wwPDB EM Validation Summary Report ⓘ

Jul 6, 2023 – 07:18 pm BST

PDB ID : 8C6J
EMDB ID : EMD-16452
Title : Human spliceosomal PM5 C* complex
Authors : Dybkov, O.; Kastner, B.; Luehrmann, R.
Deposited on : 2023-01-12
Resolution : 2.80 Å(reported)
Based on initial models : 6HYU, 5MQF, 3MDF, 7OS2, 6QDV, 5HLY, 2OK3, 4PJ3, 6BK8, .

The following versions of software and data (see references ⓘ) were used in the production of this report:

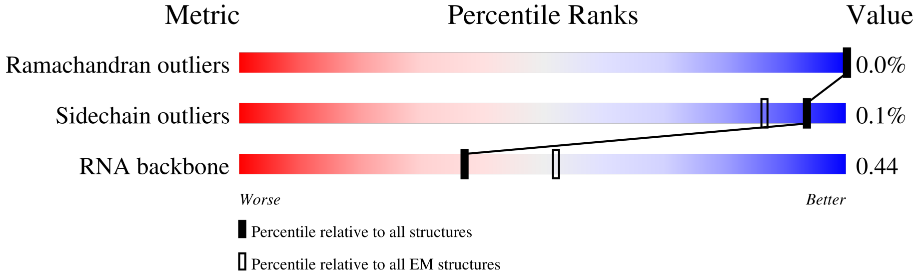
EMDB validation analysis : 0.0.1.dev50
Mogul : 1.8.4, CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.34

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

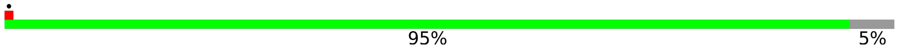
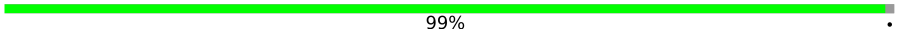
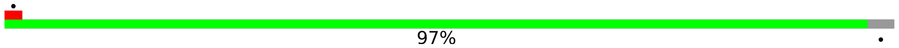
The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	2	188	
2	4	476	
3	5	116	
4	6	106	
5	7	411	
6	8	174	
7	9	146	
8	A	2335	

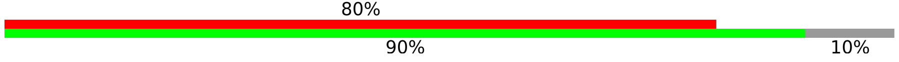
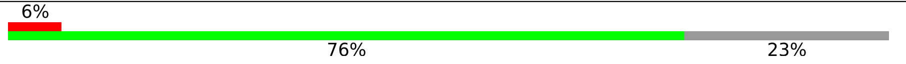
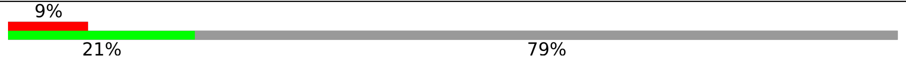
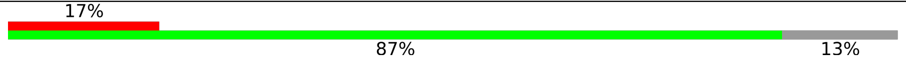
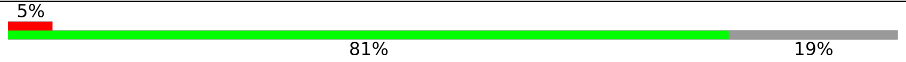
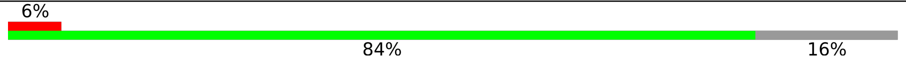
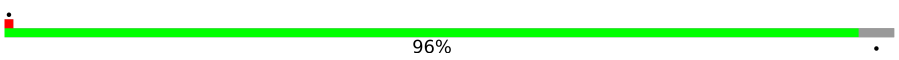
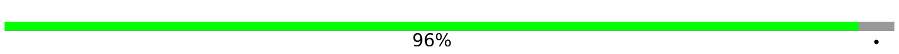
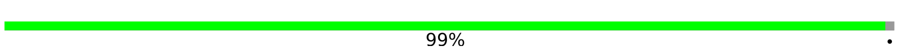
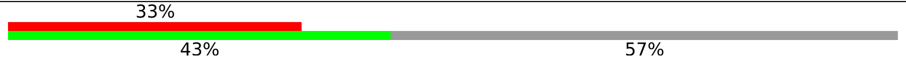
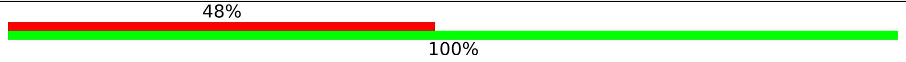
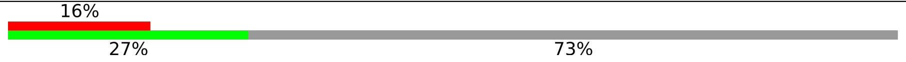
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	B	2136	53% 81% 19%
10	C	972	92% 8%
11	C3	161	61% 99% •
12	CD	622	9% 39% 61%
13	CE	301	6% 26% 74%
14	CF	339	45% 54% 45%
15	CI	285	14% 25% 75%
16	CN	301	16% 82% • 17%
17	CT	289	12% 35% 65%
18	D	184	• 67% 33%
19	E	242	5% 5% 90%
20	F	758	16% 84%
21	G	112	• 53% 47%
22	H	500	92% 8%
23	I	235	15% 20% 64%
24	J	514	6% 71% 29%
25	K	536	55% 45%
26	L	144	100%
27	M	420	69% 31%
28	N	357	86% 14%
29	O	802	19% 68% 32%
30	P	229	51% 49%
31	R	2752	• 97%
32	S	687	17% 92% 8%
33	T	855	11% 81% 19%

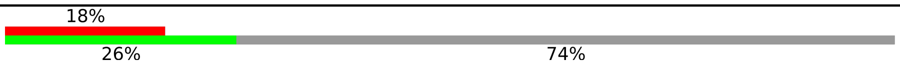
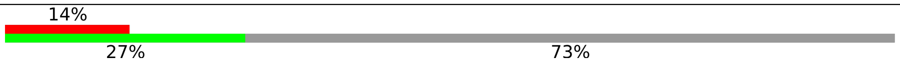
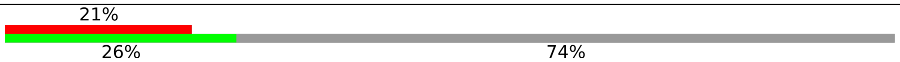
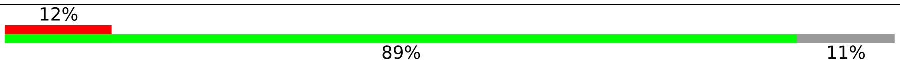
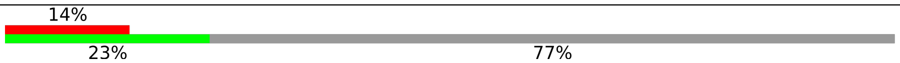
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	U	1485	
35	V	1220	
36	W	255	
37	Y	225	
38	Z	415	
39	a	118	
39	m	118	
40	b	240	
40	k	240	
41	c	119	
41	l	119	
42	d	126	
42	n	126	
43	e	92	
43	p	92	
44	f	86	
44	q	86	
45	g	76	
45	r	76	
46	h	586	
47	i	166	
48	j	222	
49	o	579	
50	s	225	
51	t	504	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
51	u	504	
51	v	504	
51	w	504	
52	y	243	
53	z	451	

2 Entry composition

There are 59 unique types of molecules in this entry. The entry contains 111233 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	2	142	3010	1346	517	1005	142	0	0

- Molecule 2 is a protein called Splicing factor ESS-2 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	4	77	636	395	107	133	1	0	0

- Molecule 3 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	5	113	2371	1064	394	801	112	0	0

- Molecule 4 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	6	97	2075	928	381	669	97	0	0

- Molecule 5 is a protein called Eukaryotic initiation factor 4A-III.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	7	390	3130	1976	546	589	19	0	0

- Molecule 6 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	8	91	730	463	122	142	3	0	0

- Molecule 7 is a protein called Protein mago nashi homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
7	9	144	595	306	144	145	0	0

- Molecule 8 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	A	2261	17716	11283	3167	3195	71	0	0

- Molecule 9 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
9	B	1726	7141	3689	1726	1726	0	0

- Molecule 10 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	C	899	7114	4552	1184	1344	34	0	0

- Molecule 11 is a protein called Peptidyl-prolyl cis-trans isomerase-like 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	C3	160	669	345	162	162	2	0

- Molecule 12 is a protein called Probable ATP-dependent RNA helicase DDX41.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
12	CD	245	1025	535	245	245	0	0

- Molecule 13 is a protein called Peptidyl-prolyl cis-trans isomerase E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	CE	79	322	164	79	79	0	0

- Molecule 14 is a protein called Protein FAM50A.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
14	CF	185	746	376	185	185	0	0

- Molecule 15 is a protein called Pre-mRNA-splicing factor ISY1 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
15	CI	70	283	143	70	70	0	0

- Molecule 16 is a protein called Nitric oxide synthase-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	CN	250	1643	991	313	327	12	0	0

- Molecule 17 is a protein called Splicing factor C9orf78.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
17	CT	102	414	210	102	102	0	0

- Molecule 18 is a protein called PRKR-interacting protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	D	123	1013	635	193	180	5	0	0

- Molecule 19 is a RNA chain called PM5 5' Exon.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
19	E	24	511	229	97	162	23	0	0

- Molecule 20 is a protein called Splicing factor Cactin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	F	122	1088	714	197	175	2	0	0

- Molecule 21 is a protein called Protein FAM32A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	G	59	494	308	94	90	2	0	0

- Molecule 22 is a protein called Pre-mRNA-splicing factor CWC22 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	H	459	3321	2079	603	621	18	0	0

- Molecule 23 is a RNA chain called PM5 Intron.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
23	I	84	1727	775	254	614	84	0	0

- Molecule 24 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	J	367	2720	1697	504	511	8	0	0

- Molecule 25 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
25	K	295	2360	1479	431	435	2	13	0	0

- Molecule 26 is a protein called Protein BUD31 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	L	144	1188	748	218	210	12	0	0

- Molecule 27 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	M	289	2327	1459	416	432	20	0	0

- Molecule 28 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	N	306	Total 2388	C 1498	N 419	O 457	S 14	0	0

- Molecule 29 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	O	548	Total 3321	C 1939	N 693	O 682	S 7	0	0

- Molecule 30 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	P	117	Total 973	C 595	N 190	O 186	S 2	0	0

- Molecule 31 is a protein called Serine/arginine repetitive matrix protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	R	76	Total 399	C 226	N 86	O 86	S 1	0	0

- Molecule 32 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	S	635	Total 3835	C 2264	N 784	O 781	S 6	0	0

- Molecule 33 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	T	689	Total 4273	C 2649	N 798	O 813	S 13	0	0

- Molecule 34 is a protein called Intron-binding protein aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	U	1330	Total 5548	C 2884	N 1332	O 1332	6	0

- Molecule 35 is a protein called ATP-dependent RNA helicase DHX8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
35	V	934	4167	2221	962	983	1	0	0

- Molecule 36 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
36	W	176	722	370	176	176	0	0

- Molecule 37 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
37	Y	92	377	193	92	92	0	0

- Molecule 38 is a protein called NF-kappa-B-activating protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
38	Z	87	454	250	99	103	2	0	0

- Molecule 39 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
39	a	103	520	314	103	103	0	0
			Total	C	N	O		
39	m	95	395	205	95	95	0	0

- Molecule 40 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
40	b	82	413	249	82	82	0	0
			Total	C	N	O		
40	k	82	340	176	82	82	0	0

- Molecule 41 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
41	c	80	Total	C	N	O	0	0
			402	242	80	80		
41	l	80	Total	C	N	O	0	0
			329	169	80	80		

- Molecule 42 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
42	d	84	Total	C	N	O	0	0
			420	252	84	84		
42	n	83	Total	C	N	O	0	0
			341	175	83	83		

- Molecule 43 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	e	79	Total	C	N	O	0	0
			393	235	79	79		
43	p	81	Total	C	N	O	0	0
			327	165	81	81		

- Molecule 44 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	f	72	Total	C	N	O	0	0
			359	215	72	72		
44	q	72	Total	C	N	O	0	0
			300	156	72	72		

- Molecule 45 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
45	g	73	Total	C	N	O	0	0
			365	218	73	74		
45	r	73	Total	C	N	O	0	0
			302	155	73	74		

- Molecule 46 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	h	276	Total	C	N	O	S	0	0
			2280	1433	408	431	8		

- Molecule 47 is a protein called Peptidyl-prolyl cis-trans isomerase-like 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
47	i	164	Total 1270	C 810	N 220	O 233	S 7	0	0

- Molecule 48 is a protein called STING ER exit protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
48	j	95	Total 398	C 208	N 95	O 95	0	0

- Molecule 49 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
49	o	513	Total 3918	C 2478	N 692	O 724	S 24	0	0

- Molecule 50 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms				AltConf	Trace
50	s	225	Total 922	C 471	N 225	O 226	0	0

- Molecule 51 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
51	t	137	Total 581	C 307	N 137	O 137	0	0
51	u	130	Total 550	C 290	N 130	O 130	0	0
51	v	138	Total 585	C 309	N 138	O 138	0	0
51	w	131	Total 554	C 292	N 131	O 131	0	0

- Molecule 52 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
52	y	216	Total 1423	C 843	N 284	O 294	S 2	0	0

- Molecule 53 is a protein called Splicing regulator SDE2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
53	z	103	606	346	133	124	3	0	0

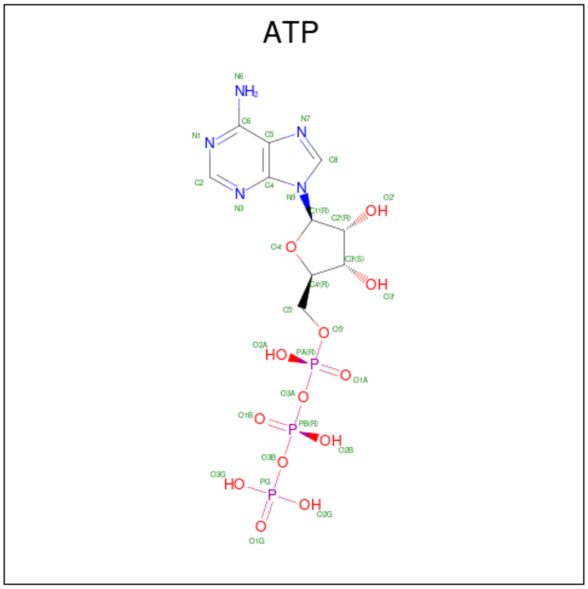
- Molecule 54 is POTASSIUM ION (three-letter code: K) (formula: K) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	K	
54	6	2	2	2	0

- Molecule 55 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

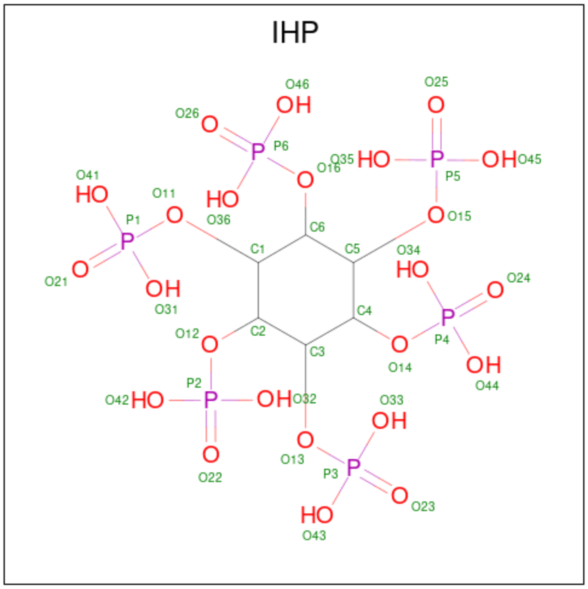
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
55	6	5	5	5	0
55	7	1	1	1	0
55	C	1	1	1	0

- Molecule 56 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



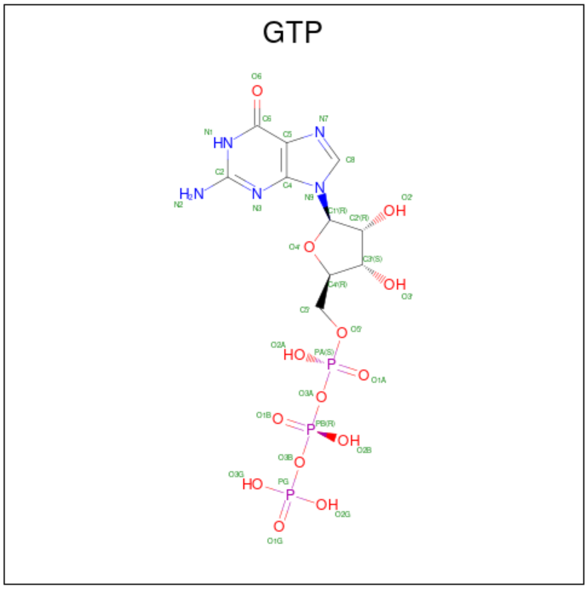
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
56	7	1	31	10	5	13	3	0

- Molecule 57 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
57	A	1	36	6	24	6	0

- Molecule 58 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
58	C	1	32	10	5	14	3	0

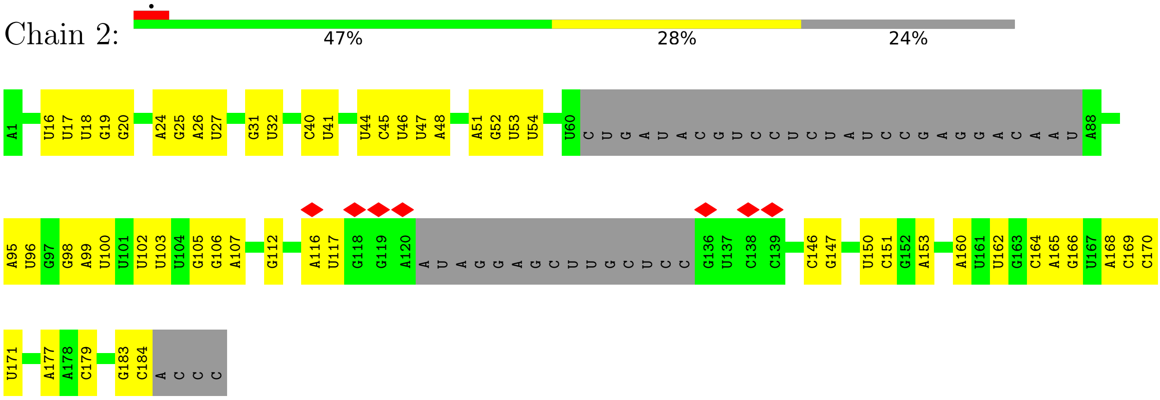
- Molecule 59 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
59	L	3	3	3	0
59	M	3	3	3	0

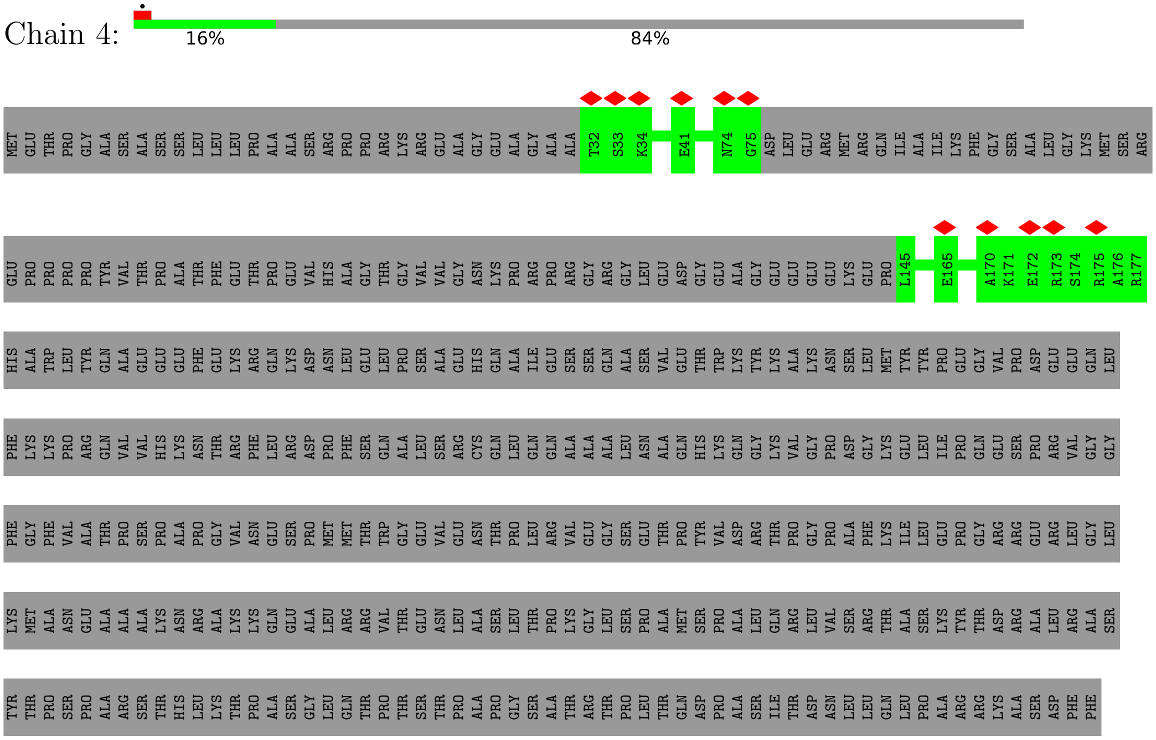
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

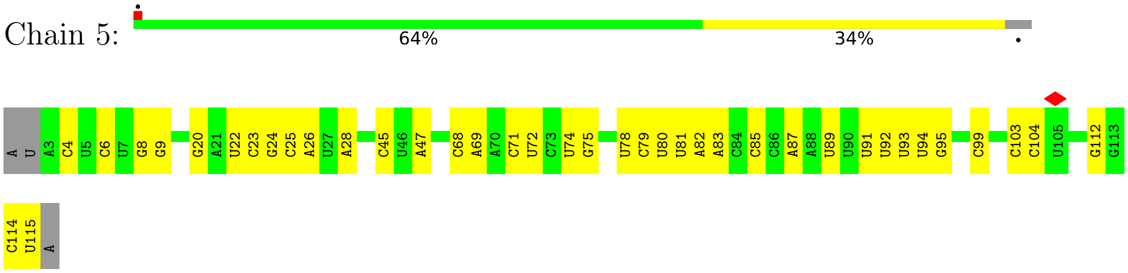
- Molecule 1: U2 snRNA



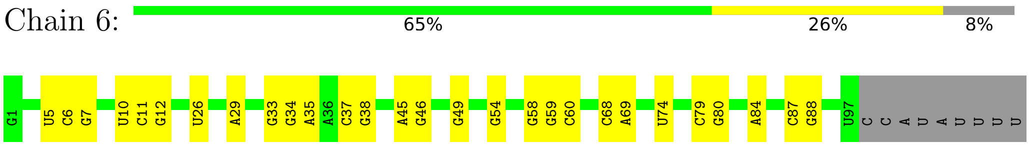
- Molecule 2: Splicing factor ESS-2 homolog



