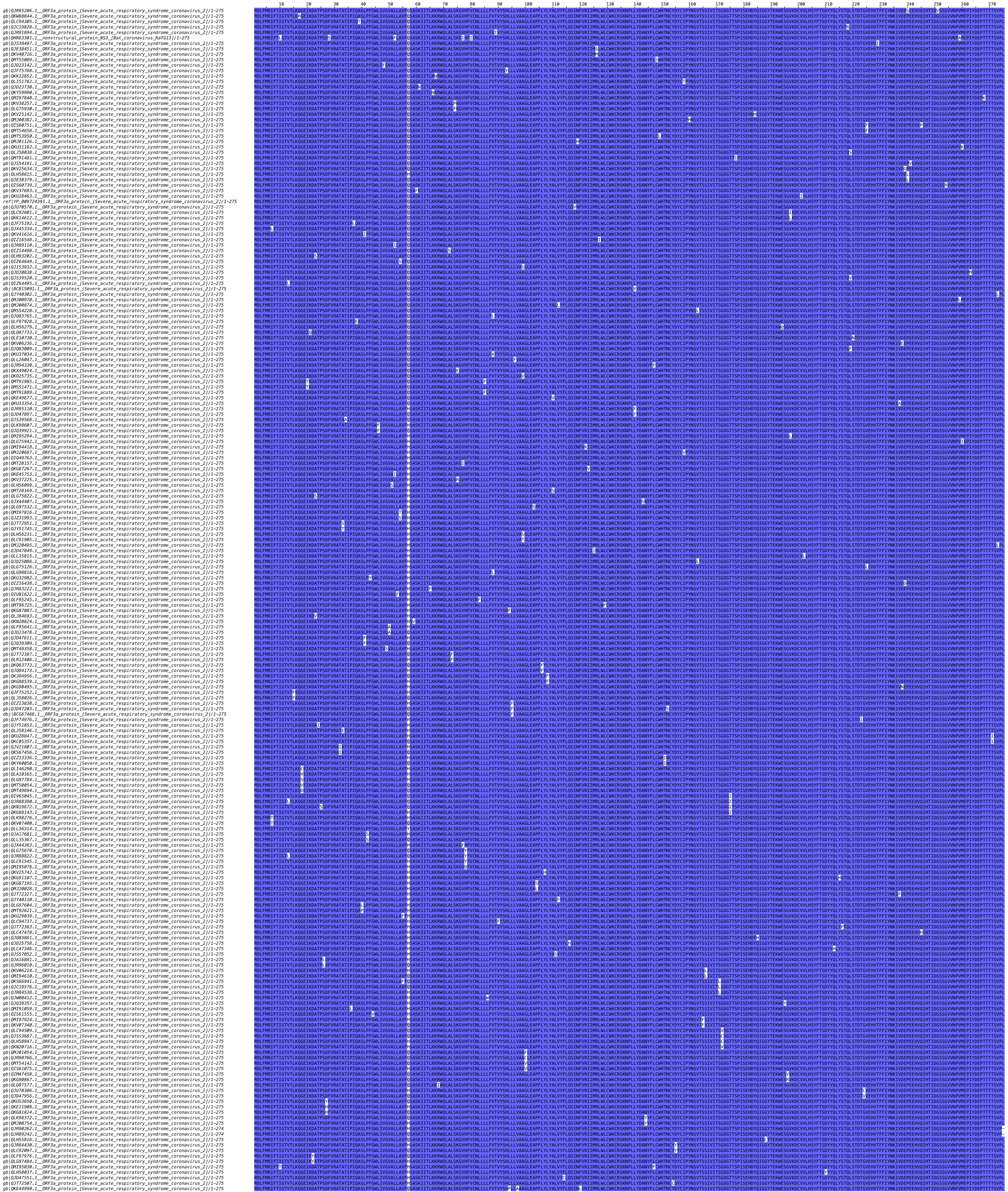

10000
1000
100
10
1
30.00
20.00
10.00
0
true +ves
new hits
min seqide
1
2
3
4
5
6
7
8
9