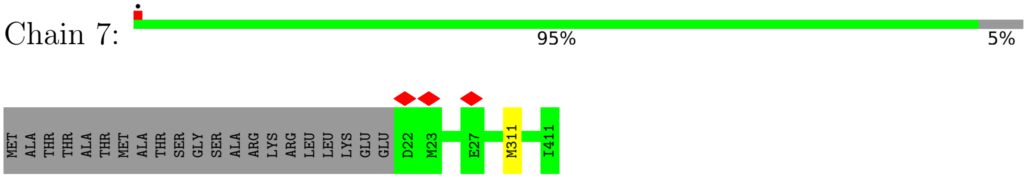
- Molecule 3: U5 snRNA



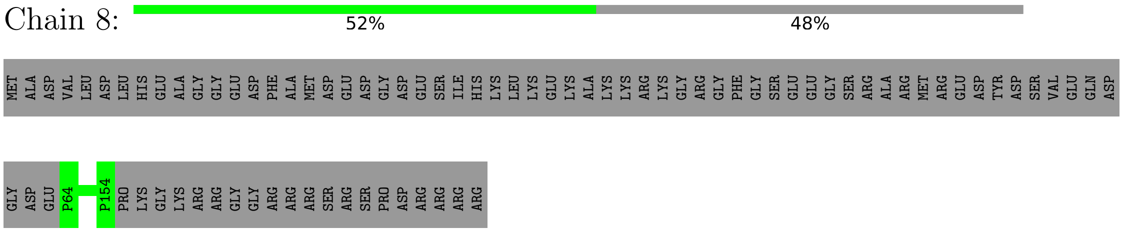
- Molecule 4: U6 snRNA



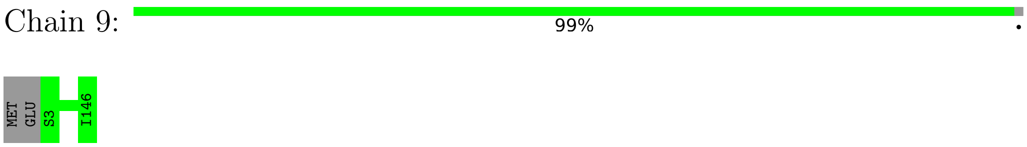
- Molecule 5: Eukaryotic initiation factor 4A-III



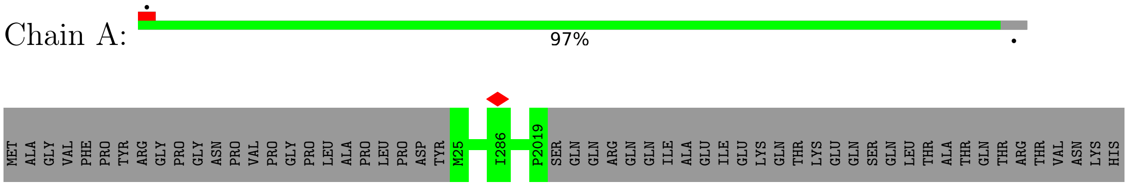
- Molecule 6: RNA-binding protein 8A

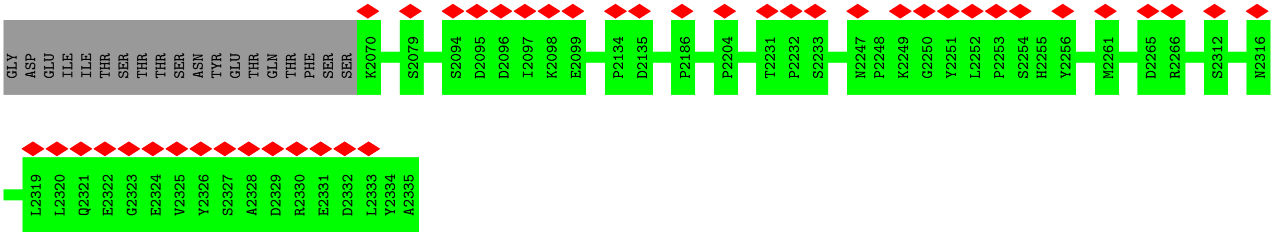


- Molecule 7: Protein mago nashi homolog

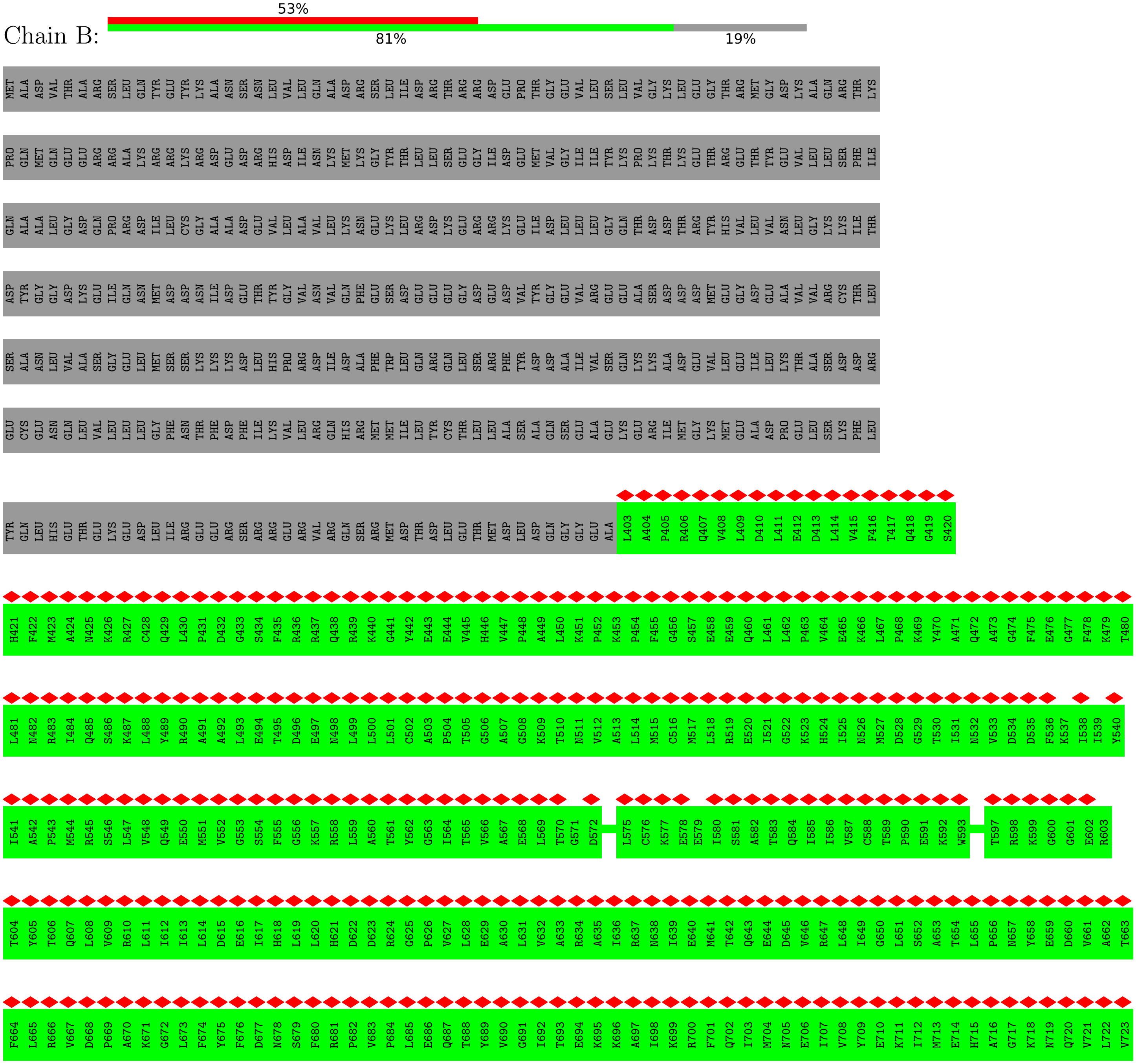


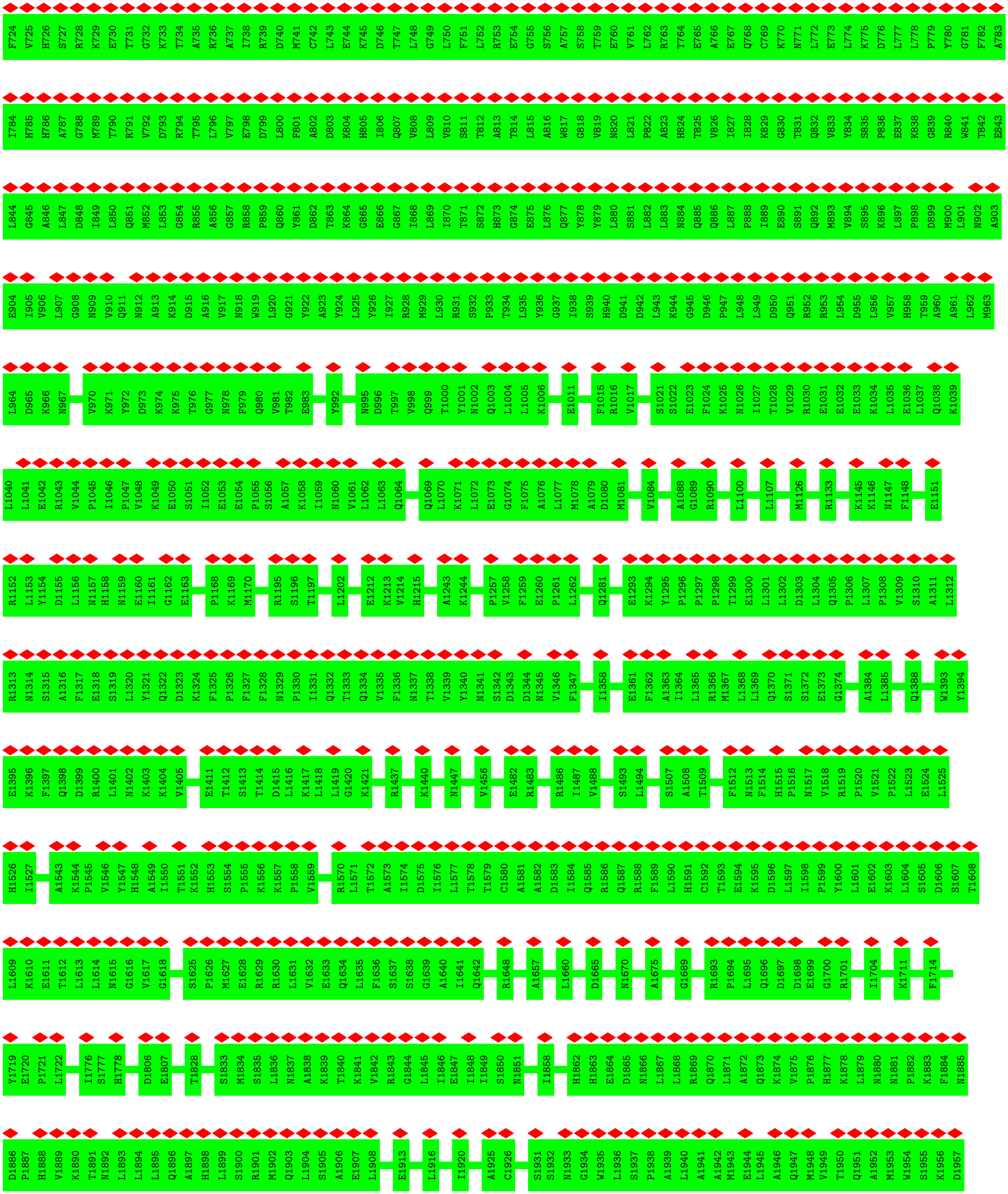
- Molecule 8: Pre-mRNA-processing-splicing factor 8

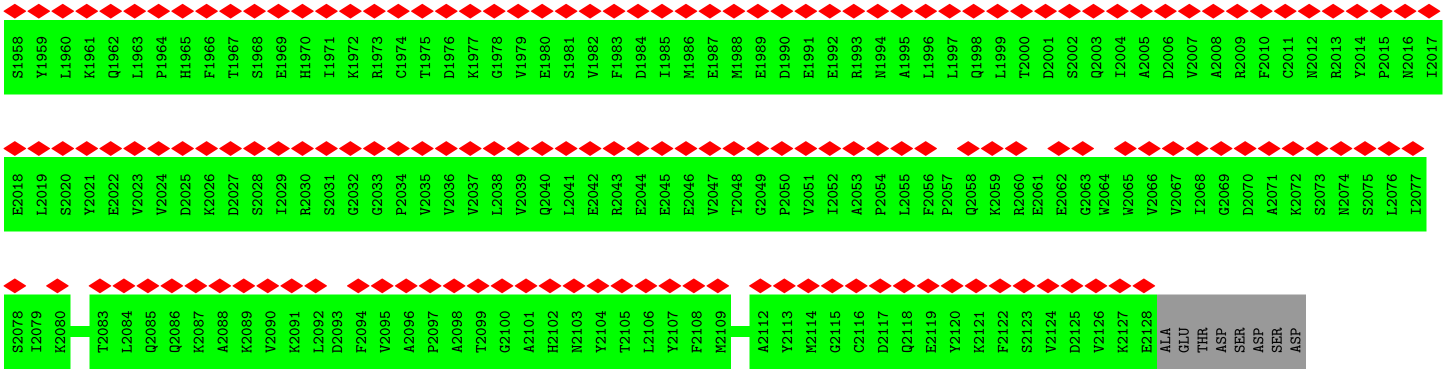




- Molecule 9: U5 small nuclear ribonucleoprotein 200 kDa helicase

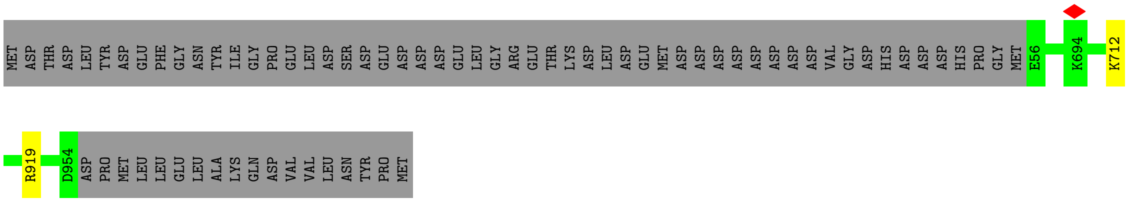



F724
V725
H726
S727
R728
K729
E730
T731
G732
K733
T734
A735
R736
A737
I738
R739
D740
M741
C742
L743
E744
K745
D746
T747
L748
G749
L750
F751
L752
R753
E754
G755
S756
A757
S758
T759
E760
V761
L762
R763
T764
E765
A766
E767
Q768
C769
K770
N771
L772
E773
L774
K775
D776
L777
L778
P779
Y780
G781
F782
A783
I784
H785
H786
A787
G788
M789
T790
R791
V792
D793
R794
T795
L796
V797
E798
D799
L800
F801
A802
D803
K804
H805
I806
Q807
V808
L809
V810
S811
T812
A813
T814
L815
A816
W817
G818
V819
N820
L821
P822
A823
H824
T825
V826
I827
I828
K829
G830
T831
Q832
V833
Y834
S835
P836
E837
K838
G839
R840
W841
T842
E843
L844
G845
A846
L847
D848
I849
L850
Q851
M852
L853
G854
R855
A856
G857
R858
P859
Q860
Y861
D862
T863
K864
G865
E866
G867
I868
L869
I870
T871
S872
H873
G874
E875
L876
Q877
Y878
I879
L880
S881
L882
L883
N884
Q885
Q886
L887
P888
I889
E890
S891
Q892
M893
V894
S895
K896
L897
P898
D899
M900
L901
N902
A903
E904
I905
V906
L907
G908
N909
V910
Q911
N912
A913
K914
D915
A916
V917
N918
W919
L920
G921
Y922
A923
Y924
L925
Y926
I927
R928
M929
L930
R931
S932
P933
T934
L935
Y936
G937
I938
S939
H940
D941
D942
L943
K944
G945
D946
P947
L948
L949
D950
Q951
R952
R953
L954
D955
L956
V957
H958
T959
A960
L961
L962
M963
L964
D965
K966
N967
V970
K971
Y972
D973
K974
K975
T976
G977
N978
F979
Q980
V981
T982
E983
Y992
N995
D996
T997
V998
Q999
T1000
Y1001
N1002
Q1003
L1004
L1005
K1006
E1011
F1015
R1016
V1017
S1021
S1022
E1023
F1024
K1025
N1026
I1027
T1028
V1029
R1030
E1031
E1032
E1033
K1034
L1035
E1036
L1037
Q1038
K1039
L1040
L1041
E1042
R1043
V1044
P1045
I1046
P1047
V1048
K1049
E1050
S1051
I1052
E1053
E1054
P1055
S1056
A1057
K1058
I1059
N1060
V1061
L1062
L1063
Q1064
Q1069
L1070
K1071
L1072
E1073
G1074
F1075
A1076
L1077
M1078
A1079
D1080
M1081
V1084
A1088
G1089
R1090
L1100
L1107
M1126
R1133
K1145
K1146
N1147
F1148
E1151
R1152
L1153
Y1154
D1155
L1156
N1157
H1158
N1159
E1160
I1161
G1162
E1163
P1168
K1169
M1170
R1195
S1196
T1197
L1202
E1212
K1213
V1214
H1215
A1243
K1244
P1257
V1258
F1259
E1260
P1261
L1262
Q1281
E1293
K1294
Y1295
P1296
P1297
P1298
T1299
E1300
L1301
L1302
D1303
L1304
Q1305
P1306
L1307
P1308
V1309
S1310
A1311
L1312
R1313
N1314
S1315
A1316
F1317
E1318
S1319
L1320
Y1321
Q1322
D1323
K1324
F1325
P1326
F1327
F1328
N1329
P1330
I1331
Q1332
T1333
Q1334
V1335
F1336
N1337
T1338
V1339
Y1340
N1341
S1342
D1343
D1344
N1345
V1346
F1347
I1358
E1361
F1362
A1363
I1364
L1365
R1366
M1367
L1368
L1369
Q1370
S1371
S1372
E1373
G1374
A1384
L1385
Q1388
W1393
Y1394
E1395
K1396
F1397
Q1398
D1399
R1400
L1401
N1402
K1403
K1404
V1405
E1411
T1412
S1413
T1414
D1415
L1416
K1417
L1418
L1419
G1420
K1421
R1437
K1440
N1447
V1456
E1482
R1483
R1486
I1487
V1488
S1493
L1494
S1507
A1508
T1509
F1512
N1513
F1514
H1515
P1516
N1517
V1518
R1519
P1520
V1521
P1522
L1523
E1524
L1525
H1526
I1527
A1543
K1544
P1545
V1546
Y1547
H1548
A1549
I1550
T1551
K1552
H1553
S1554
P1555
K1556
K1557
P1558
V1559
R1570
L1571
T1572
A1573
I1574
D1575
I1576
L1577
T1578
T1579
C1580
A1581
A1582
D1583
I1584
Q1585
R1586
Q1587
R1588
F1589
L1590
H1591
C1592
T1593
E1594
K1595
D1596
L1597
I1598
P1599
Y1600
L1601
E1602
K1603
L1604
S1605
D1606
S1607
T1608
L1609
K1610
E1611
T1612
L1613
L1614
N1615
G1616
V1617
G1618
S1625
P1626
M1627
E1628
R1629
R1630
L1631
V1632
E1633
Q1634
L1635
F1636
S1637
S1638
G1639
A1640
I1641
Q1642
R1648
A1657
L1660
D1665
N1670
A1675
G1689
R1693
P1694
L1695
Q1696
D1697
D1698
E1699
G1700
R1701
I1704
K1711
F1714
Y1719
E1720
P1721
L1722
I1776
S1777
H1778
D1806
E1807
T1828
S1833
M1834
S1835
L1836
N1837
A1838
K1839
T1840
K1841
V1842
R1843
G1844
L1845
I1846
E1847
I1848
I1849
S1850
N1851
I1858
H1862
H1863
E1864
D1865
N1866
L1867
L1868
R1869
Q1870
L1871
A1872
Q1873
K1874
V1875
P1876
H1877
K1878
I1879
N1880
N1881
P1882
K1883
F1884
N1885
D1886
P1887
H1888
V1889
K1890
T1891
N1892
L1893
L1894
L1895
Q1896
A1897
H1898
L1899
S1900
R1901
M1902
Q1903
L1904
S1905
A1906
E1907
L1908
E1913
L1916
I1920
A1925
C1926
S1931
S1932
N1933
G1934
W1935
L1936
S1937
P1938
A1939
L1940
A1941
A1942
M1943
E1944
L1945
A1946
Q1947
M1948
V1949
T1950
Q1951
A1952
M1953
W1954
S1955
K1956
D1957



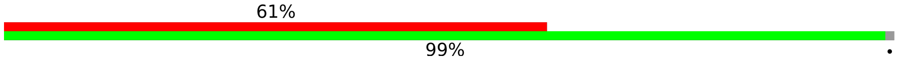
- Molecule 10: 116 kDa U5 small nuclear ribonucleoprotein component

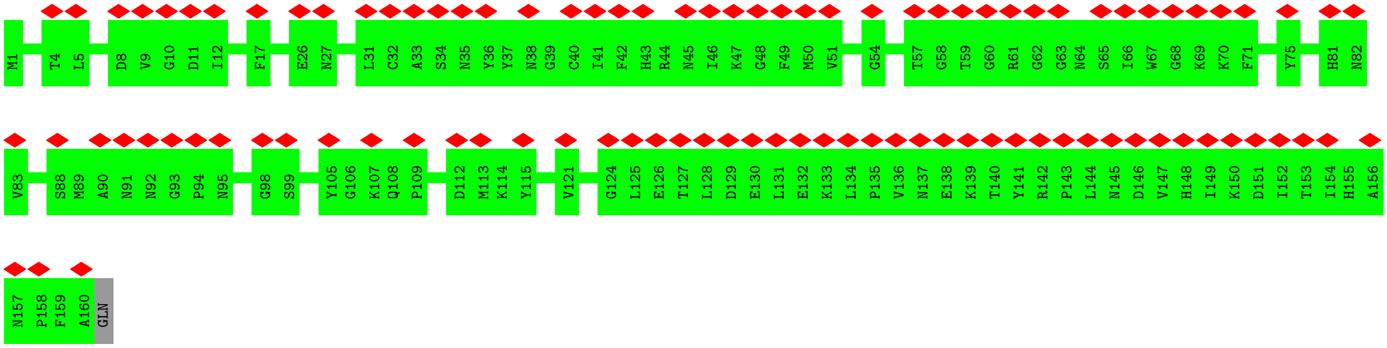
Chain C:



- Molecule 11: Peptidyl-prolyl cis-trans isomerase-like 3

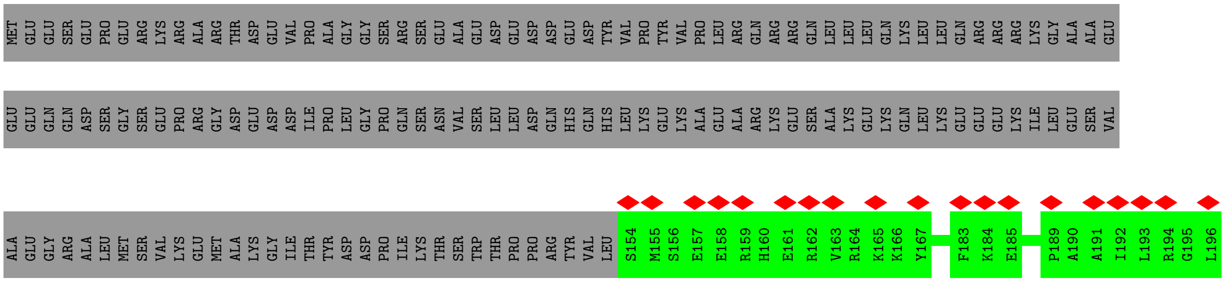
Chain C3:

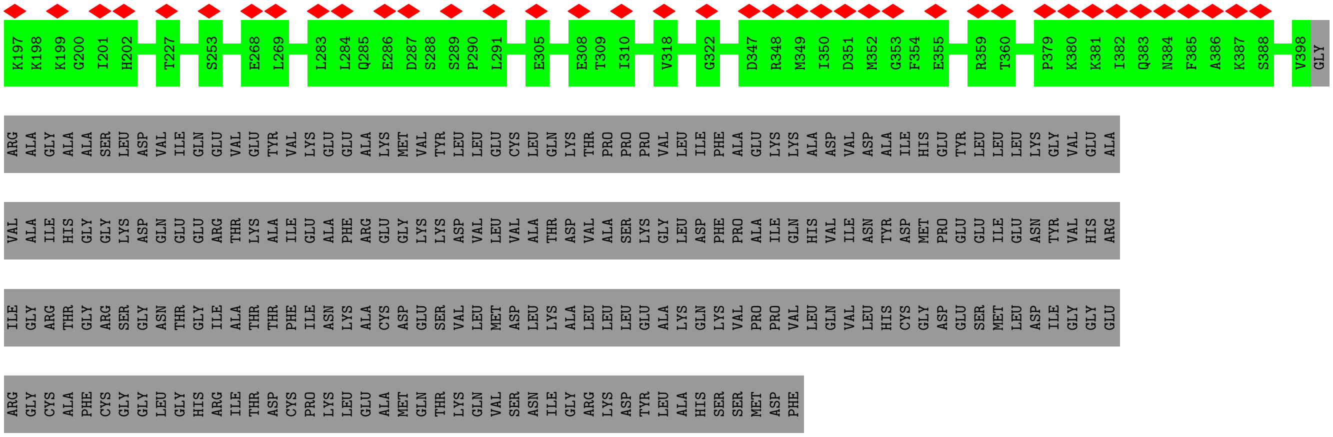




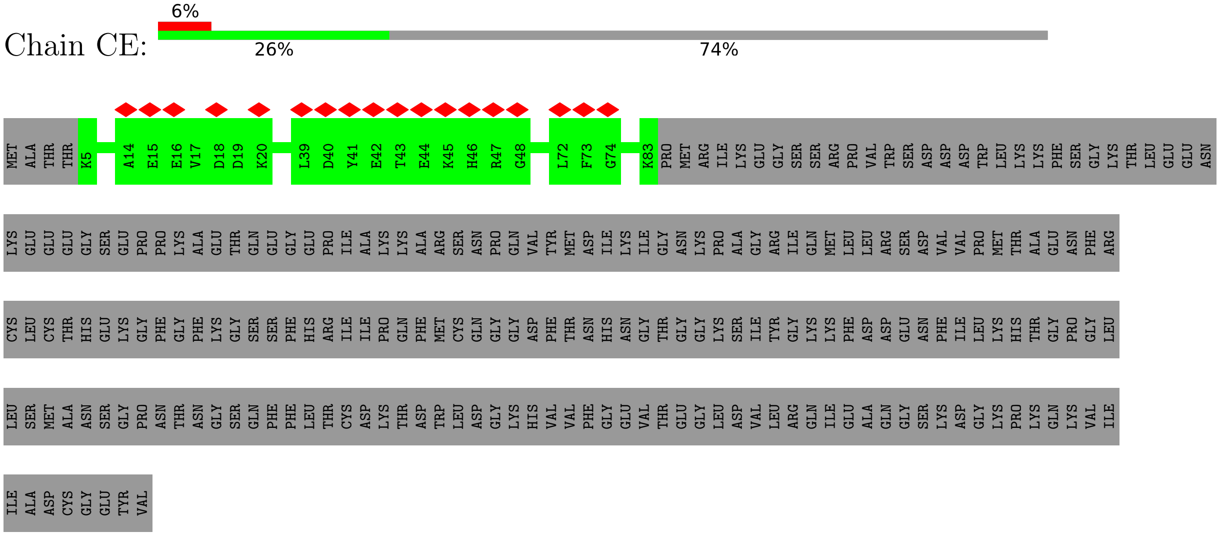
- Molecule 12: Probable ATP-dependent RNA helicase DDX41

Chain CD:

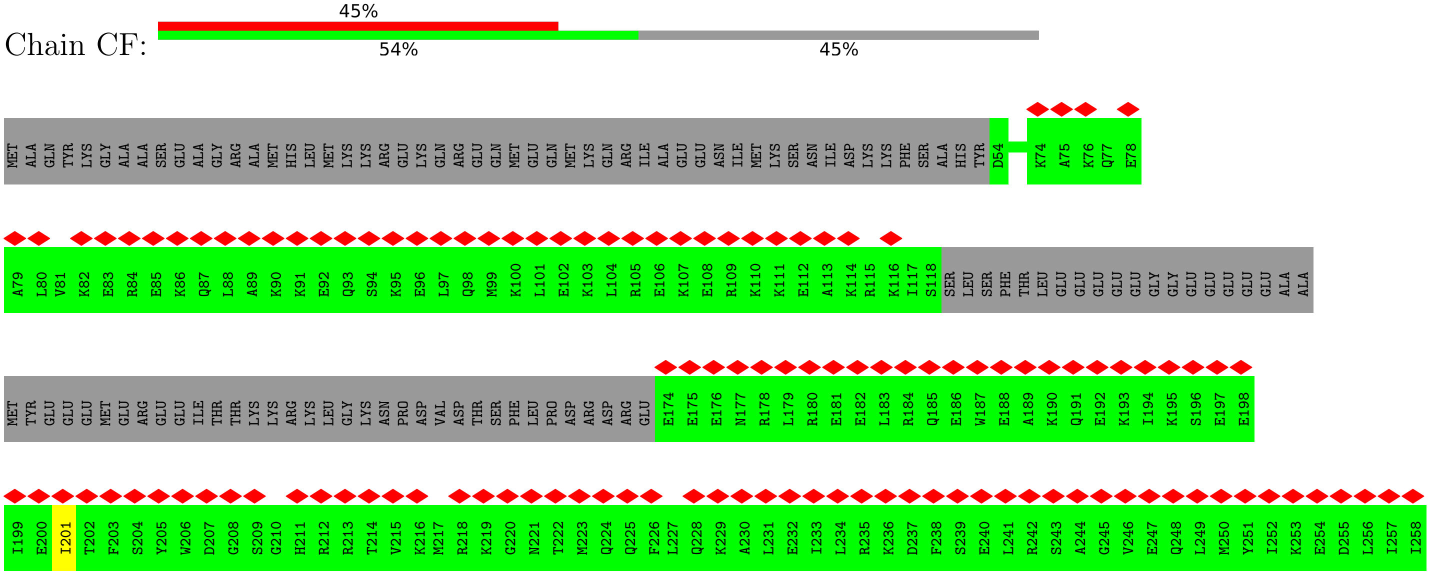


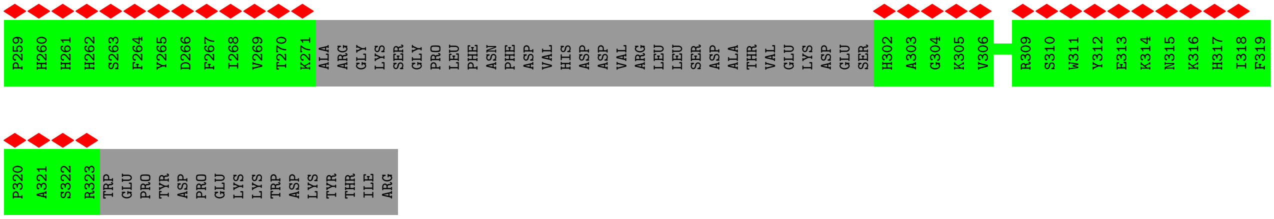


- Molecule 13: Peptidyl-prolyl cis-trans isomerase E

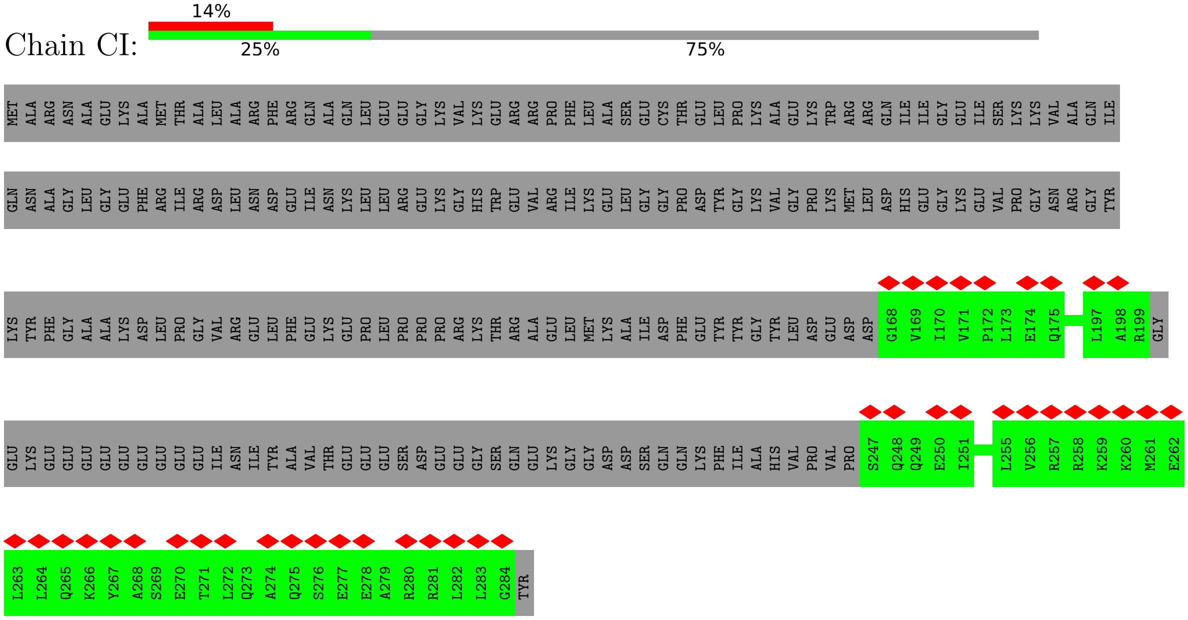


- Molecule 14: Protein FAM50A

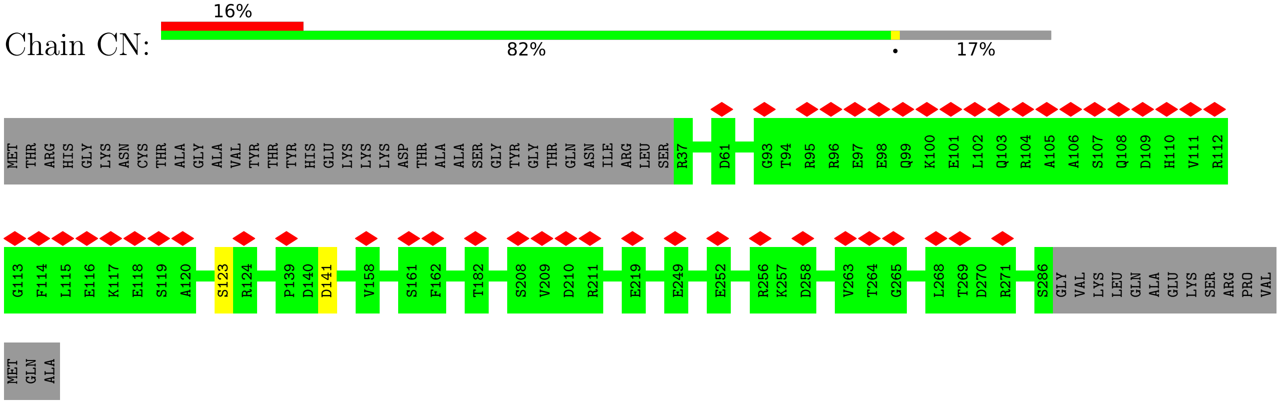




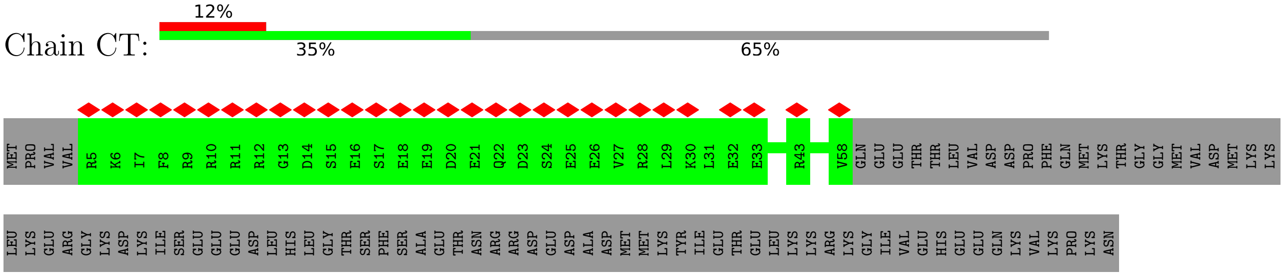
- Molecule 15: Pre-mRNA-splicing factor ISY1 homolog

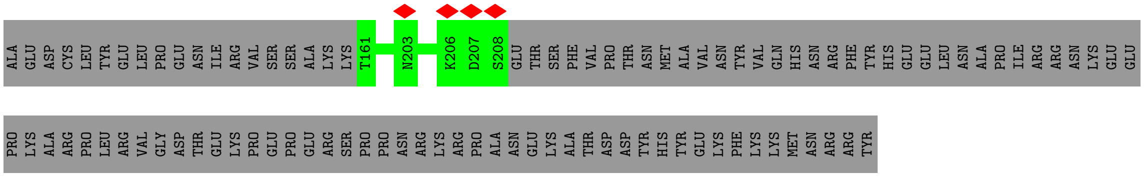


- Molecule 16: Nitric oxide synthase-interacting protein



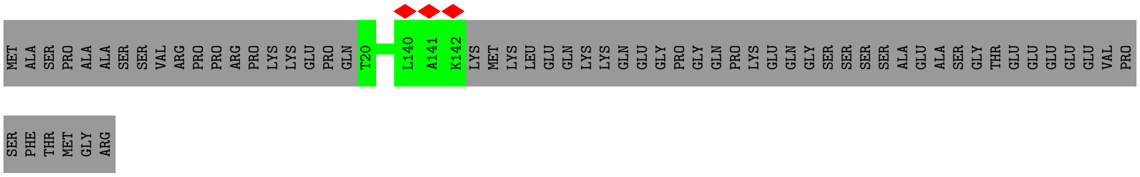
- Molecule 17: Splicing factor C9orf78





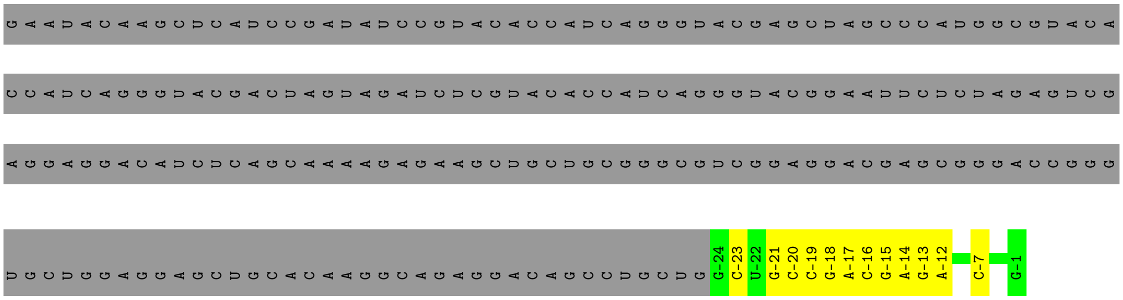
- Molecule 18: PRKR-interacting protein 1

Chain D:



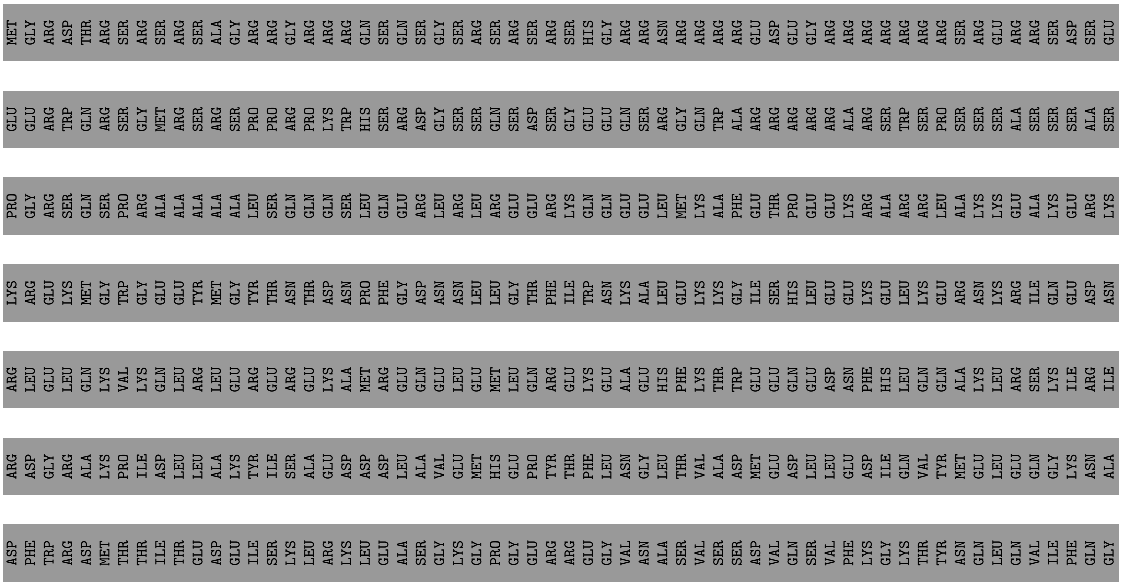
- Molecule 19: PM5 5’ Exon

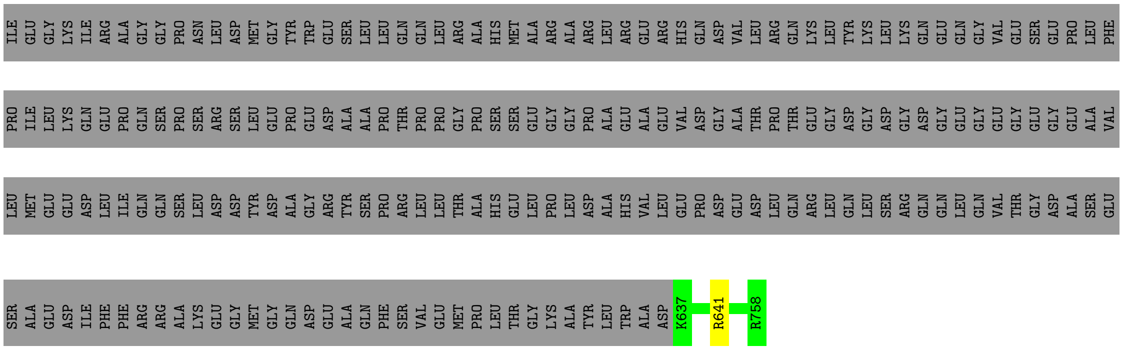
Chain E:



- Molecule 20: Splicing factor Cactin

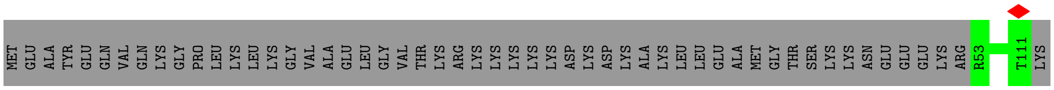
Chain F:





- Molecule 21: Protein FAM32A

Chain G:

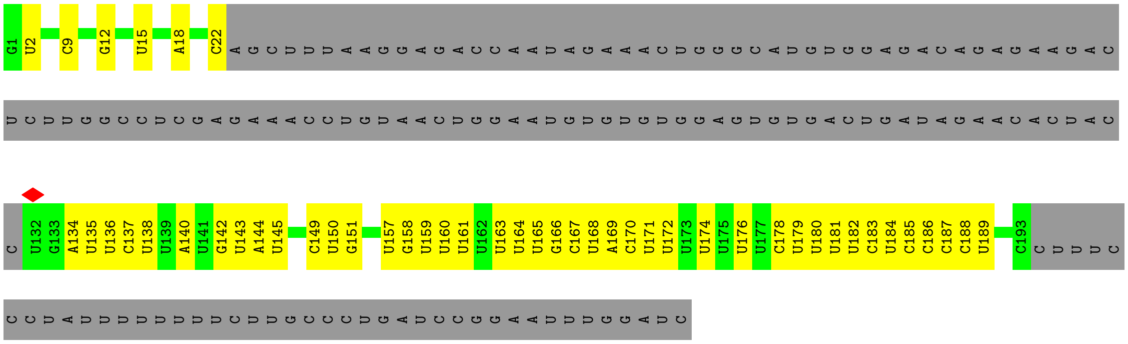



- Molecule 22: Pre-mRNA-splicing factor CWC22 homolog

Chain H:

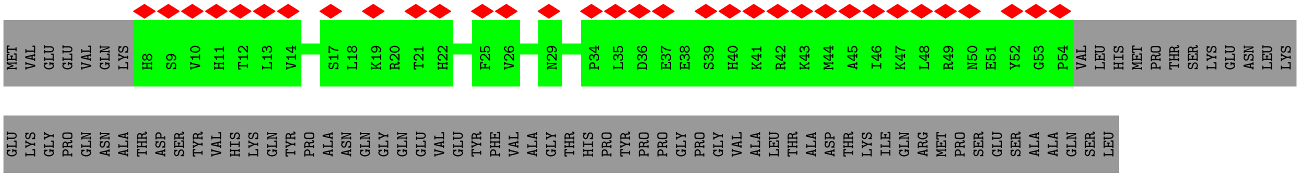

- Molecule 23: PM5 Intron

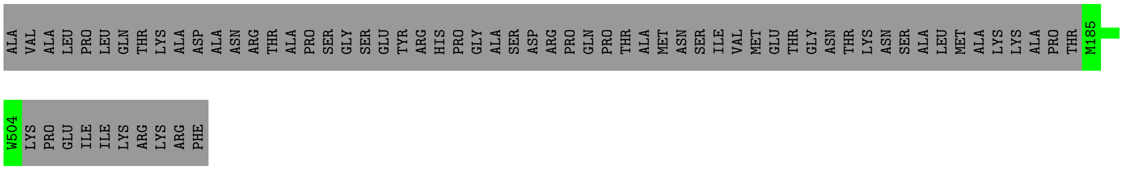
Chain I:

- Molecule 24: Pleiotropic regulator 1

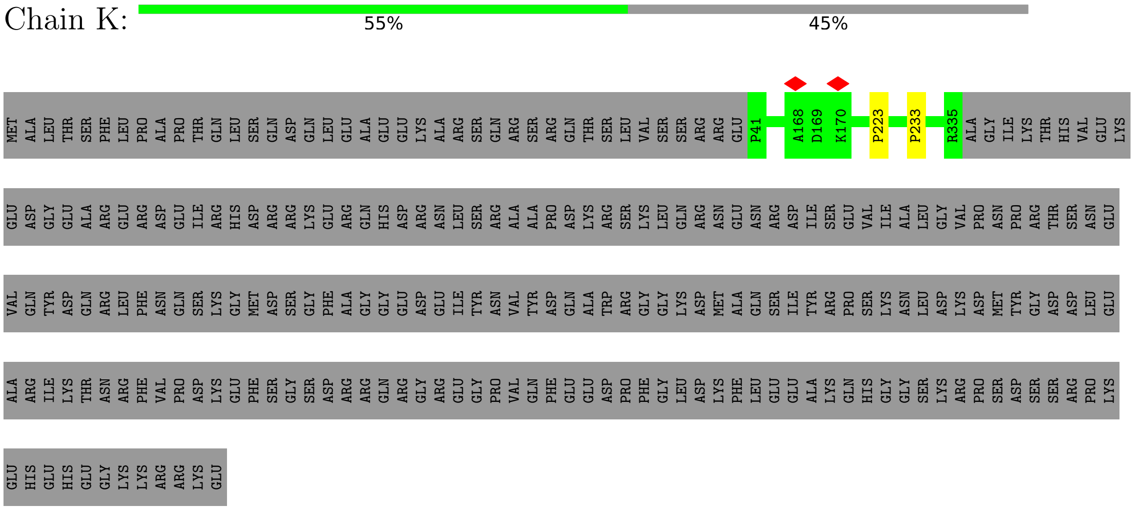
Chain J:



- Molecule 25: SNW domain-containing protein 1

Chain K:



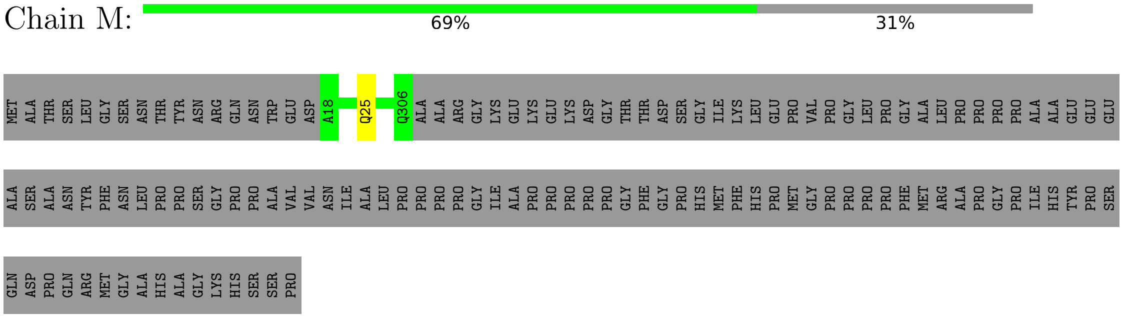
- Molecule 26: Protein BUD31 homolog

Chain L:

There are no outlier residues recorded for this chain.