Figure S2



10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270

Figure S3

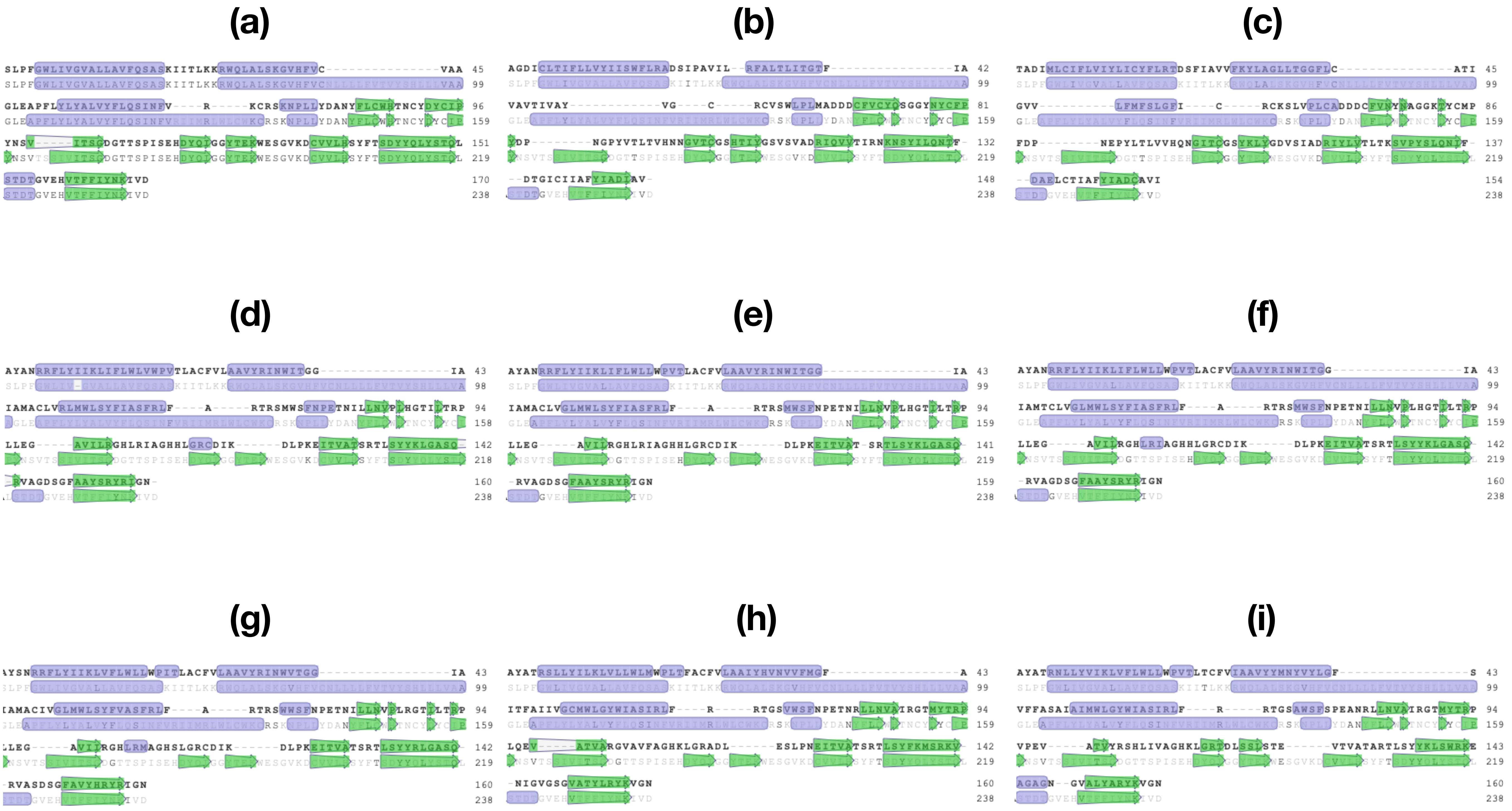


Figure S4