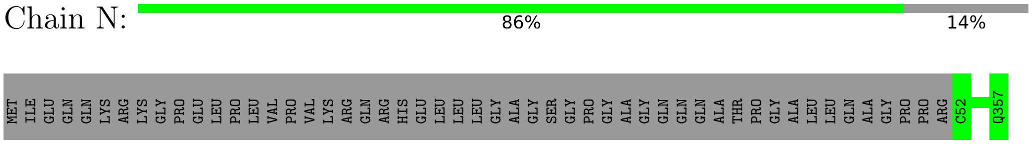
- Molecule 27: Pre-mRNA-splicing factor RBM22

Chain M:

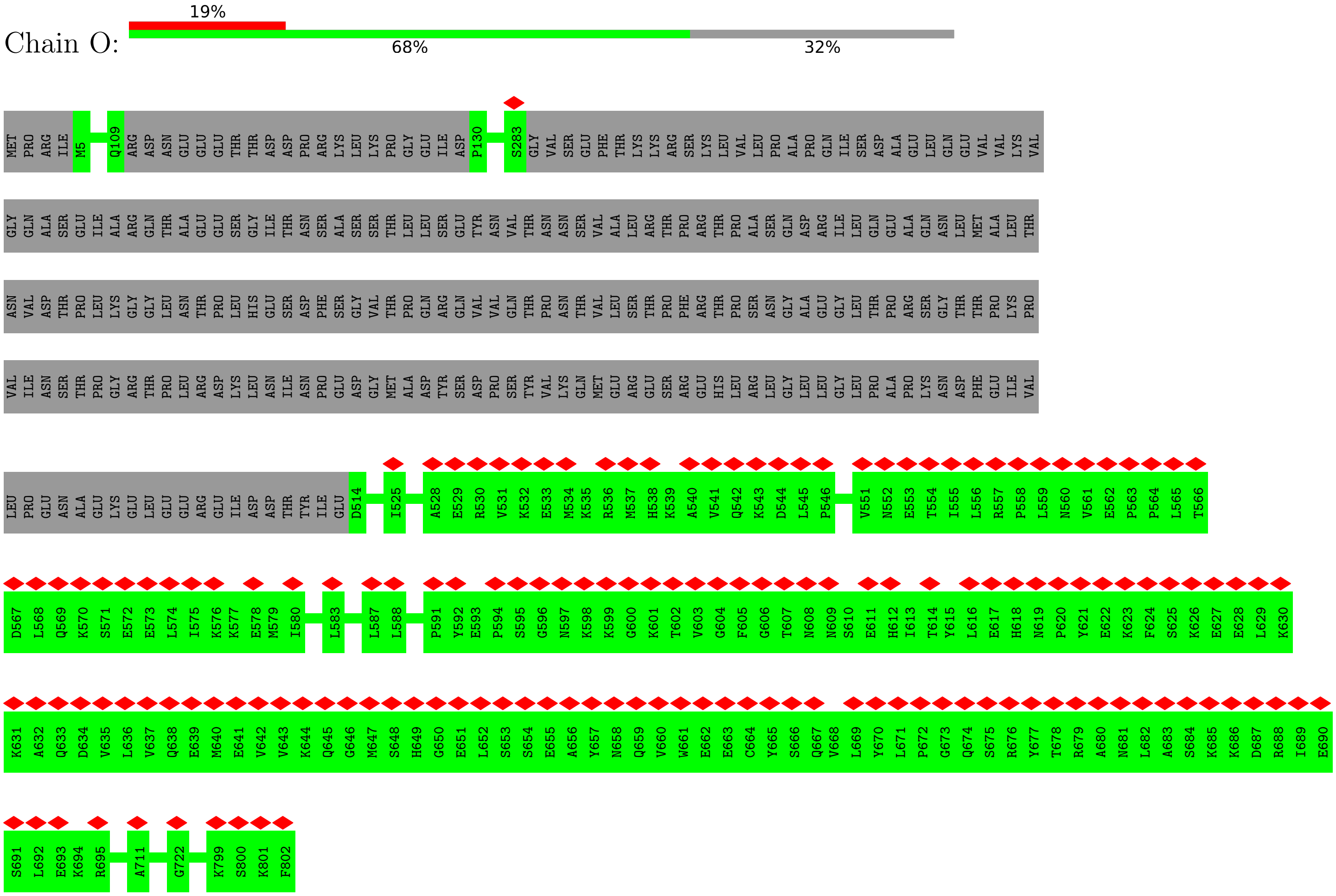


- Molecule 28: U5 small nuclear ribonucleoprotein 40 kDa protein

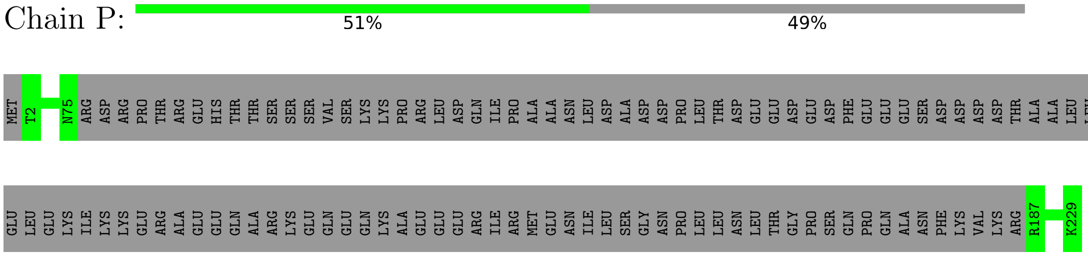
Chain N:



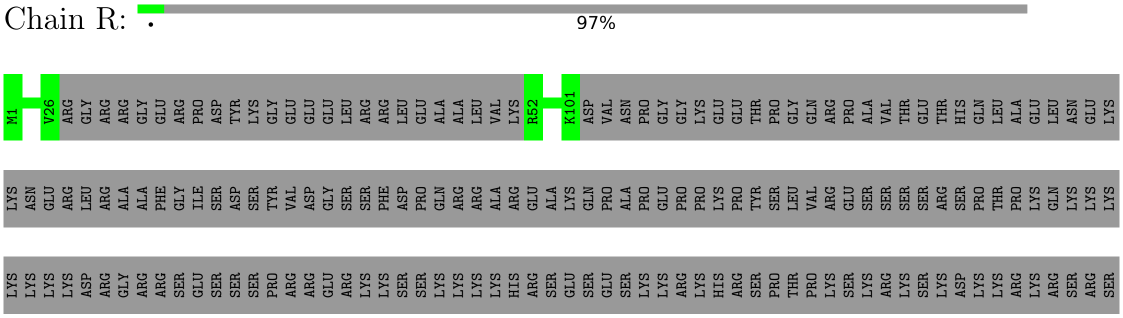
- Molecule 29: Cell division cycle 5-like protein



- Molecule 30: Spliceosome-associated protein CWC15 homolog



- Molecule 31: Serine/arginine repetitive matrix protein 2



THR THR PRO ALA PRO LYS SER ARG ARG ALA HIS ARG SER THR SER ALA ASP SER ALA SER SER SER ASP THR SER ARG SER ARG SER ARG SER ALA ALA ALA LYS THR HIS THR THR ALA LEU ALA GLY ARG SER PRO SER PRO ALA SER GLY ARG ARG GLY GLU GLY ASP ALA PRO PHE

SER GLU PRO GLY THR THR SER THR GLN ARG PRO SER SER PRO GLU THR ALA THR LYS GLN PRO SER SER PRO TYR GLU ASP LYS ASP LYS ASP LYS LYS GLU LYS SER ALA THR ARG PRO SER PRO SER PRO GLU ARG SER SER THR GLY PRO GLU PRO PRO ALA PRO THR PRO LEU LEU

ALA GLU ARG HIS GLY GLY SER PRO GLN PRO LEU ALA THR THR PRO LEU SER GLN GLU PRO VAL ASN PRO PRO SER GLU ALA SER PRO THR ARG ASP ARG SER PRO PRO LYS SER PRO GLU LYS LEU PRO GLN SER SER SER SER GLU SER SER PRO PRO SER PRO GLN PRO THR LYS VAL

SER ARG HIS ALA SER SER SER PRO GLU SER PRO LYS PRO ALA PRO ALA PRO GLY SER HIS ARG GLU ILE SER SER SER PRO THR SER LYS ASN ARG SER HIS GLY ARG ALA LYS ARG ASP LYS SER HIS SER HIS THR PRO SER ARG ARG MET GLY ARG SER ARG SER PRO ALA THR ALA

LYS ARG GLY ARG SER ARG SER ARG THR PRO THR LYS ARG GLY HIS SER ARG SER ARG SER PRO GLN TRP ARG ARG SER ARG SER ALA GLN ARG TRP GLY ARG SER ARG SER PRO GLN ARG ARG GLY ARG SER ARG SER PRO GLN ARG PRO GLY TRP SER ARG SER ARG ASN THR GLN ARG

ARG GLY ARG SER ARG SER ALA ARG ARG GLY ARG SER HIS SER ARG SER PRO ALA THR ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG ARG SER ARG SER ARG THR PRO THR ARG ARG ARG SER ARG SER ARG THR PRO ALA

ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG ARG SER ARG THR ARG SER PRO VAL ARG ARG ARG SER ARG SER ARG SER PRO ALA ARG ARG SER GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG

SER ARG THR PRO ALA ARG ARG SER GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ARG ARG GLY ARG SER ARG SER ARG SER LEU VAL ARG ARG GLY ARG SER HIS SER ARG THR PRO GLN ARG ARG GLY ARG SER GLY SER SER SER GLU ARG

LYS ASN LYS SER ARG THR SER GLN ARG ARG SER ARG SER ASN SER SER PRO GLU MET LYS LYS SER ARG ILE SER SER ARG ARG SER ARG SER LEU SER SER PRO ARG SER LYS ALA LYS SER ARG LEU SER LEU ARG ARG SER LEU SER GLY SER SER PRO CYS PRO LYS GLN LYS SER

GLN THR PRO PRO ARG ARG SER ARG SER GLY SER SER GLN PRO LYS ALA LYS SER ARG THR PRO PRO ARG ARG SER ARG SER SER SER SER PRO PRO PRO LYS GLN LYS SER LYS THR PRO SER ARG GLN SER HIS SER SER SER SER PRO HIS PRO LYS VAL LYS SER GLY THR PRO PRO

ARG GLN GLY SER ILE THR SER PRO GLN ALA ASN GLU GLN SER VAL THR PRO GLN ARG ARG SER CYS PHE GLU SER SER PRO ASP PRO GLU LEU LYS SER ARG THR PRO SER ARG HIS SER CYS SER GLY SER SER PRO PRO ARG VAL LYS SER SER THR PRO PRO ARG GLN SER PRO SER

ARG SER SER SER PRO GLN PRO LYS VAL LYS ALA ILE ILE SER PRO ARG GLN ARG SER HIS SER GLY SER SER SER PRO SER PRO SER ARG VAL THR SER ARG THR THR PRO ARG ARG SER ARG SER VAL SER PRO CYS SER ASN VAL GLU SER ARG LEU LEU PRO ARG TYR SER HIS SER

GLY SER SER SER PRO ASP THR LYS VAL LYS PRO GLU THR PRO PRO ARG GLN SER HIS SER GLY SER ILE SER PRO TYR PRO LYS VAL LYS ALA GLN THR PRO PRO GLY PRO SER LEU SER GLY SER LYS SER PRO CYS PRO GLN GLU LYS SER LYS ASP SER LEU VAL GLN SER CYS PRO

GLY SER LEU SER LEU CYS ALA GLY VAL LYS SER SER THR PRO PRO GLY GLU SER TYR PHE GLY VAL SER SER LEU GLN LEU LYS GLY GLN SER GLN THR SER PRO ASP HIS ARG SER ASP THR SER SER PRO GLU VAL ARG GLN SER HIS SER GLU SER PRO SER LEU GLN SER LYS SER

GLN THR SER PRO LYS GLY GLY ARG SER ARG SER SER SER PRO VAL THR GLU LEU ALA SER ARG SER PRO ILE ARG GLN ASP ARG GLY GLU PHE SER ALA SER PRO MET LEU LYS SER GLY MET SER PRO GLU GLN SER ARG PHE GLN SER ASP SER SER SER TYR PRO THR VAL ASP SER

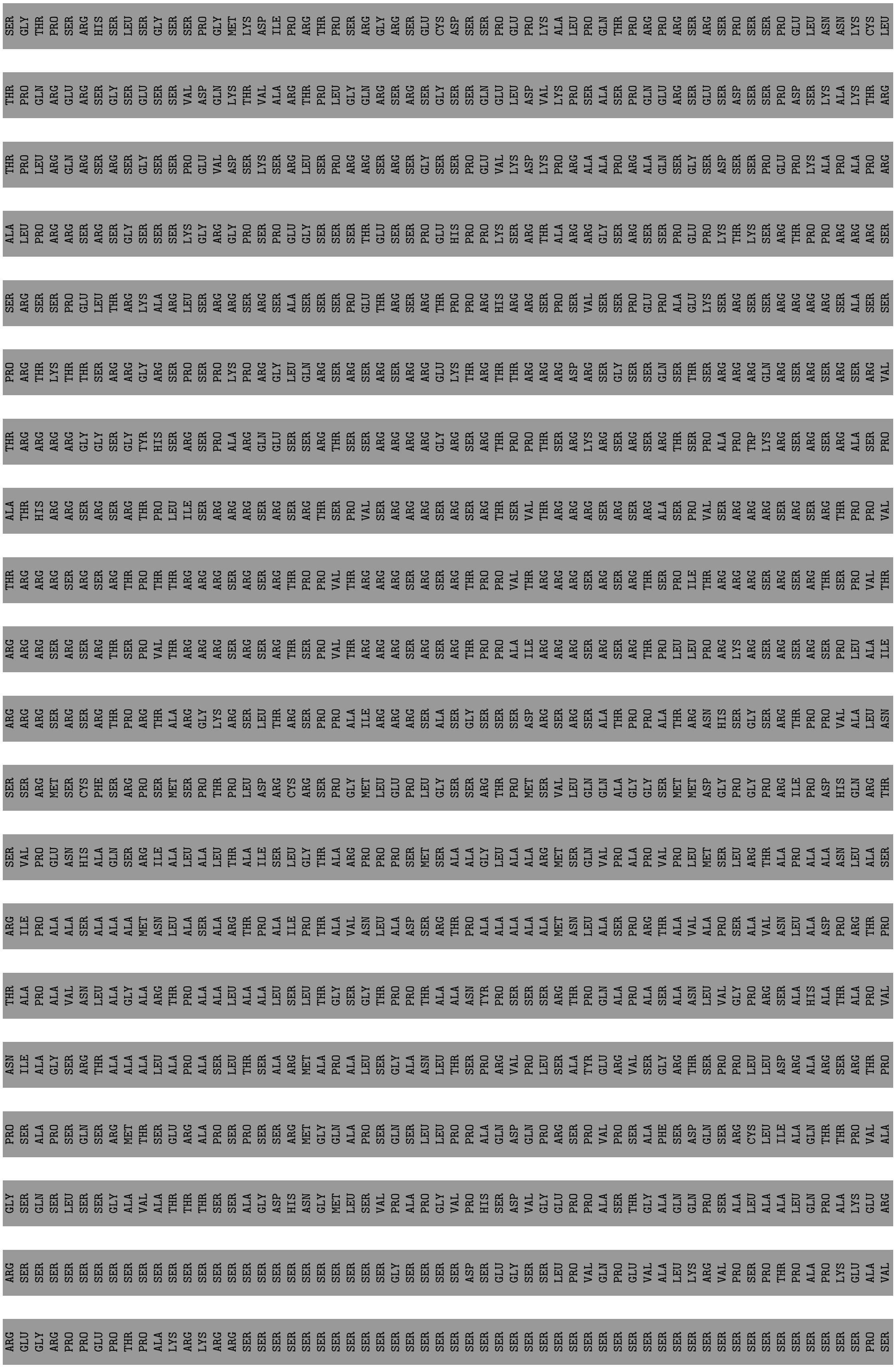
ASN SER LEU LEU GLY GLN SER ARG LEU GLU THR ALA GLU SER LYS GLU LYS MET ALA LEU PRO PRO GLN GLU ASP ALA THR ALA SER PRO PRO ARG GLN LYS ASP LYS PHE SER PRO PHE PRO VAL GLN ASP ARG PRO GLU SER SER LEU VAL PHE LYS ASP THR LEU ARG THR PRO PRO

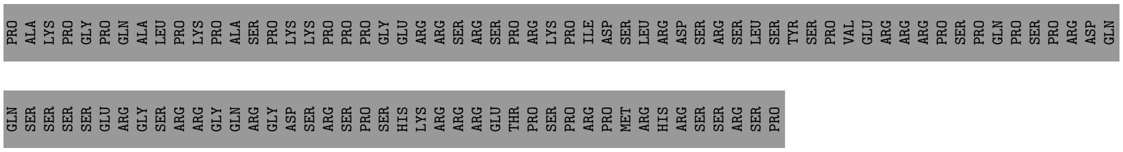
ARG GLU ARG SER GLY ALA GLY SER SER PRO GLU THR LYS GLU GLN ASN SER ALA LEU PRO THR SER SER GLN ASP GLU GLU LEU MET GLU VAL VAL GLU LYS SER GLU GLU PRO ALA GLY GLN ILE LEU SER HIS LEU SER SER GLU LEU LYS GLU MET SER THR SER ASN PHE GLU SER

SER PRO GLU VAL GLU GLU ARG PRO ALA VAL SER LEU THR LEU ASP GLN SER GLN SER GLN ALA SER LEU GLU ALA VAL GLU VAL PRO SER MET ALA SER SER TRP GLY GLY PRO HIS PHE SER PRO GLU HIS LYS GLU LEU SER ASN SER PRO LEU ARG GLU ASN SER PHE GLY SER PRO

LEU GLU PHE ARG ASN SER GLY PRO LEU GLY THR GLU MET ASN THR GLY PHE SER SER GLU VAL LYS GLU ASP LEU ASN GLY PRO PHE LEU ASN GLN LEU GLU THR ASP PRO SER LEU ASP MET LYS GLU GLN SER THR ARG SER SER GLY HIS SER SER SER GLU LEU SER PRO ASP ALA

VAL GLU LYS ALA GLY MET SER SER ASN GLN SER ILE SER SER PRO VAL LEU ASP ALA VAL PRO ARG THR PRO SER ARG GLU ARG SER SER SER ALA SER SER PRO GLU MET LYS ASP GLY LEU PRO ARG THR PRO SER ARG ARG SER ARG SER GLY SER SER PRO GLY LEU ARG ASP GLY


SER GLY THR PRO SER ARG HIS SER LEU SER GLY SER SER PRO GLY MET LYS ASP ILE PRO ARG THR PRO SER ARG GLY ARG SER GLU CYS ASP SER SER PRO GLU PRO LYS ALA LEU PRO GLN THR PRO ARG PRO ARG SER ARG SER PRO SER SER PRO GLU LEU ASN ASN LYS CYS LEU
THR PRO GLN ARG GLU ARG SER GLY SER GLU SER SER VAL ASP GLN LYS THR VAL ALA ARG THR PRO LEU GLY GLN ARG SER ARG SER GLY SER SER GLN GLU LEU ASP VAL LYS PRO SER ALA SER PRO GLN GLU ARG SER GLU SER ASP SER SER PRO ASP SER LYS ALA LYS THR ARG
THR PRO LEU ARG GLN ARG SER ARG SER GLY SER SER PRO GLU VAL ASP SER LYS SER ARG LEU SER PRO ARG ARG SER ARG SER GLY SER SER PRO GLU VAL LYS ASP LYS PRO ARG ALA ALA PRO ARG ALA GLN SER GLY SER ASP SER SER PRO GLU PRO LYS ALA PRO ALA PRO ARG
ALA LEU PRO ARG ARG SER ARG SER GLY SER SER SER LYS GLY ARG GLY PRO SER PRO GLU GLY SER SER SER THR GLU SER SER PRO GLU HIS PRO PRO LYS SER ARG THR ALA ARG ARG GLY SER ARG SER SER PRO GLU PRO LYS THR LYS SER ARG THR PRO PRO ARG ARG ARG SER
SER ARG SER SER PRO GLU LEU THR ARG LYS ALA ARG LEU SER ARG ARG SER ARG SER ALA SER SER SER PRO GLU THR ARG SER ARG THR PRO PRO ARG HIS ARG ARG SER PRO SER VAL SER SER PRO GLU PRO ALA GLU LYS SER ARG SER SER ARG ARG ARG ARG SER ALA SER SER
PRO ARG THR LYS THR THR SER ARG ARG GLY ARG SER PRO SER PRO LYS PRO ARG GLY LEU GLN ARG SER ARG SER ARG SER ARG ARG GLU LYS THR ARG THR THR ARG ARG ARG ASP ARG SER GLY SER SER GLN SER THR SER ARG ARG ARG GLN ARG SER ARG SER ARG SER ARG VAL
THR ARG ARG ARG ARG GLY GLY SER GLY TYR HIS SER ARG SER PRO ALA ARG GLN GLU SER SER ARG THR SER SER ARG ARG ARG ARG GLY ARG SER ARG THR PRO PRO THR SER ARG LYS ARG SER ARG SER ARG THR SER PRO ALA PRO TRP LYS ARG SER ARG SER ARG ALA SER PRO
ALA THR HIS ARG ARG SER ARG SER ARG THR PRO LEU ILE SER ARG ARG ARG SER ARG SER ARG THR SER PRO VAL SER ARG ARG ARG SER ARG SER ARG THR SER VAL THR ARG ARG ARG SER ARG SER ARG ALA SER PRO VAL SER ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL
THR ARG ARG ARG SER ARG SER ARG THR PRO THR THR ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL THR ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL THR ARG ARG ARG SER ARG SER ARG THR SER PRO ILE THR ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR
ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR ARG ARG ARG SER ARG SER ARG THR PRO PRO ALA ILE ARG ARG ARG SER ARG SER ARG THR PRO LEU LEU PRO ARG LYS ARG SER ARG SER ARG SER PRO LEU ALA ILE
ARG ARG ARG SER ARG SER ARG THR PRO ARG THR ALA ARG GLY LYS ARG SER LEU THR ARG SER PRO PRO ALA ILE ARG ARG ARG SER ALA SER GLY SER SER SER ASP ARG SER ARG SER ALA THR PRO PRO ALA THR ARG ASN HIS SER GLY SER ARG THR PRO PRO VAL ALA LEU ASN
SER SER ARG MET SER CYS PHE SER ARG PRO SER MET SER PRO THR PRO LEU ASP ARG CYS ARG SER PRO GLY MET LEU GLU PRO LEU GLY SER SER ARG THR PRO MET SER VAL LEU GLN GLN ALA GLY GLY SER MET MET ASP GLY PRO GLY PRO ARG ILE PRO ASP HIS GLN ARG THR
SER VAL PRO GLU ASN HIS ALA GLN SER ARG ILE ALA LEU ALA LEU THR ALA ILE SER LEU GLY THR ALA ARG PRO PRO PRO SER MET SER ALA ALA GLY LEU ALA ALA ARG MET SER GLN VAL PRO ALA PRO VAL PRO LEU MET SER LEU ARG THR ALA PRO ALA ALA ASN LEU ALA SER
ARG ILE PRO ALA ALA SER ALA ALA ALA MET ASN LEU ALA SER ALA ARG THR PRO ALA ILE PRO THR ALA VAL ASN LEU ALA ASP SER ARG THR PRO ALA ALA ALA ALA ALA MET ASN LEU ALA SER PRO ARG THR ALA VAL ALA PRO SER ALA VAL ASN LEU ALA ASP PRO ARG THR PRO
THR ALA PRO ALA VAL ASN LEU ALA GLY ALA ARG THR PRO ALA ALA LEU ALA ALA LEU SER LEU THR GLY SER GLY THR PRO PRO THR ALA ALA ASN TYR PRO SER SER SER ARG THR PRO GLN ALA PRO ALA SER ALA ASN LEU VAL GLY PRO ARG SER ALA HIS ALA THR ALA PRO VAL
ASN ILE ALA GLY SER ARG THR ALA ALA ALA LEU ALA PRO ALA SER LEU THR SER ALA ARG MET ALA PRO ALA LEU SER GLY ALA ASN LEU THR SER PRO ARG VAL PRO LEU SER ALA TYR GLU ARG VAL SER GLY ARG THR SER PRO PRO LEU LEU ASP ARG ALA ARG SER ARG THR PRO
PRO SER ALA PRO SER GLN SER ARG MET THR SER GLU ARG ALA PRO SER PRO SER SER ARG MET GLY GLN ALA PRO SER GLN SER LEU LEU PRO PRO ALA GLN ASP GLN PRO ARG SER PRO VAL PRO SER ALA PHE SER ASP GLN SER ARG CYS LEU ILE ALA GLN THR THR PRO VAL ALA
GLY SER GLN SER LEU SER SER GLY ALA VAL ALA THR THR THR SER SER ALA GLY ASP HIS ASN GLY MET LEU SER VAL PRO ALA PRO GLY VAL PRO HIS SER ASP VAL GLY GLU PRO PRO ALA SER THR GLY ALA GLN GLN PRO SER ALA LEU ALA ALA LEU GLN PRO ALA LYS GLU ARG
ARG SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER GLY SER SER SER SER ASP SER GLU GLY SER SER LEU PRO VAL GLN PRO GLU VAL ALA LEU LYS ARG VAL PRO SER PRO THR PRO ALA PRO LYS GLU ALA VAL
ARG GLU GLY ARG PRO PRO GLU PRO THR PRO ALA LYS ARG LYS ARG ARG SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER PRO SER

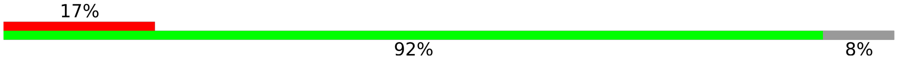


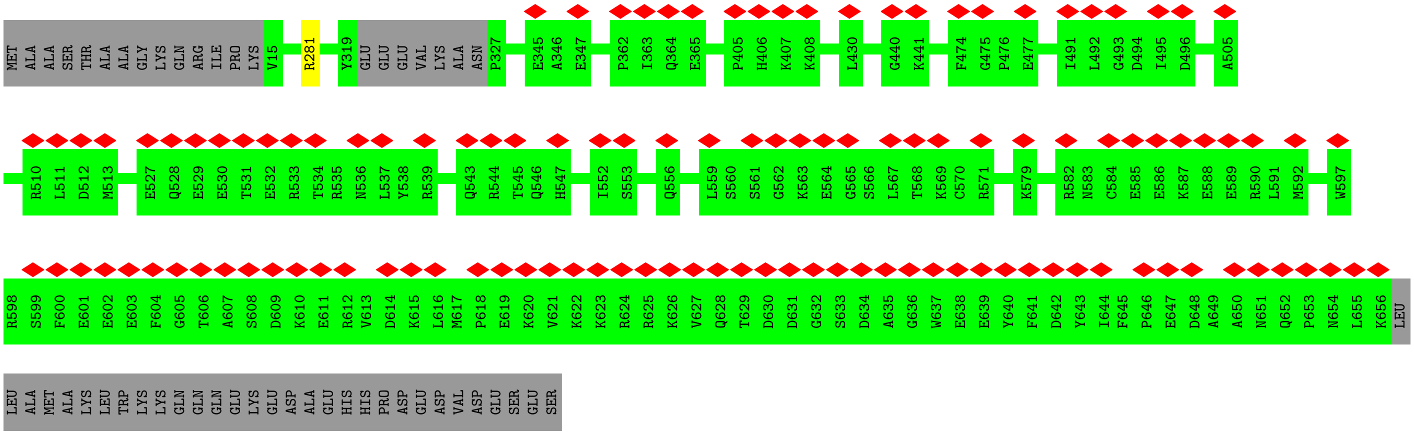
PRO ALA LYS PRO GLY PRO GLN ALA LEU PRO LYS PRO ALA SER PRO LYS LYS PRO PRO PRO GLY GLU ARG ARG SER ARG SER PRO ARG LYS PRO ILE ASP SER LEU ARG ASP SER ARG SER LEU SER TYR SER PRO VAL GLU ARG ARG ARG PRO SER PRO GLN PRO SER PRO ARG ASP GLN

GLN SER SER SER SER GLU ARG GLY SER ARG ARG GLY GLN ARG GLY ASP SER ARG SER PRO SER HIS LYS ARG ARG ARG GLU THR PRO SER PRO ARG PRO MET ARG HIS ARG SER SER ARG SER PRO

- Molecule 32: Crooked neck-like protein 1

Chain S:





MET ALA ALA SER THR ALA ALA GLY LYS GLN ARG ILE PRO LYS V15 R281 Y319 GLU GLU GLU VAL LYS ALA ASN P327 E345 A346 E347 P362 I363 Q364 E365 P405 H406 K407 K408 L430 G440 K441 F474 G475 P476 E477 I491 L492 G493 D494 I495 D496 A505

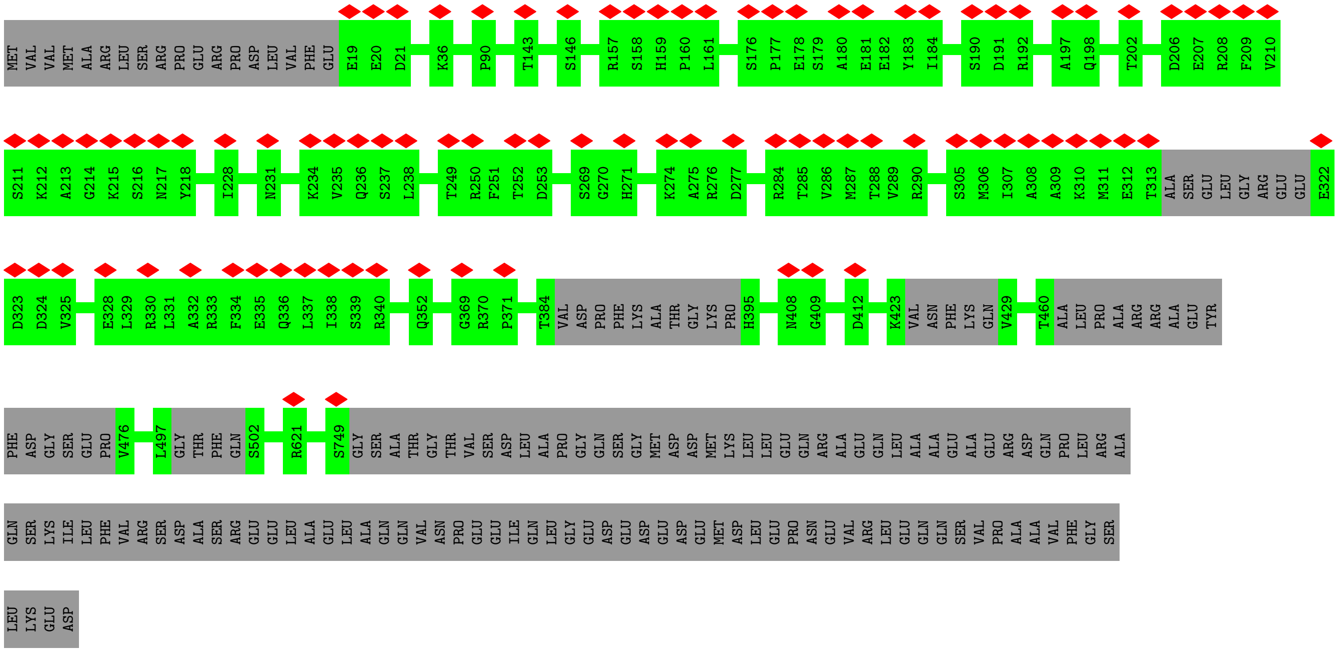
R510 L511 D512 M513 E527 Q528 E529 E530 T531 E532 R533 T534 R535 N536 L537 Y538 R539 Q543 R544 T545 Q546 H547 I552 S553 Q556 L559 S560 S561 G562 K563 E564 G565 S566 L567 T568 K569 C570 R571 K579 R582 N583 C584 E585 E586 K587 E588 E589 R590 L591 M592 W597

R598 S599 F600 E601 E602 E603 F604 G605 T606 A607 S608 D609 K610 E611 R612 V613 D614 K615 L616 M617 P618 E619 K620 V621 K622 K623 R624 R625 K626 V627 Q628 T629 D630 D631 G632 S633 D634 A635 G636 W637 E638 E639 Y640 F641 D642 Y643 I644 F645 P646 E647 D648 A649 A650 N651 Q652 P653 N654 L655 K656 LEU

LEU ALA MET ALA LYS LEU TRP LYS LYS GLN GLN GLN GLU LYS GLU ASP ALA GLU HIS HIS PRO ASP GLU ASP VAL ASP GLU SER GLU SER

- Molecule 33: Pre-mRNA-splicing factor SYF1

Chain T:



MET VAL VAL MET ALA ARG LEU SER ARG PRO GLU ARG PRO ASP LEU VAL PHE GLU E19 E20 D21 K36 P90 T143 S146 R157 S158 H159 P160 L161 S176 P177 E178 S179 A180 E181 E182 Y183 I184 S190 D191 R192 A197 Q198 T202 D206 E207 R208 F209 V210

S211 K212 A213 G214 K215 S216 N217 Y218 I228 N231 K234 V235 Q236 S237 L238 T249 R250 F251 T252 D253 S269 G270 H271 K274 A275 R276 D277 R284 T285 V286 M287 T288 V289 R290 S305 M306 I307 A308 A309 K310 M311 E312 T313 ALA SER GLU LEU GLY ARG GLU GLU E322

D323 D324 V325 E328 L329 R330 L331 A332 R333 F334 E335 Q336 L337 I338 S339 R340 Q352 G369 R370 P371 T384 VAL ASP PRO PHE LYS ALA THR GLY LYS PRO H395 N408 G409 D412 K423 VAL ASN PHE LYS GLN V429 T460 ALA LEU PRO ALA ARG ARG ALA GLU TYR

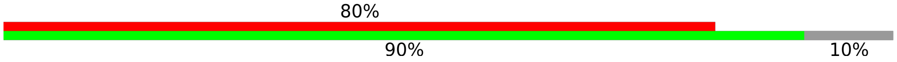
PHE ASP GLY SER GLU PRO V476 L497 GLY THR PHE GLN S502 R621 S749 GLY SER ALA THR GLY THR VAL SER ASP LEU ALA PRO GLY GLN SER GLY MET ASP ASP MET LYS LEU LEU GLU GLN ARG ALA GLU GLN LEU ALA ALA GLU ALA GLU ARG ASP GLN PRO LEU ARG ALA

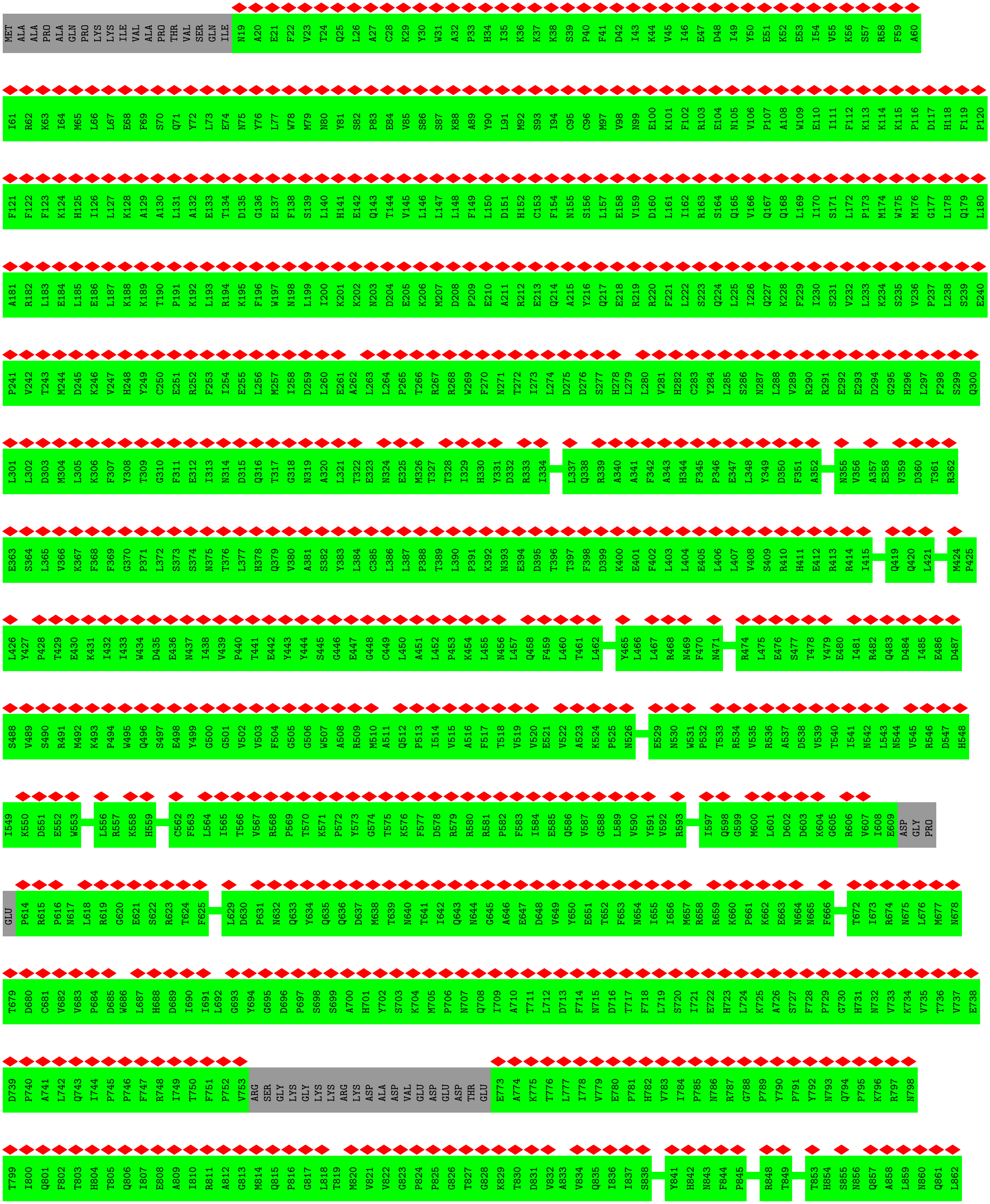
GLN SER LYS ILE LEU PHE VAL ARG SER ASP ALA SER ARG GLU GLU LEU ALA GLU LEU ALA GLN GLN VAL ASN PRO GLU GLU ILE GLN LEU GLY GLU ASP GLU ASP GLU ASP GLU MET ASP LEU GLU PRO ASN GLU VAL ARG LEU GLU GLN GLN SER VAL PRO ALA ALA VAL PHE GLY SER

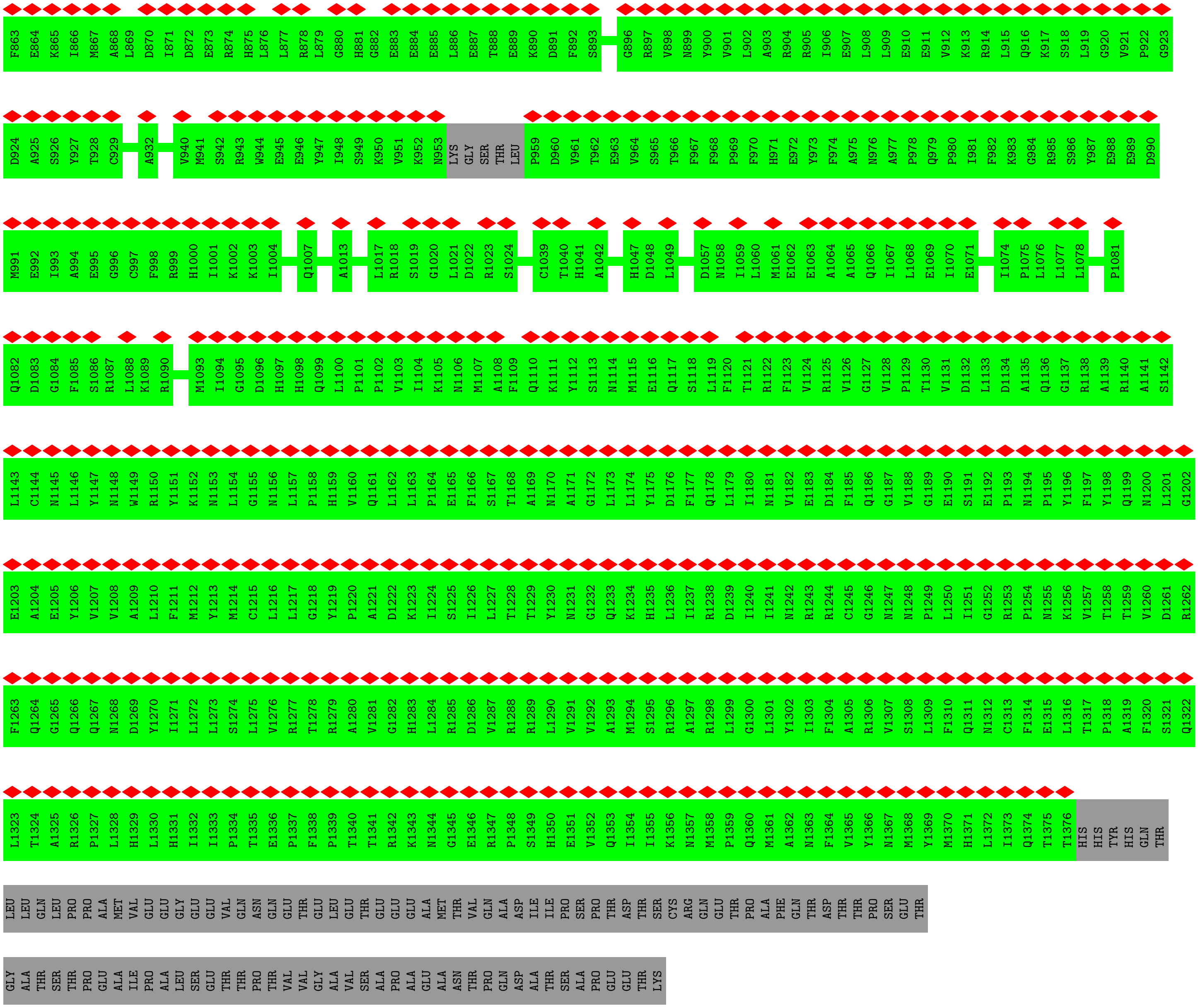
LEU LYS GLU ASP

- Molecule 34: Intron-binding protein aquarius

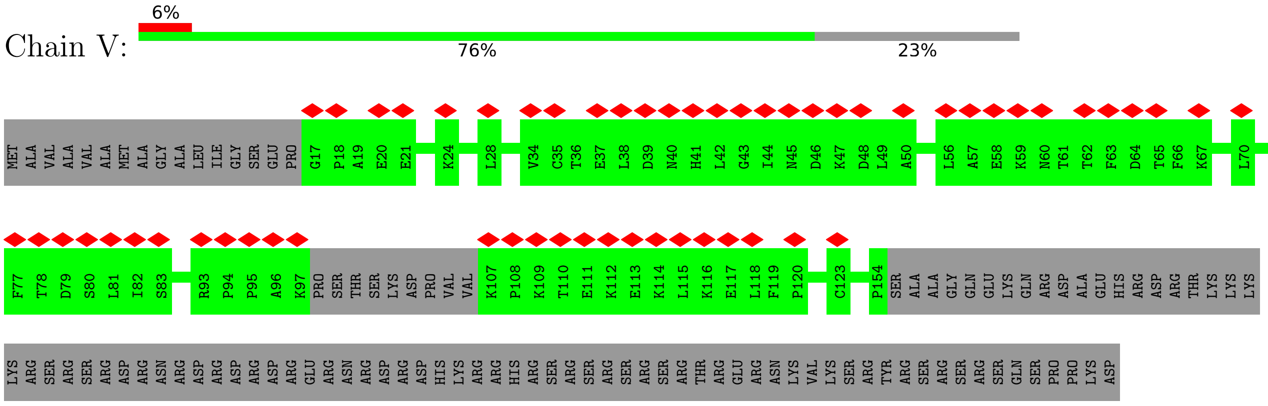
Chain U:

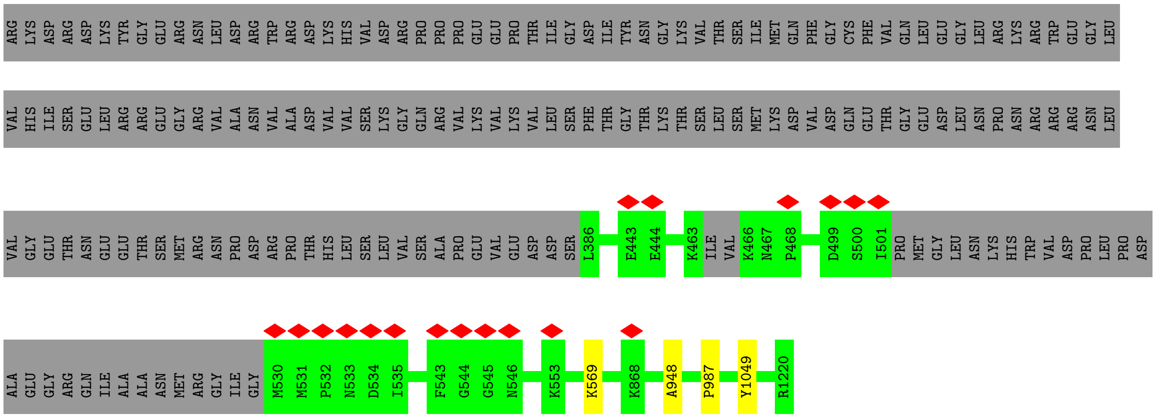



MET ALA ALA PRO ALA GLN PRO LYS LYS ILE VAL ALA PRO THR VAL SER GLN ILE N19 A20 E21 F22 V23 T24 Q25 L26 A27 C28 K29 Y30 W31 A32 P33 H34 I35 K36 K37 K38 S39 P40 F41 D42 I43 K44 V45 I46 E47 D48 I49 Y50 E51 K52 E53 I54 V55 K56 S57 R58 F59 A60
I61 R62 K63 I64 M65 L66 L67 E68 F69 S70 Q71 Y72 L73 E74 N75 Y76 L77 W78 M79 N80 Y81 S82 P83 E84 V85 S86 S87 K88 A89 Y90 L91 M92 S93 I94 C95 C96 M97 V98 N99 E100 K101 F102 R103 E104 N105 V106 P107 A108 W109 E110 I111 F112 K113 K114 K115 P116 D117 H118 F119 P120
F121 F122 F123 K124 H125 I126 L127 K128 A129 A130 L131 A132 E133 T134 D135 G136 E137 F138 S139 L140 H141 E142 Q143 T144 V145 L146 L147 L148 F149 L150 D151 H152 C153 F154 N155 S156 L157 E158 V159 D160 L161 I162 R163 S164 Q165 V166 Q167 Q168 L169 I170 S171 L172 P173 M174 W175 M176 G177 L178 Q179 L180
A181 R182 L183 E184 L185 E186 L187 K188 K189 T190 P191 K192 L193 R194 K195 F196 W197 N198 L199 I200 K201 K202 N203 D204 E205 K206 M207 D208 P209 E210 A211 R212 E213 Q214 A215 Y216 Q217 E218 R219 R220 F221 L222 S223 Q224 L225 I226 Q227 K228 F229 I230 S231 V232 L233 K234 S235 V236 P237 L238 S239 E240
P241 V242 T243 M244 D245 K246 V247 H248 Y249 C250 E251 R252 F253 I254 E255 L256 M257 I258 D259 L260 E261 A262 L263 L264 P265 T266 R267 R268 W269 F270 N271 T272 I273 L274 D275 D276 S277 H278 L279 L280 V281 H282 C283 Y284 L285 S286 N287 L288 V289 R290 R291 E292 E293 D294 G295 H296 L297 F298 S299 Q300
L301 L302 D303 M304 L305 K306 F307 Y308 T309 G310 F311 E312 I313 N314 D315 Q316 T317 G318 N319 A320 L321 T322 E323 N324 E325 M326 T327 T328 I329 H330 Y331 D332 R333 I334 L337 Q338 R339 A340 A341 F342 A343 H344 F345 P346 E347 L348 Y349 D350 F351 A352 N355 V356 A357 E358 V359 D360 T361 R362
E363 S364 L365 V366 K367 F368 F369 G370 P371 L372 S373 S374 N375 T376 L377 H378 Q379 V380 A381 S382 Y383 L384 C385 L386 L387 P388 T389 L390 P391 K392 N393 E394 D395 T396 T397 F398 D399 K400 E401 F402 L403 L404 E405 L406 L407 V408 S409 R410 H411 E412 R413 R414 I415 Q419 Q420 L421 M424 P425
L426 Y427 P428 T429 E430 K431 I432 I433 W434 D435 E436 N437 I438 V439 P440 T441 E442 Y443 Y444 S445 G446 E447 G448 C449 L450 A451 L452 P453 K454 L455 N456 L457 Q458 F459 L460 T461 L462 Y465 L466 L467 R468 N469 F470 N471 R474 L475 E476 S477 T478 Y479 E480 I481 R482 Q483 D484 I485 E486 D487
S488 V489 S490 R491 M492 K493 P494 W495 Q496 S497 E498 Y499 G500 G501 V502 V503 F504 G505 G506 W507 A508 R509 M510 A511 Q512 P513 I514 V515 A516 F517 T518 V519 V520 E521 V522 A523 K524 P525 N526 E529 N530 W531 P532 T533 R534 V535 R536 A537 D538 V539 T540 I541 N542 L543 N544 V545 R546 D547 H548
I549 K550 D551 E552 W553 L556 R557 K558 H559 C562 F563 L564 I565 T566 V567 R568 P569 T570 K571 P572 Y573 G574 T575 K576 F577 D578 R579 R580 R581 P582 F583 I584 E585 Q586 V587 G588 L589 V590 Y591 V592 R593 I597 Q598 G599 M600 L601 D602 D603 K604 G605 R606 V607 I608 E609 ASP GLY PRO
GLU P614 R615 P616 N617 L618 R619 G620 E621 S622 R623 T624 F625 L629 D630 P631 N632 Q633 Y634 Q635 Q636 D637 M638 T639 N640 T641 I642 Q643 N644 G645 A646 E647 D648 V649 Y650 E651 T652 F653 N654 I655 I656 M657 R658 R659 K660 P661 K662 E663 N664 N665 F666 T672 I673 R674 N675 L676 M677 N678
T679 D680 C681 V682 V683 P684 D685 W686 L687 H688 D689 I690 I691 L692 G693 Y694 G695 D696 P697 S698 S699 A700 H701 Y702 S703 K704 M705 P706 N707 Q708 I709 A710 T711 L712 D713 F714 N715 D716 T717 F718 L719 S720 I721 E722 H723 L724 K725 A726 S727 F728 P729 G730 H731 N732 V733 K734 V735 T736 V737 E738
D739 P740 A741 L742 Q743 I744 P745 P746 F747 R748 I749 T750 F751 P752 V753 ARG SER GLY LYS GLY LYS LYS ARG LYS ASP ALA ASP VAL GLU ASP GLU ASP THR GLU E773 A774 K775 T776 L777 I778 V779 E780 P781 H782 V783 I784 P785 N786 R787 G788 P789 Y790 P791 Y792 N793 Q794 P795 K796 R797 N798
T799 I800 Q801 F802 T803 H804 T805 Q806 I807 E808 A809 I810 R811 A812 G813 M814 Q815 P816 G817 L818 T819 M820 V821 V822 G823 P824 P825 G826 T827 G828 K829 T830 D831 V832 A833 V834 Q835 I836 I837 S838 Y841 H842 N843 F844 P845 R848 T849 T853 H854 S855 N856 Q857 A858 L859 N860 Q861 L862

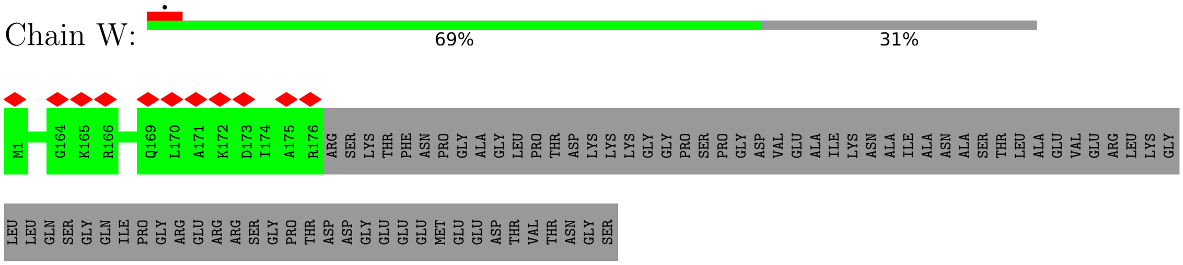


- Molecule 35: ATP-dependent RNA helicase DHX8

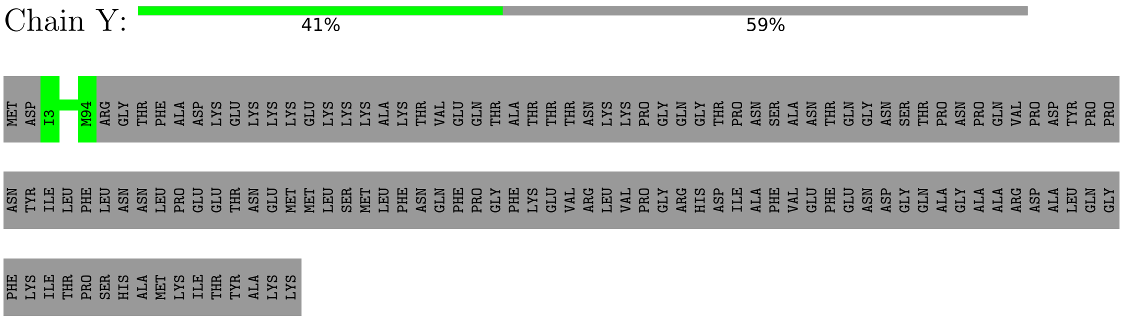




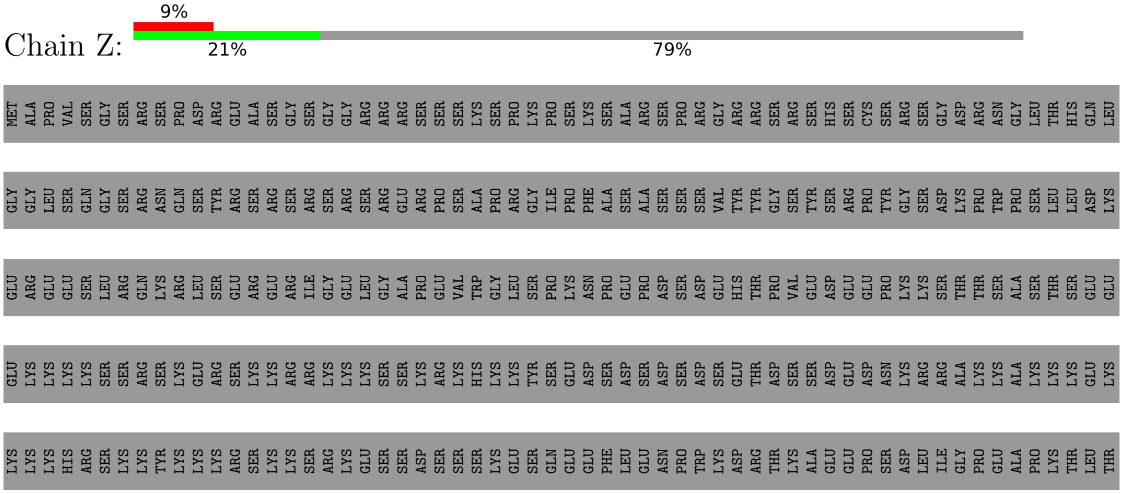
- Molecule 36: U2 small nuclear ribonucleoprotein A'

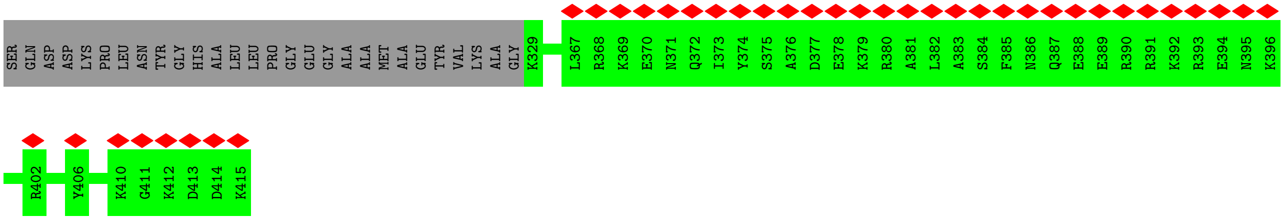


- Molecule 37: U2 small nuclear ribonucleoprotein B”



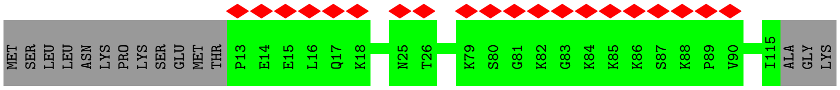
- Molecule 38: NF-kappa-B-activating protein





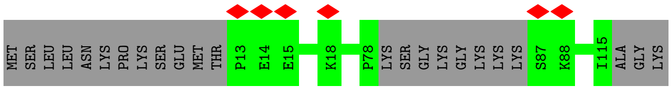
- Molecule 39: Small nuclear ribonucleoprotein Sm D2

Chain a:



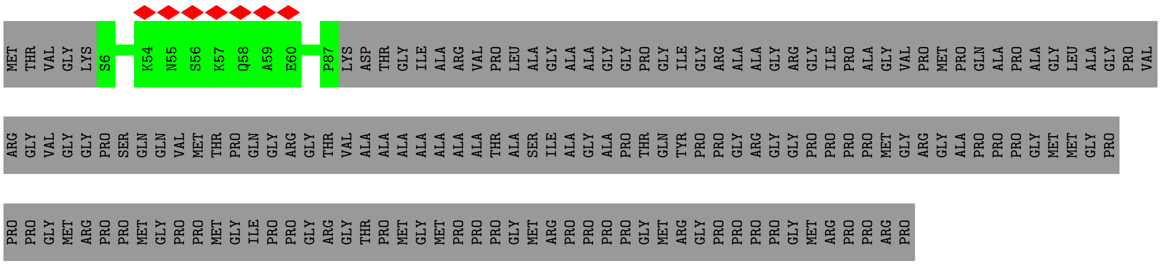
- Molecule 39: Small nuclear ribonucleoprotein Sm D2

Chain m:



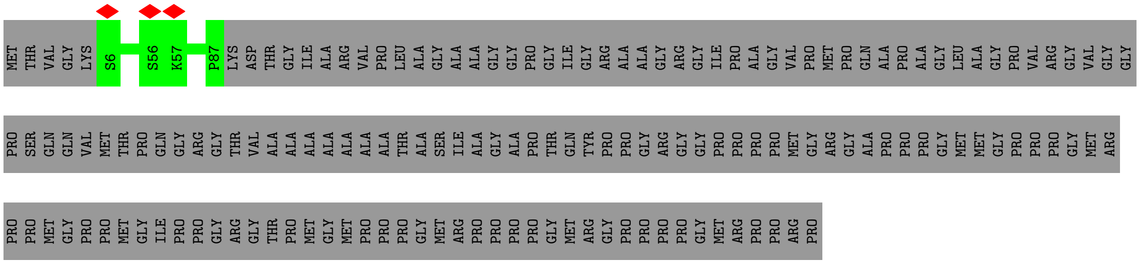
- Molecule 40: Small nuclear ribonucleoprotein-associated proteins B and B’

Chain b:



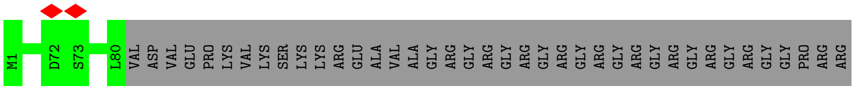
- Molecule 40: Small nuclear ribonucleoprotein-associated proteins B and B’

Chain k:



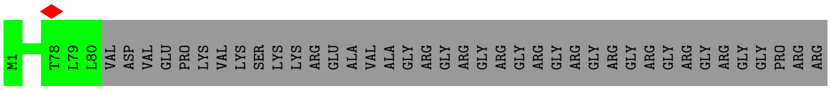
- Molecule 41: Small nuclear ribonucleoprotein Sm D1

Chain c:



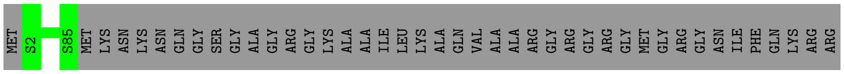
- Molecule 41: Small nuclear ribonucleoprotein Sm D1

Chain l:



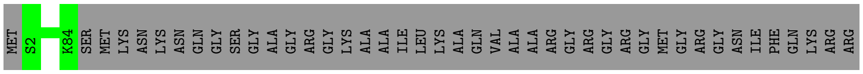
- Molecule 42: Small nuclear ribonucleoprotein Sm D3

Chain d:



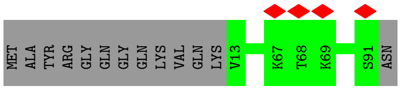
- Molecule 42: Small nuclear ribonucleoprotein Sm D3

Chain n:



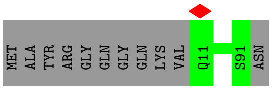
- Molecule 43: Small nuclear ribonucleoprotein E

Chain e:



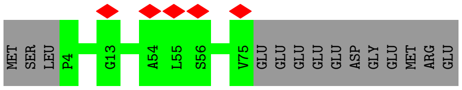
- Molecule 43: Small nuclear ribonucleoprotein E

Chain p:



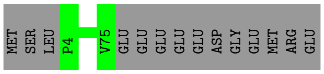
- Molecule 44: Small nuclear ribonucleoprotein F

Chain f:



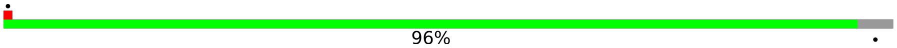
- Molecule 44: Small nuclear ribonucleoprotein F

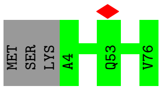
Chain q:



- Molecule 45: Small nuclear ribonucleoprotein G

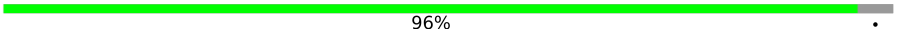
Chain g:

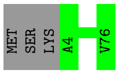




- Molecule 45: Small nuclear ribonucleoprotein G

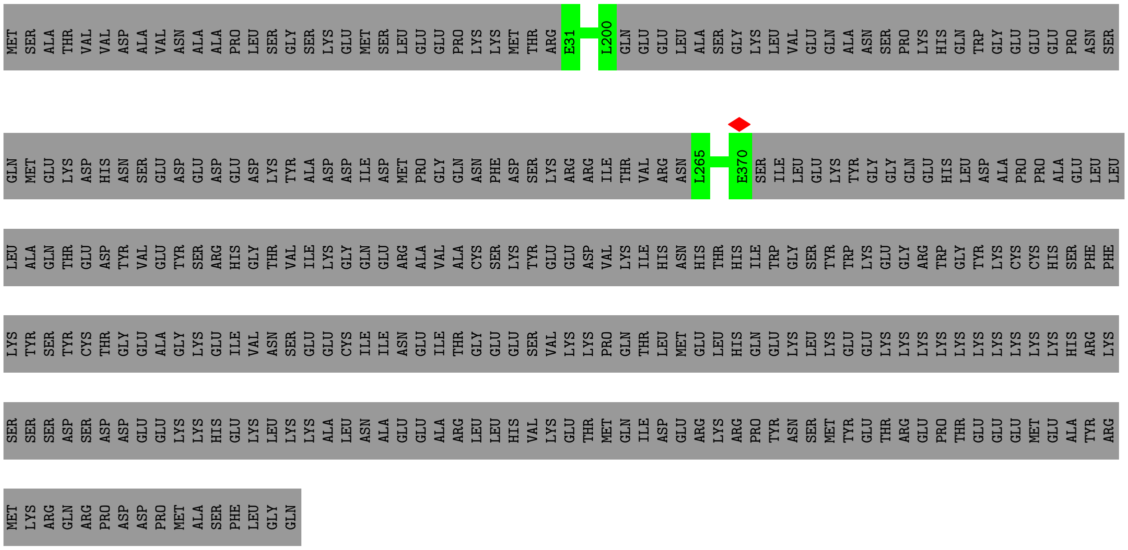
Chain r:





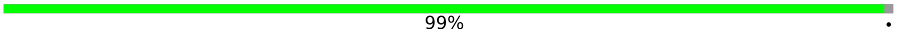
- Molecule 46: Pre-mRNA-splicing factor SLU7

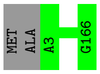
Chain h:



- Molecule 47: Peptidyl-prolyl cis-trans isomerase-like 1

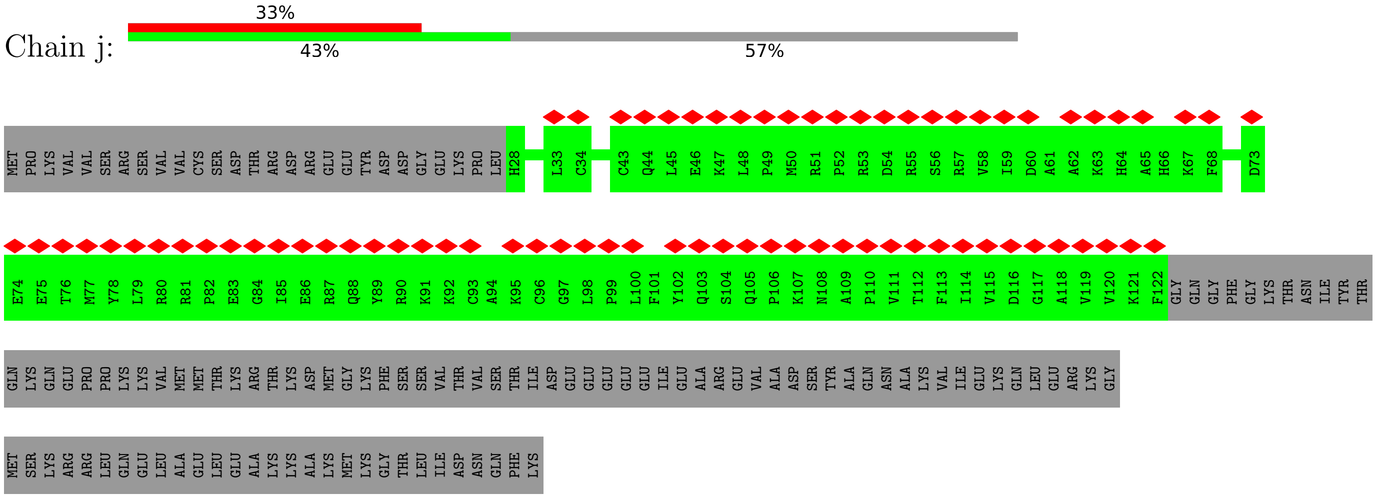
Chain i:





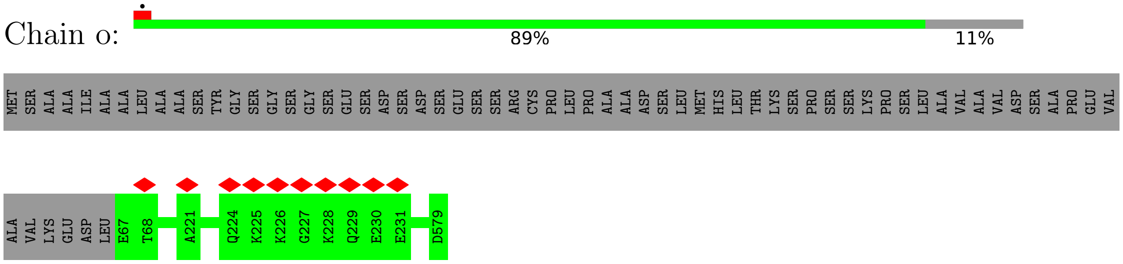
- Molecule 48: STING ER exit protein

Chain j:



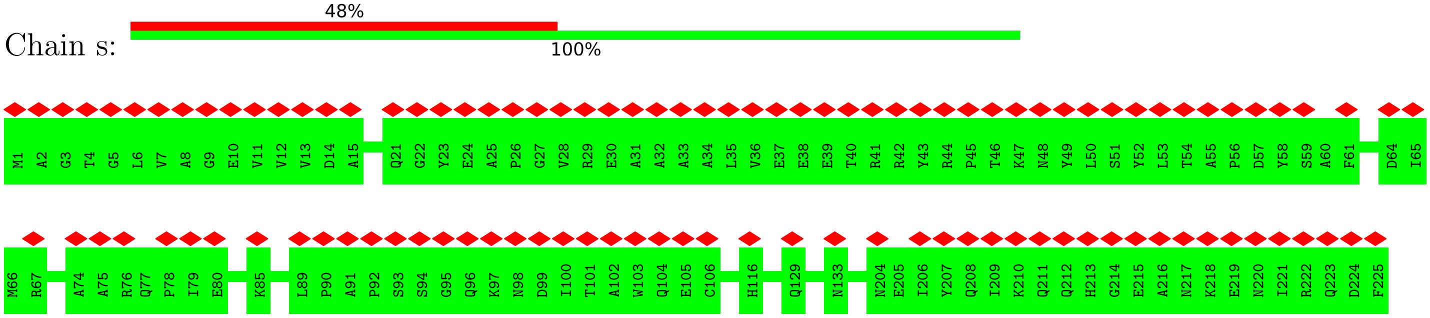
- Molecule 49: Pre-mRNA-processing factor 17

Chain o:



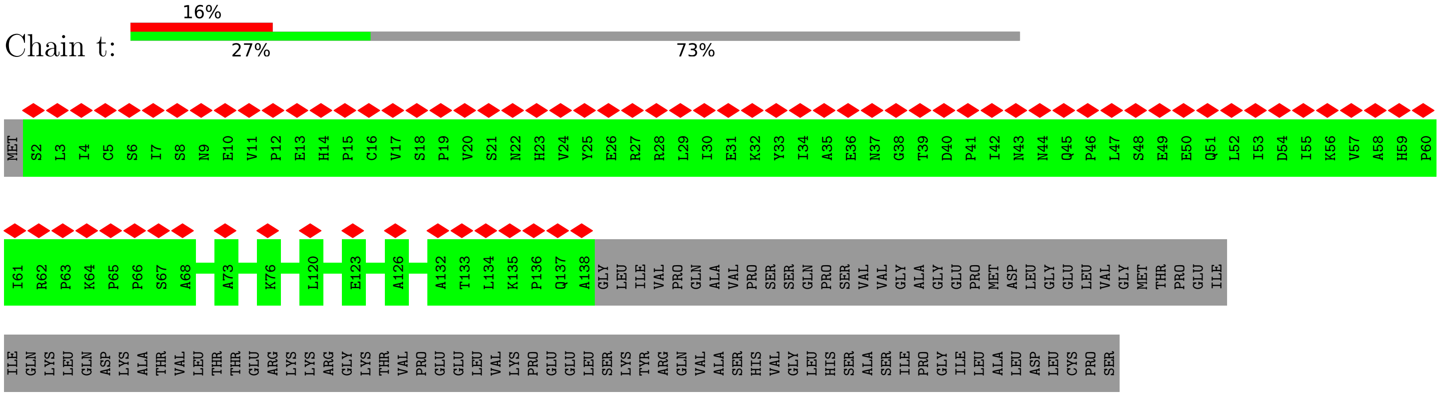
- Molecule 50: Pre-mRNA-splicing factor SPF27