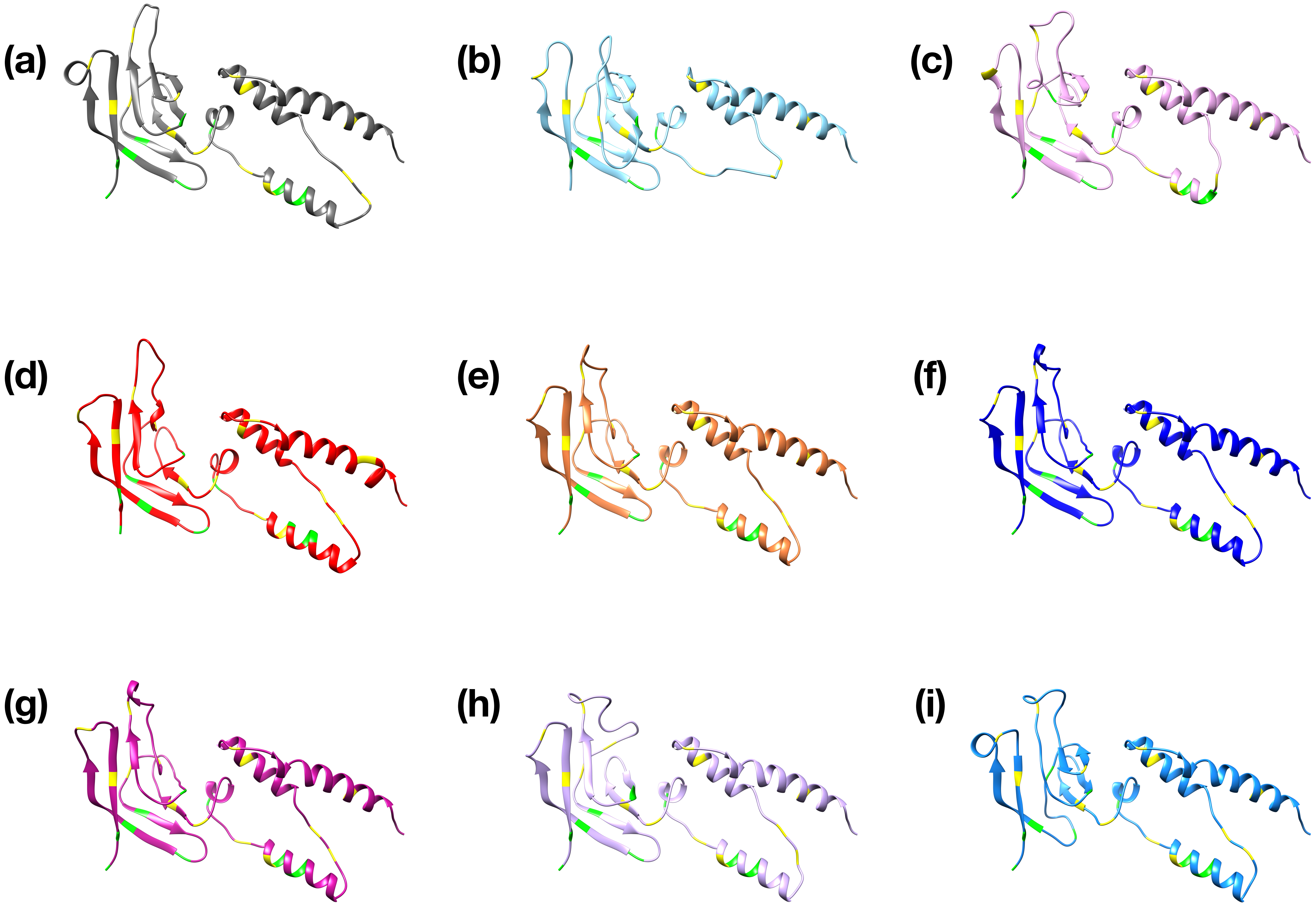


Figure S5

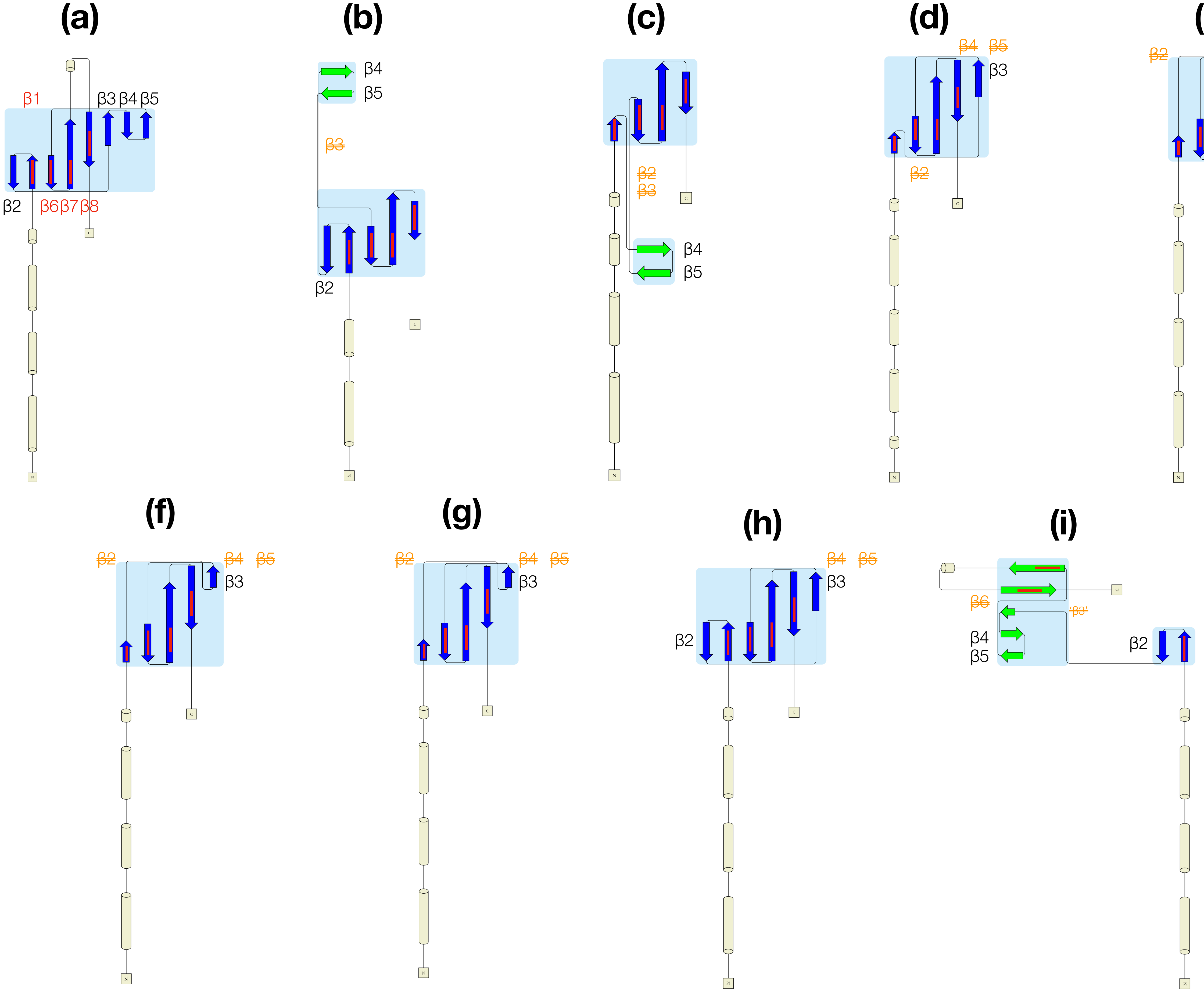

(a)
β1
β3 β4 β5
β2
β6 β7 β8
(b)
β4
β5
~~β3~~
β2
(c)
~~β2~~
~~β3~~
β4
β5
(d)
~~β4~~ ~~β5~~
β3
~~β2~~
(e)
~~β2~~
~~β4~~ ~~β5~~
β3
(f)
~~β2~~
~~β4~~ ~~β5~~
β3
(g)
~~β2~~
~~β4~~ ~~β5~~
β3
(h)
~~β4~~ ~~β5~~
β3
β2
(i)
~~β6~~
~~'β3'~~
β4
β5
β2