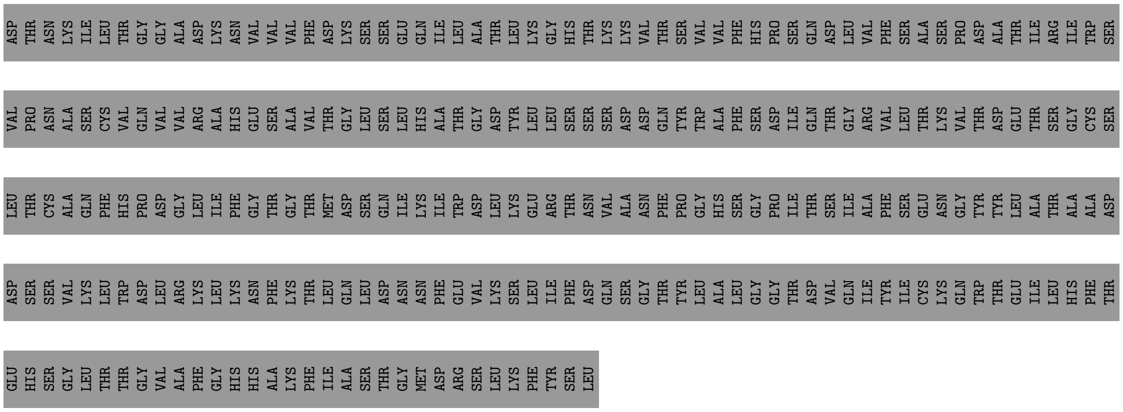
Chain s:



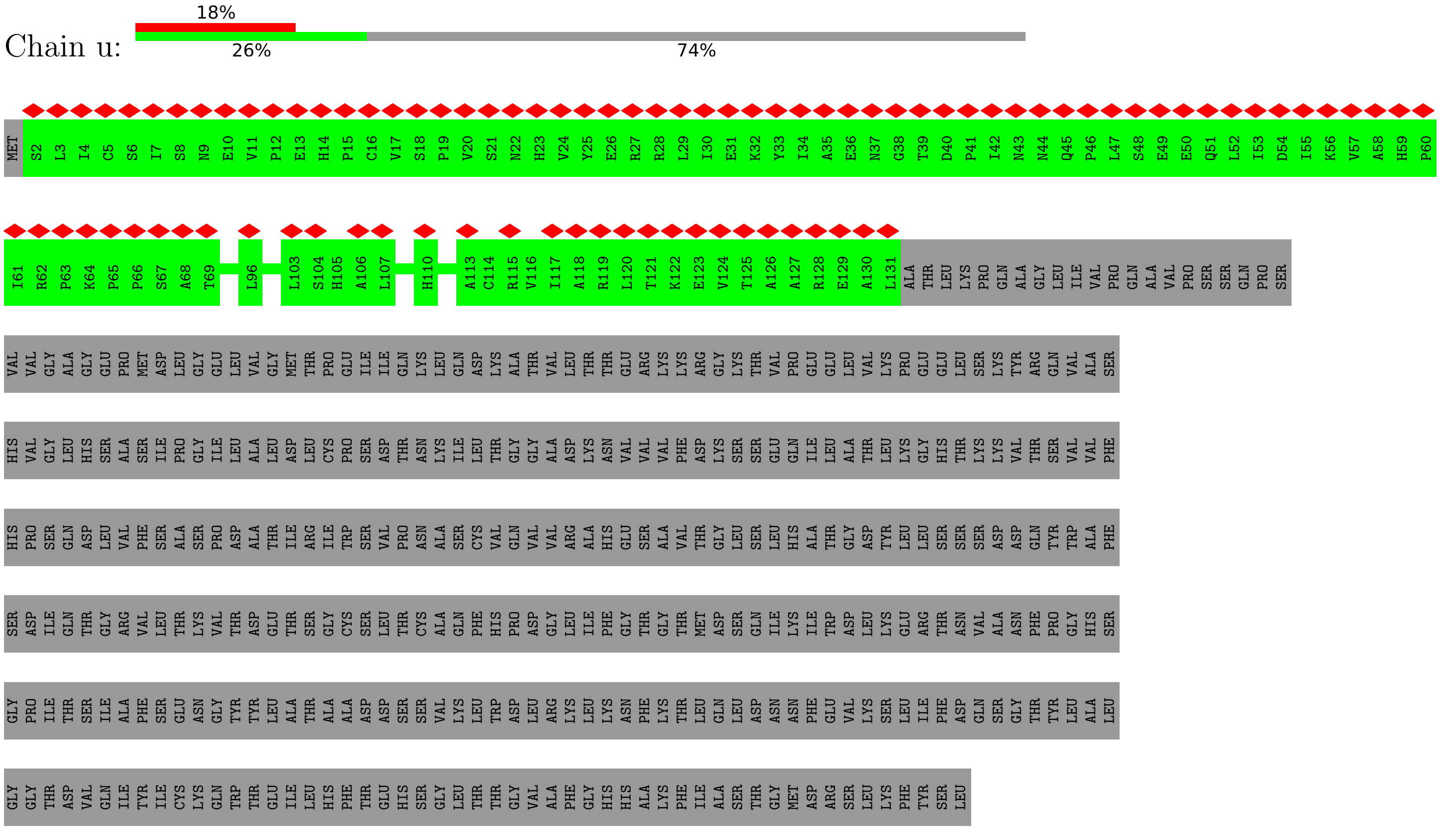
- Molecule 51: Pre-mRNA-processing factor 19

Chain t:

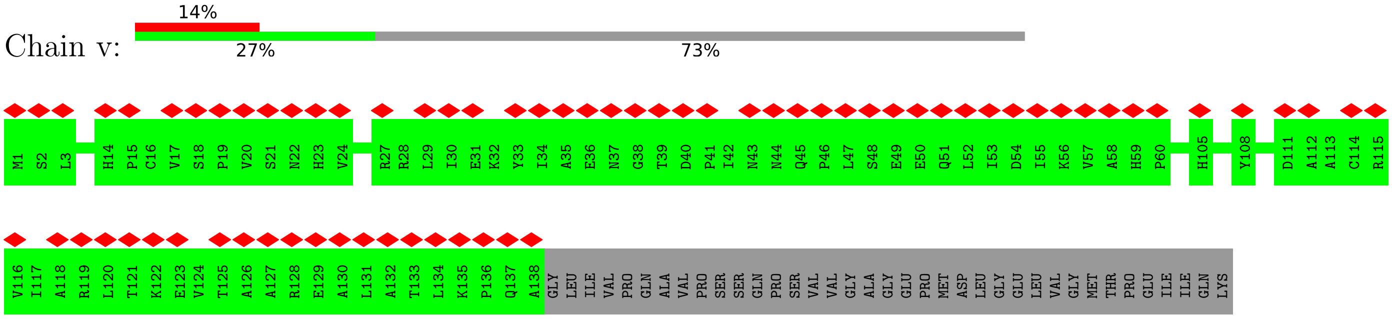


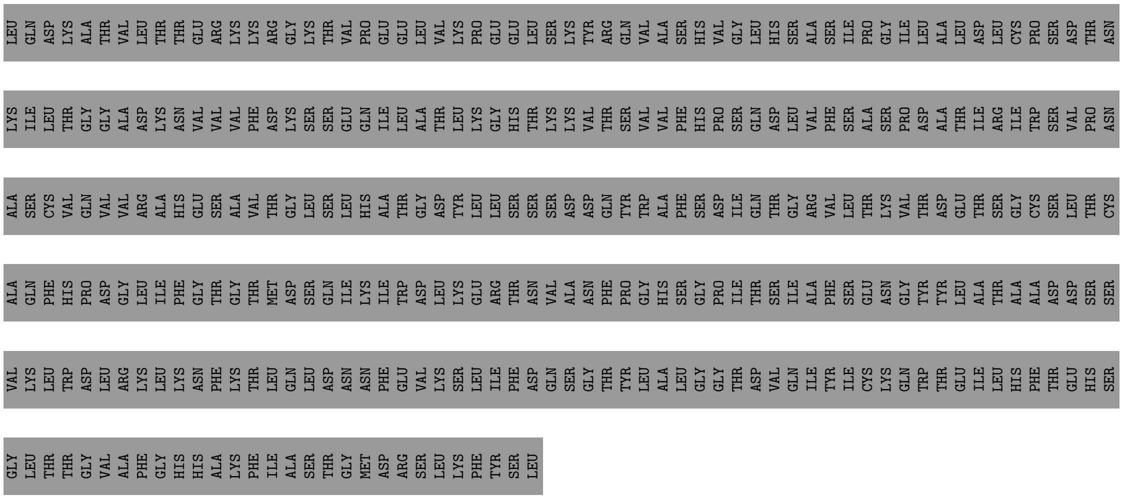


- Molecule 51: Pre-mRNA-processing factor 19

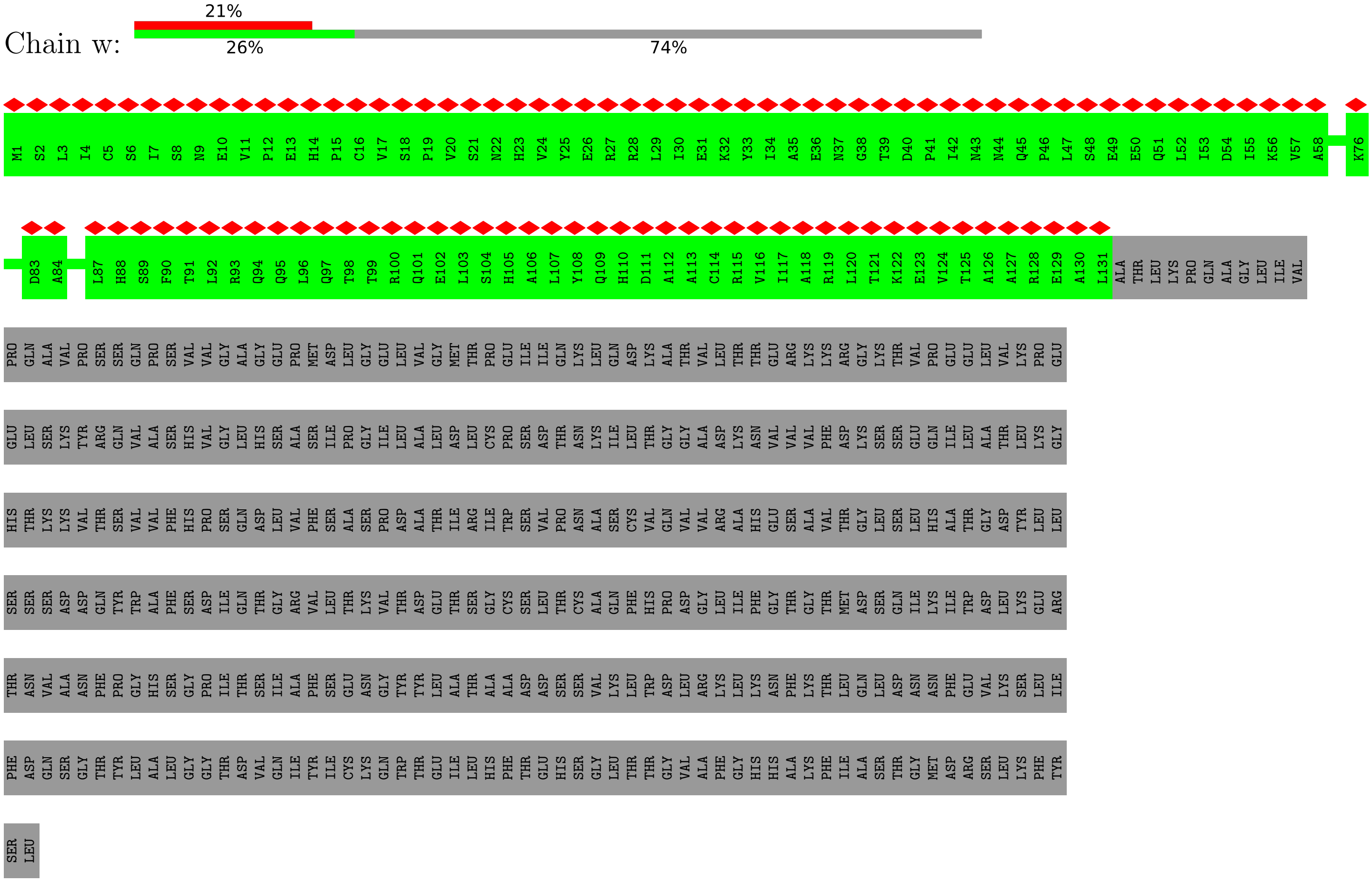


- Molecule 51: Pre-mRNA-processing factor 19

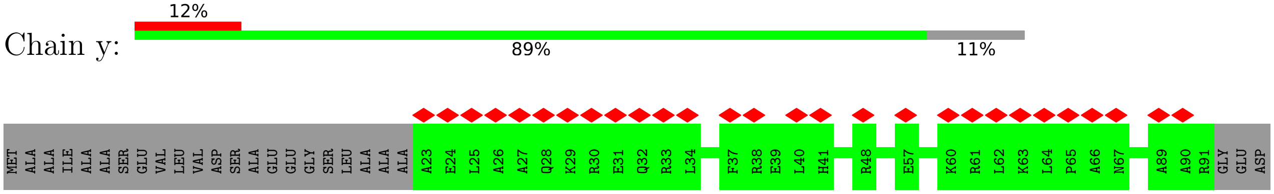


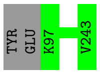


- Molecule 51: Pre-mRNA-processing factor 19



- Molecule 52: Pre-mRNA-splicing factor SYF2

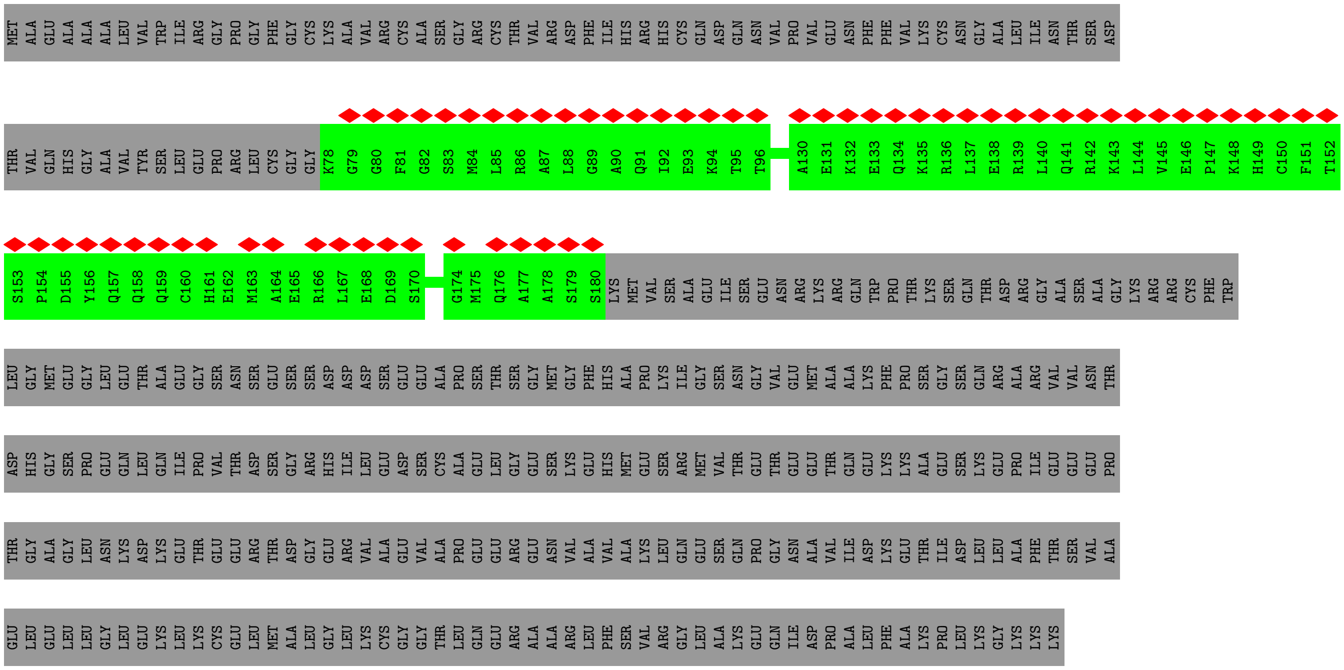




TYR GLU K97 V243

- Molecule 53: Splicing regulator SDE2

Chain z:



MET ALA GLU ALA ALA ALA LEU VAL TRP ILE ARG GLY PRO GLY PHE GLY CYS LYS ALA VAL ARG CYS ALA SER GLY ARG CYS THR VAL ARG ASP PHE ILE HIS ARG HIS CYS GLN ASP GLN ASN VAL PRO VAL GLU ASN PHE PHE VAL LYS CYS ASN GLY ALA LEU ILE ASN THR SER ASP

THR VAL GLN HIS GLY ALA VAL TYR SER LEU GLU PRO ARG LEU CYS GLY GLY K78 G79 G80 F81 G82 S83 M84 L85 R86 A87 L88 G89 A90 Q91 I92 E93 K94 T95 T96 A130 E131 K132 E133 Q134 K135 R136 L137 E138 R139 L140 Q141 R142 K143 L144 V145 E146 P147 K148 H149 C150 F151 T152

S153 P154 D155 Y156 Q157 Q158 Q159 C160 H161 E162 M163 A164 E165 R166 L167 E168 D169 S170 G174 M175 Q176 A177 A178 S179 S180 LYS MET VAL SER ALA GLU ILE SER GLU ASN ARG LYS ARG GLN TRP PRO THR LYS SER GLN THR ASP ARG GLY ALA SER ALA GLY LYS ARG ARG CYS PHE TRP

LEU GLY MET GLU GLY LEU GLU THR ALA GLU GLY SER ASN SER GLU SER SER ASP ASP SER GLU GLU ALA PRO SER THR SER GLY MET GLY PHE HIS ALA PRO LYS ILE GLY SER ASN GLY VAL GLU MET ALA ALA LYS PHE PRO SER GLY SER GLN ARG ALA ARG VAL VAL ASN THR

ASP HIS GLY SER PRO GLU GLN LEU GLN ILE PRO VAL THR ASP SER GLY ARG HIS ILE LEU GLU ASP SER CYS ALA GLU LEU GLY GLU SER LYS GLU HIS MET GLU SER ARG MET VAL THR GLU THR GLU GLU THR GLN GLU LYS LYS ALA GLU SER LYS GLU PRO ILE GLU GLU PRO

THR GLY ALA GLY LEU ASN LYS ASP LYS GLU THR GLU GLU ARG THR ASP GLY GLU ARG VAL ALA GLU VAL ALA PRO GLU GLU ARG GLU ASN VAL ALA VAL ALA LYS LEU GLN GLU SER GLN PRO GLY ASN ALA VAL ILE ASP LYS GLU THR ILE ASP LEU LEU ALA PHE THR SER VAL ALA

GLU LEU GLU LEU LEU GLY LEU GLU LYS LEU LYS CYS GLU LEU MET ALA LEU GLY LEU LYS CYS GLY GLY THR LEU GLN GLU ARG ALA ALA ARG LEU PHE SER VAL ARG GLY LEU ALA LYS GLU GLN ILE ASP PRO ALA LEU PHE ALA LYS PRO LEU LYS GLY LYS LYS LYS

4 Experimental information ⓘ

Property	**Value**	**Source**
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	1150057	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	44.24	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	1.260	Depositor
Minimum map value	-0.359	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.026	Depositor
Recommended contour level	0.115	Depositor
Map size (Å)	609.0, 609.0, 609.0	wwPDB
Map dimensions	580, 580, 580	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, IHP, ZN, K, SEP, ATP, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z|>5$	RMSZ	$\#|Z|>5$
1	2	0.21	0/3359	0.77	0/5224
2	4	0.24	0/643	0.48	0/862
3	5	0.28	0/2643	0.82	0/4110
4	6	0.28	0/2323	0.74	0/3619
5	7	0.25	0/3179	0.50	0/4291
6	8	0.25	0/748	0.50	0/1012
7	9	0.22	0/600	0.45	0/759
8	A	0.27	0/18188	0.47	0/24579
9	B	0.22	0/7219	0.43	0/9180
10	C	0.27	0/7275	0.49	0/9884
11	C3	0.23	0/675	0.50	0/856
12	CD	0.23	0/1039	0.43	0/1327
13	CE	0.24	0/323	0.48	0/406
14	CF	0.21	0/745	0.39	0/930
15	CI	0.21	0/282	0.31	0/351
16	CN	0.27	0/1674	0.50	0/2247
17	CT	0.21	0/414	0.39	0/518
18	D	0.25	0/1030	0.49	0/1371
19	E	0.25	0/571	0.80	0/889
20	F	0.27	0/1133	0.52	0/1530
21	G	0.24	0/503	0.48	0/672
22	H	0.25	0/3378	0.44	0/4509
23	I	0.22	0/1915	0.82	0/2967
24	J	0.28	0/2791	0.52	0/3788
25	K	0.28	0/2387	0.52	0/3205
26	L	0.26	0/1214	0.50	0/1627
27	M	0.26	0/2375	0.51	0/3204
28	N	0.25	0/2442	0.50	0/3309
29	O	0.23	0/3365	0.45	0/4411
30	P	0.26	0/988	0.53	0/1315
31	R	0.25	0/403	0.43	0/526
32	S	0.24	0/3907	0.44	0/5142

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
33	T	0.24	0/4328	0.42	0/5947
34	U	0.22	0/5628	0.43	0/7170
35	V	0.23	0/4217	0.42	0/5405
36	W	0.22	0/727	0.43	0/919
37	Y	0.22	0/379	0.43	0/478
38	Z	0.25	0/456	0.45	0/586
39	a	0.23	0/524	0.43	0/732
39	m	0.23	0/398	0.46	0/503
40	b	0.24	0/416	0.44	0/581
40	k	0.23	0/343	0.50	0/435
41	c	0.23	0/404	0.44	0/564
41	l	0.22	0/331	0.49	0/418
42	d	0.24	0/422	0.44	0/588
42	n	0.23	0/343	0.48	0/433
43	e	0.23	0/393	0.44	0/547
43	p	0.22	0/327	0.47	0/409
44	f	0.24	0/362	0.45	0/502
44	q	0.23	0/303	0.48	0/384
45	g	0.23	0/367	0.46	0/509
45	r	0.23	0/304	0.51	0/383
46	h	0.27	0/2334	0.46	0/3138
47	i	0.26	0/1304	0.48	0/1767
48	j	0.23	0/403	0.48	0/514
49	o	0.26	0/4024	0.49	0/5426
50	s	0.23	0/928	0.41	0/1171
51	t	0.23	0/591	0.40	0/759
51	u	0.23	0/559	0.41	0/717
51	v	0.22	0/595	0.37	0/764
51	w	0.22	0/563	0.40	0/722
52	y	0.25	0/1444	0.45	0/1904
53	z	0.23	0/611	0.47	0/797
All	All	0.25	0/114059	0.51	0/153792