Figure S6

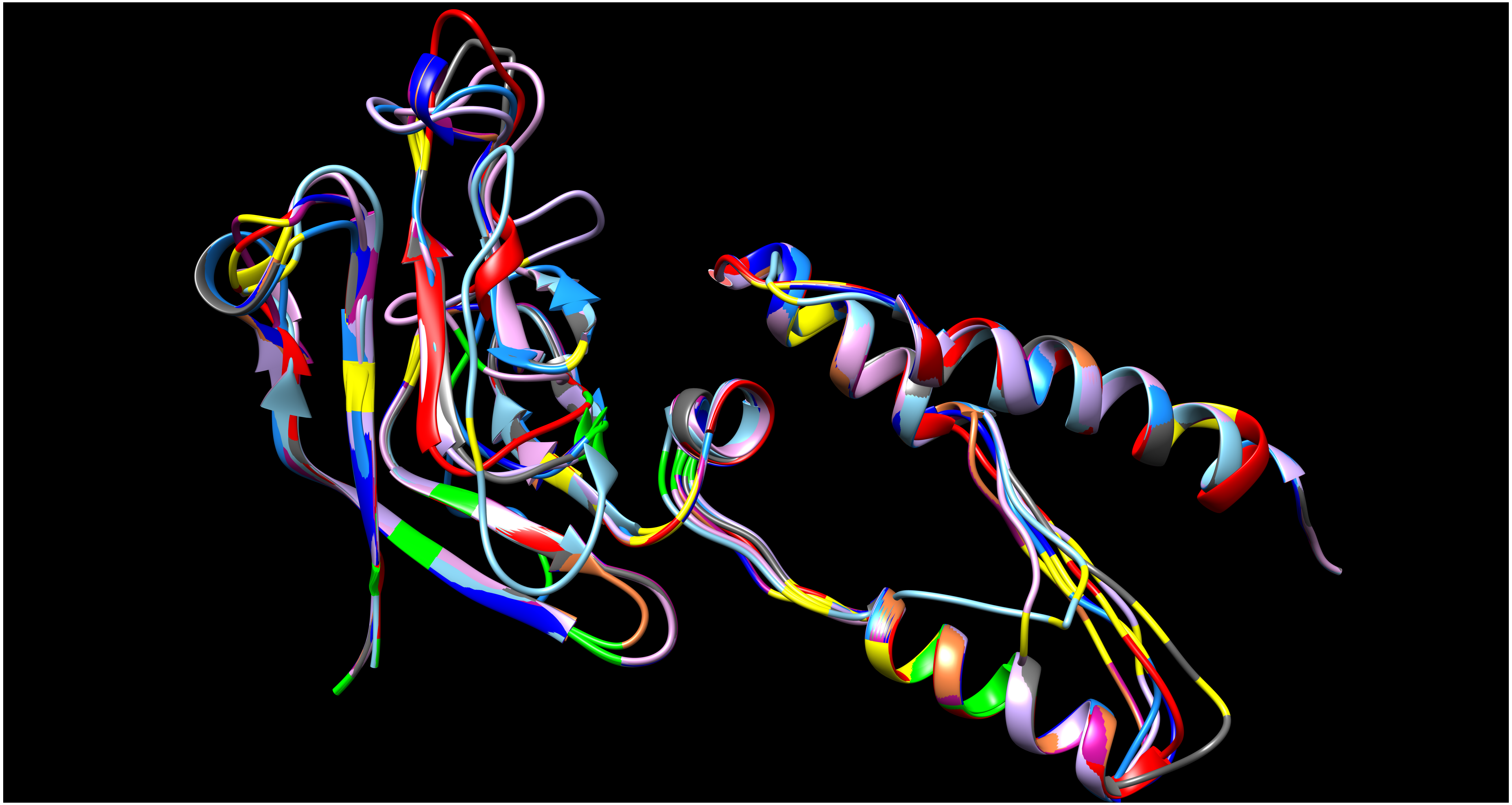


Figure S7

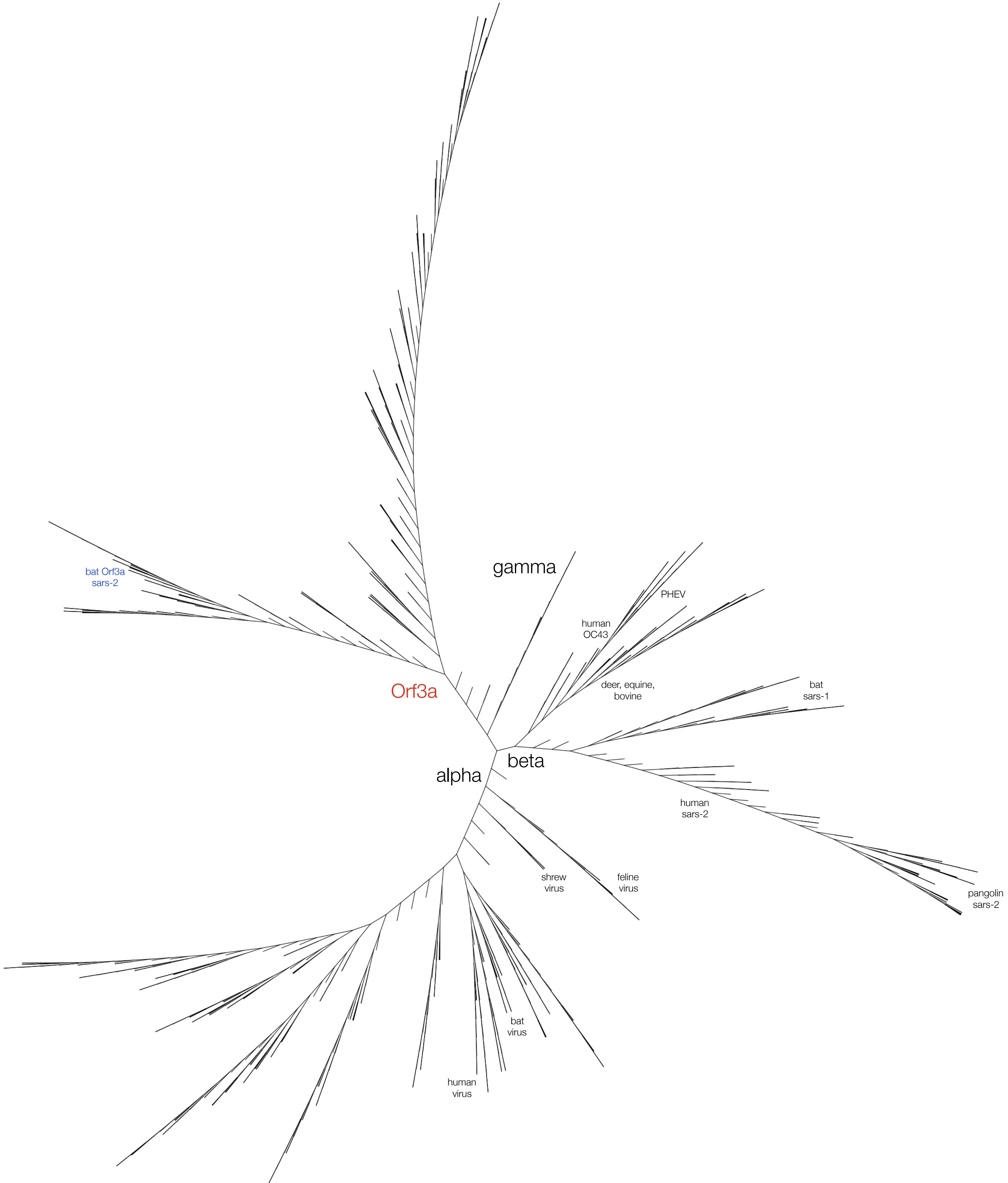
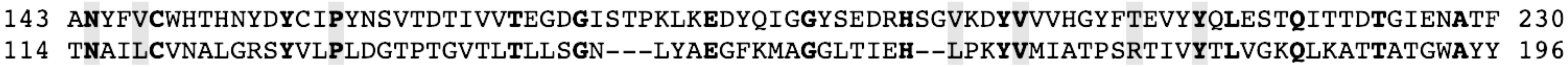

bat Orf3a
sars-2
Orf3a
gamma
human
OC43
PHEV
deer, equine,
bovine
bat
sars-1
beta
alpha
human
sars-2
pangolin
sars-2
shrew
virus
feline
virus
bat
virus
human
virus

Figure S8

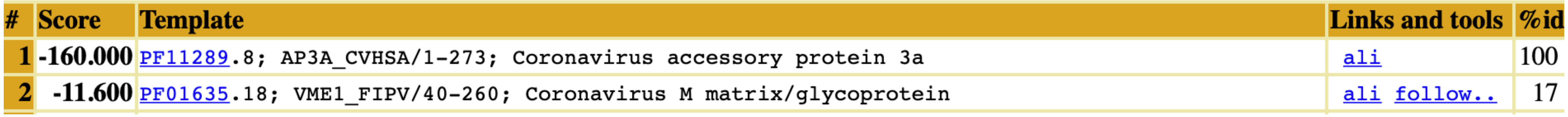
Query: PF11289.8; AP3A_CVHSA/1-273; Coronavirus accessory protein 3a
Template: PF01635.18; VME1_FIPV/40-260; Coronavirus M matrix/glycoprotein



```
143 ANYFVCWHTHNYDYCIPYNSVTDTIVVTEGDGISTPKLKEDYQIGGYSEDRHSGVKDYVVVHGYFTEVYYQLESTQITTDTGIENATF 230
114 TNAILCVNALGRSYVLPLDGTPTGVTLTLLSGN---LYAEGFKMAGGLTIEH--LPKYVMIATPSRTIVYTLVGKQLKATTATGWAYY 196
```

◆◆

http://ffas.godziklab.org/ffas-cgi/cgi/get_mu.pl?ses=&rv=PfamA32U*PfamA32U*re*PF11289&lv=



#	Score	Template	Links and tools	%id
1	-160.000	PF11289.8; AP3A_CVHSA/1-273; Coronavirus accessory protein 3a	ali	100
2	-11.600	PF01635.18; VME1_FIPV/40-260; Coronavirus M matrix/glycoprotein	ali follow..	17

Table S1

Run #	true +ves	new hits	all hits	excluded	final hits	PDB ids	min seqide	precision	comment	pfam	save archive
1	898	0	898	0	898	6xdc	28.10	100.00	orf3a homologs	PF11289	KGB7F59A014
2	898	7	905	0	905	na	20.14	100.00	hypothetical bat orf3a’s	na	KGBF0TZE014
3	905	13	918	0	918	na	17.36	100.00	orf3a homologs only	na	KGC2RBC0014
4	918	3	921	0	921	na	16.24	100.00	first ‘feline’ M proteins	PF01635	KGCB6UR6014
5	921	1034	1955	0	1955	na	10.05	100.00	almost 1000 M proteins	na	KGCTVD93014
6	1955	1069	3024	0	3024	na	8.00	100.00	another 1000 M proteins	na	KGD3TX9B016
7	3024	12	3036	25	3011	na	7.00	100.00	partial sequences (excluded)	na	KGER1U38014
8	3011	19	3030	0	3025	na	5.88	100.00	more partial sequences	na	KGFXSXGP016
9	3025	5*	3025	0	3025	na	6.67	100.00	no hits, convergence	na	KGGMKT62014

Table S2

sequence id	description	model	length	seq identity	QMEAN	C-beta	all-atom	solvation	torsion
a-6xdcA	SARS-CoV-2 Orf3a	3fnkwV	170	100.00	-4.53	-3.99	-1.14	-3.06	-2.64
b-ADY17912.1	Zaria bat CoV ZBCoV Orf3	Krxq5Y	148	20.95	-5.61	-3.51	-3.02	-6.10	-2.79
c-YP_009072441.1	Bat Hp Beta CoV Zhejiang2013	4cdPqa	154	20.78	-7.41	-4.43	-2.18	-6.94	-4.36
d-QIG55988.1	SARS-CoV-2 M	JBtQPZ	160	5.66	-6.13	-3.95	-1.98	-5.01	-3.67
e-QMJ19959.1	SARS-CoV-2 M	3dyh3c	159	12.58	-7.00	-3.92	-2.50	-4.98	-4.53
f-QIQ54051.1	Pangolin CoV M	KrDHuh	160	13.13	-6.11	-3.42	-2.29	-5.05	-3.73
g-YP_003858587.1	Bat CoV BM48-31/BGR2008 M	2kGQSz	160	13.13	-6.61	-4.30	-2.41	-4.72	-4.17
h-YP_009072443.1	Bat Hp Beta CoV Zhejiang2013 M	TJjzr2	160	13.75	-5.65	-4.44	-1.61	-5.09	-3.05
i-ADY69165.1	Zaria bat CoV ZBCoV M	F2S4AZ	160	16.25	-7.31	-3.95	-2.55	-5.42	-4.72
average of a-i			159	24.03	-6.26	-3.99	-2.19	-5.15	-3.74
native 6xdc	monomer of 6xdc	na	238	100.00	-3.22	-3.46	0.72	-1.30	-2.05
native 6xdc	dimeric form (as in PDB)	na	238	100.00	-2.97	-3.63	0.67	-0.90	-2.23