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	4	73/476 (15%)	71 (97%)	2 (3%)	0	100	100
5	7	388/411 (94%)	380 (98%)	8 (2%)	0	100	100
6	8	89/174 (51%)	86 (97%)	3 (3%)	0	100	100
7	9	142/146 (97%)	141 (99%)	1 (1%)	0	100	100
8	A	2257/2335 (97%)	2206 (98%)	51 (2%)	0	100	100
9	B	1724/2136 (81%)	1702 (99%)	22 (1%)	0	100	100
10	C	897/972 (92%)	842 (94%)	55 (6%)	0	100	100
11	C3	160/161 (99%)	159 (99%)	1 (1%)	0	100	100
12	CD	243/622 (39%)	241 (99%)	2 (1%)	0	100	100
13	CE	77/301 (26%)	77 (100%)	0	0	100	100
14	CF	179/339 (53%)	177 (99%)	1 (1%)	1 (1%)	25	56
15	CI	66/285 (23%)	66 (100%)	0	0	100	100
16	CN	248/301 (82%)	230 (93%)	16 (6%)	2 (1%)	19	49
17	CT	98/289 (34%)	96 (98%)	2 (2%)	0	100	100
18	D	121/184 (66%)	119 (98%)	2 (2%)	0	100	100
20	F	120/758 (16%)	112 (93%)	8 (7%)	0	100	100
21	G	57/112 (51%)	55 (96%)	2 (4%)	0	100	100
22	H	455/500 (91%)	443 (97%)	12 (3%)	0	100	100
24	J	363/514 (71%)	348 (96%)	15 (4%)	0	100	100
25	K	291/536 (54%)	272 (94%)	18 (6%)	1 (0%)	41	72
26	L	142/144 (99%)	138 (97%)	4 (3%)	0	100	100
27	M	287/420 (68%)	280 (98%)	7 (2%)	0	100	100
28	N	304/357 (85%)	286 (94%)	18 (6%)	0	100	100
29	O	542/802 (68%)	533 (98%)	9 (2%)	0	100	100
30	P	113/229 (49%)	109 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
31	R	72/2752 (3%)	70 (97%)	2 (3%)	0	100	100
32	S	631/687 (92%)	615 (98%)	16 (2%)	0	100	100
33	T	677/855 (79%)	655 (97%)	22 (3%)	0	100	100
34	U	1328/1485 (89%)	1320 (99%)	8 (1%)	0	100	100
35	V	924/1220 (76%)	893 (97%)	27 (3%)	4 (0%)	34	66
36	W	174/255 (68%)	170 (98%)	4 (2%)	0	100	100
37	Y	90/225 (40%)	90 (100%)	0	0	100	100
38	Z	85/415 (20%)	83 (98%)	2 (2%)	0	100	100
39	a	101/118 (86%)	100 (99%)	1 (1%)	0	100	100
39	m	91/118 (77%)	89 (98%)	2 (2%)	0	100	100
40	b	80/240 (33%)	79 (99%)	1 (1%)	0	100	100
40	k	80/240 (33%)	77 (96%)	3 (4%)	0	100	100
41	c	78/119 (66%)	77 (99%)	1 (1%)	0	100	100
41	l	78/119 (66%)	78 (100%)	0	0	100	100
42	d	82/126 (65%)	80 (98%)	2 (2%)	0	100	100
42	n	81/126 (64%)	79 (98%)	2 (2%)	0	100	100
43	e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
43	p	79/92 (86%)	79 (100%)	0	0	100	100
44	f	70/86 (81%)	70 (100%)	0	0	100	100
44	q	70/86 (81%)	69 (99%)	1 (1%)	0	100	100
45	g	71/76 (93%)	70 (99%)	1 (1%)	0	100	100
45	r	71/76 (93%)	70 (99%)	1 (1%)	0	100	100
46	h	272/586 (46%)	263 (97%)	9 (3%)	0	100	100
47	i	162/166 (98%)	155 (96%)	7 (4%)	0	100	100
48	j	93/222 (42%)	93 (100%)	0	0	100	100
49	o	511/579 (88%)	476 (93%)	35 (7%)	0	100	100
50	s	223/225 (99%)	220 (99%)	3 (1%)	0	100	100
51	t	135/504 (27%)	133 (98%)	2 (2%)	0	100	100
51	u	128/504 (25%)	123 (96%)	5 (4%)	0	100	100
51	v	136/504 (27%)	136 (100%)	0	0	100	100
51	w	129/504 (26%)	127 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
52	y	212/243 (87%)	209 (99%)	3 (1%)	0	100	100
53	z	101/451 (22%)	98 (97%)	3 (3%)	0	100	100
All	All	16628/27600 (60%)	16191 (97%)	429 (3%)	8 (0%)	100	100

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
14	CF	201	ILE
16	CN	123	SER
35	V	987	PRO
35	V	1049	TYR
16	CN	141	ASP

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	4	70/395 (18%)	70 (100%)	0	100	100
5	7	345/361 (96%)	344 (100%)	1 (0%)	92	98
6	8	76/143 (53%)	76 (100%)	0	100	100
7	9	6/134 (4%)	6 (100%)	0	100	100
8	A	1825/2108 (87%)	1825 (100%)	0	100	100
9	B	79/1908 (4%)	79 (100%)	0	100	100
10	C	799/866 (92%)	797 (100%)	2 (0%)	92	98
11	C3	7/141 (5%)	7 (100%)	0	100	100
12	CD	15/533 (3%)	15 (100%)	0	100	100
13	CE	2/252 (1%)	2 (100%)	0	100	100
14	CF	2/304 (1%)	2 (100%)	0	100	100
15	CI	1/240 (0%)	1 (100%)	0	100	100
16	CN	155/260 (60%)	155 (100%)	0	100	100
17	CT	2/263 (1%)	2 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
18	D	106/157 (68%)	106 (100%)	0	100	100
20	F	111/655 (17%)	110 (99%)	1 (1%)	78	94
21	G	53/99 (54%)	53 (100%)	0	100	100
22	H	324/455 (71%)	324 (100%)	0	100	100
24	J	279/441 (63%)	279 (100%)	0	100	100
25	K	246/457 (54%)	245 (100%)	1 (0%)	91	97
26	L	130/130 (100%)	130 (100%)	0	100	100
27	M	258/361 (72%)	257 (100%)	1 (0%)	91	97
28	N	262/300 (87%)	262 (100%)	0	100	100
29	O	232/709 (33%)	232 (100%)	0	100	100
30	P	103/203 (51%)	103 (100%)	0	100	100
31	R	23/2432 (1%)	23 (100%)	0	100	100
32	S	252/617 (41%)	251 (100%)	1 (0%)	91	97
33	T	225/749 (30%)	225 (100%)	0	100	100
34	U	72/1336 (5%)	72 (100%)	0	100	100
35	V	104/1085 (10%)	104 (100%)	0	100	100
36	W	6/218 (3%)	6 (100%)	0	100	100
37	Y	3/195 (2%)	3 (100%)	0	100	100
38	Z	24/366 (7%)	24 (100%)	0	100	100
39	a	5/110 (4%)	5 (100%)	0	100	100
39	m	5/110 (4%)	5 (100%)	0	100	100
40	b	4/177 (2%)	4 (100%)	0	100	100
40	k	4/177 (2%)	4 (100%)	0	100	100
41	c	3/101 (3%)	3 (100%)	0	100	100
41	l	3/101 (3%)	3 (100%)	0	100	100
42	d	3/101 (3%)	3 (100%)	0	100	100
42	n	3/101 (3%)	3 (100%)	0	100	100
43	e	1/84 (1%)	1 (100%)	0	100	100
43	p	1/84 (1%)	1 (100%)	0	100	100
44	f	4/74 (5%)	4 (100%)	0	100	100
44	q	4/74 (5%)	4 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
45	g	3/66 (4%)	3 (100%)	0	100	100
45	r	3/66 (4%)	3 (100%)	0	100	100
46	h	243/520 (47%)	243 (100%)	0	100	100
47	i	133/134 (99%)	133 (100%)	0	100	100
48	j	6/198 (3%)	6 (100%)	0	100	100
49	o	400/502 (80%)	400 (100%)	0	100	100
50	s	7/196 (4%)	7 (100%)	0	100	100
51	t	11/435 (2%)	11 (100%)	0	100	100
51	u	10/435 (2%)	10 (100%)	0	100	100
51	v	11/435 (2%)	11 (100%)	0	100	100
51	w	10/435 (2%)	10 (100%)	0	100	100
52	y	114/209 (54%)	114 (100%)	0	100	100
53	z	40/371 (11%)	40 (100%)	0	100	100
All	All	7228/24169 (30%)	7221 (100%)	7 (0%)	93	98

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
20	F	641	ARG
25	K	233	PRO
32	S	281	ARG
27	M	25	GLN
10	C	919	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
52	y	219	ASN
52	y	215	ASN
47	i	120	GLN
28	N	101	ASN
52	y	203	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	139/188 (73%)	51 (36%)	3 (2%)
19	E	23/242 (9%)	11 (47%)	2 (8%)
23	I	82/235 (34%)	47 (57%)	3 (3%)
3	5	112/116 (96%)	39 (34%)	3 (2%)
4	6	96/106 (90%)	27 (28%)	2 (2%)
All	All	452/887 (50%)	175 (38%)	13 (2%)

5 of 175 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	16	U
1	2	18	U
1	2	19	G
1	2	20	G
1	2	24	A

5 of 13 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	6	58	G
19	E	-19	C
23	I	183	C
23	I	150	U
23	I	181	U

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
25	SEP	K	232	25	8,9,10	0.63	0	8,12,14	0.66	0
25	SEP	K	224	25	8,9,10	0.61	0	8,12,14	0.67	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
25	SEP	K	232	25	-	0/5/8/10	-
25	SEP	K	224	25	-	3/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
25	K	224	SEP	CB-OG-P-O2P
25	K	224	SEP	CB-OG-P-O3P
25	K	224	SEP	CB-OG-P-O1P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 18 ligands modelled in this entry, 15 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$\lvert Z\rvert > 2$	Counts	RMSZ	#$\lvert Z\rvert > 2$
56	ATP	7	502	55	26,33,33	0.66	0	31,52,52	0.88	1 (3%)
57	IHP	A	3000	-	36,36,36	0.72	0	54,60,60	0.91	3 (5%)
58	GTP	C	1502	55	26,34,34	1.03	3 (11%)	32,54,54	0.80	1 (3%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
56	ATP	7	502	55	-	4/18/38/38	0/3/3/3
57	IHP	A	3000	-	-	2/30/54/54	0/1/1/1
58	GTP	C	1502	55	-	1/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
58	C	1502	GTP	C5-C6	-2.76	1.41	1.47
58	C	1502	GTP	C8-N7	-2.24	1.31	1.35
58	C	1502	GTP	C5-C4	-2.04	1.37	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
57	A	3000	IHP	C6-C1-C2	2.79	116.51	110.41
57	A	3000	IHP	C5-C4-C3	-2.41	105.14	110.41
56	7	502	ATP	C5-C6-N6	2.32	123.87	120.35
57	A	3000	IHP	C6-C5-C4	-2.30	105.37	110.41
58	C	1502	GTP	O6-C6-C5	2.01	128.29	124.37

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
56	7	502	ATP	C5'-O5'-PA-O1A
56	7	502	ATP	PB-O3A-PA-O5'
57	A	3000	IHP	C4-O14-P4-O24
56	7	502	ATP	O4'-C4'-C5'-O5'
56	7	502	ATP	C3'-C4'-C5'-O5'

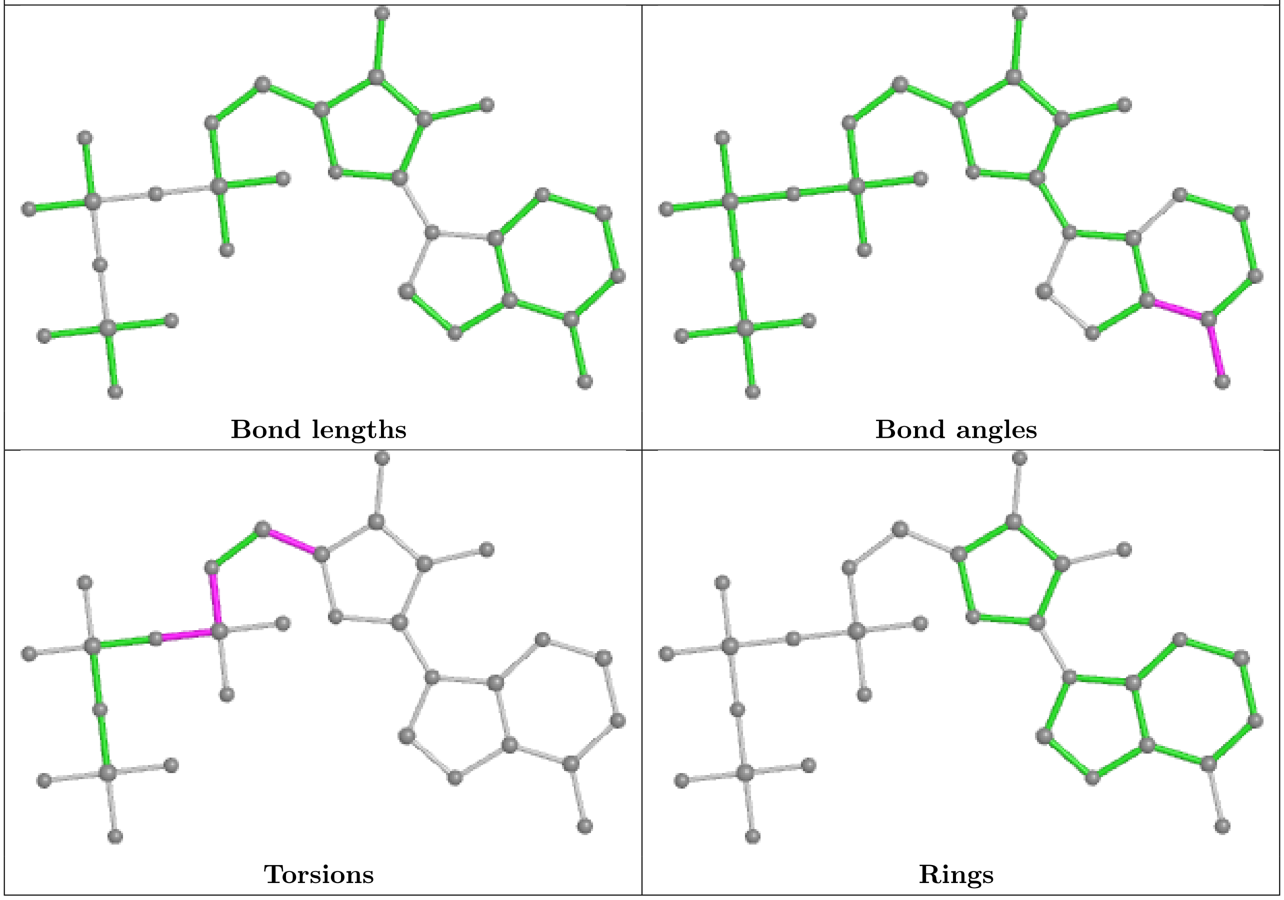
There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

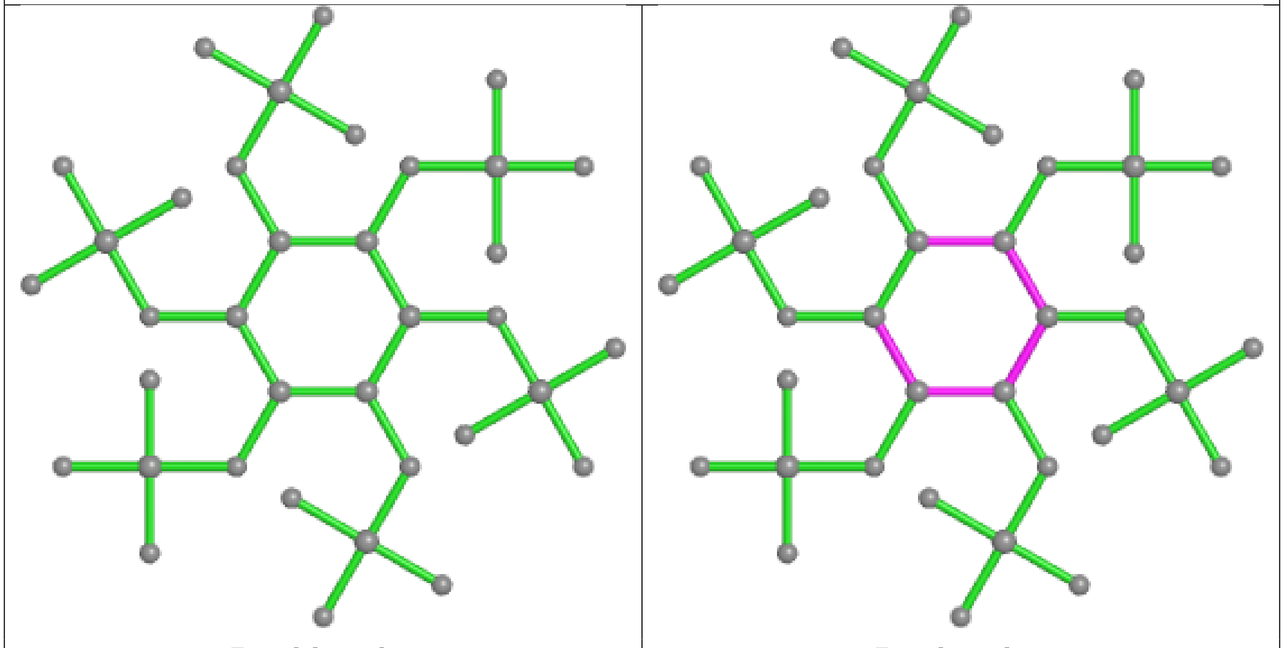
Ligand ATP 7 502



Bond lengths **Bond angles**

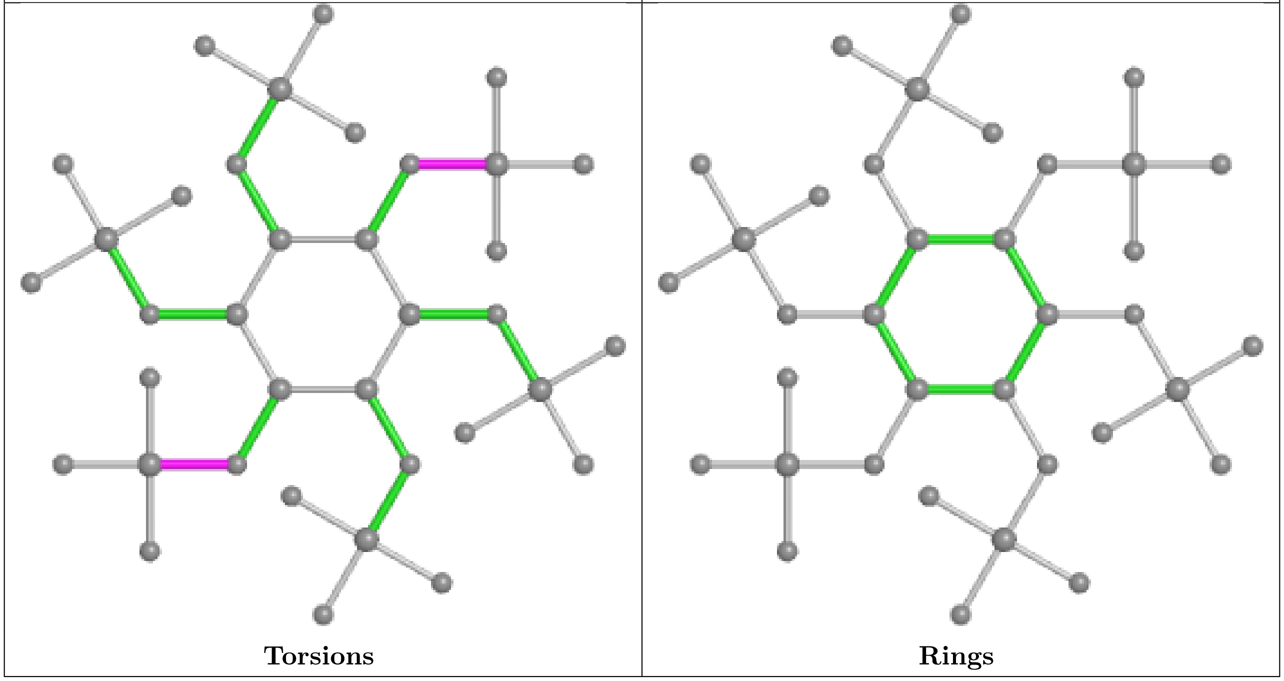
Torsions **Rings**

Ligand IHP A 3000



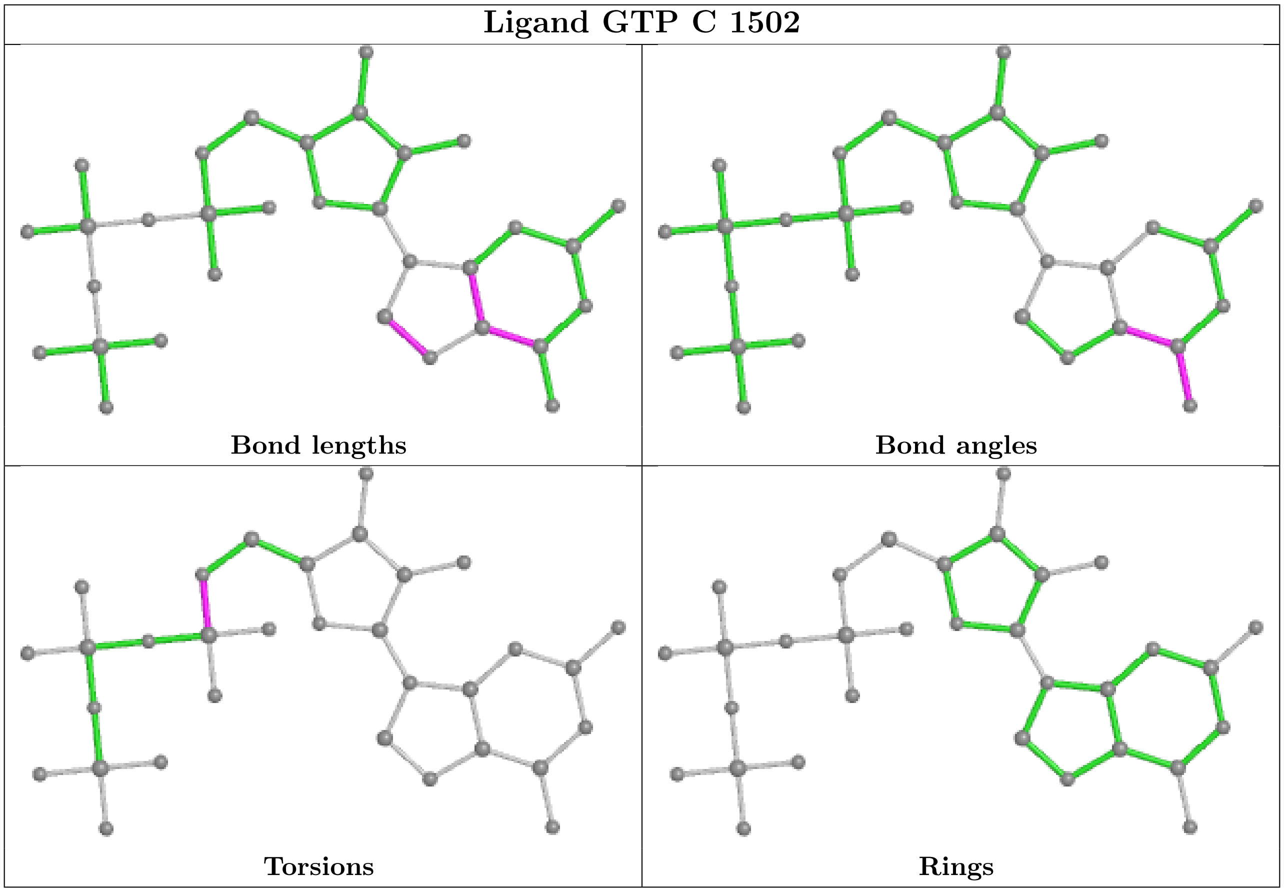
Bond lengths

Bond angles



Torsions

Rings



5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

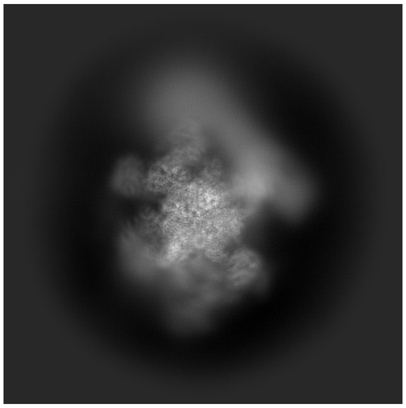
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-16452. These allow visual inspection of the internal detail of the map and identification of artifacts.

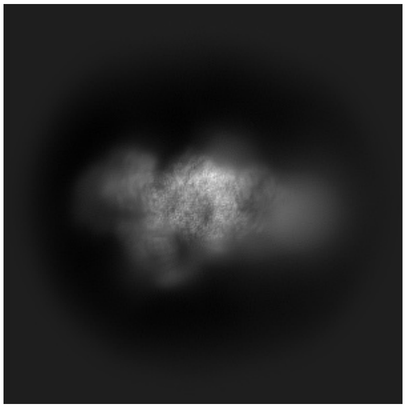
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

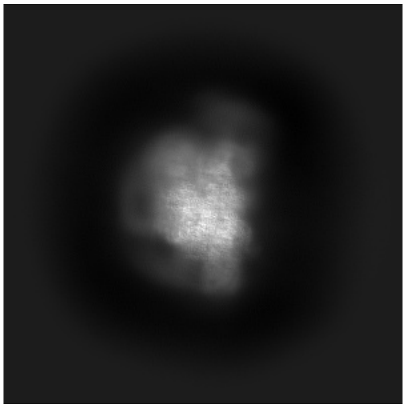
6.1.1 Primary map



X

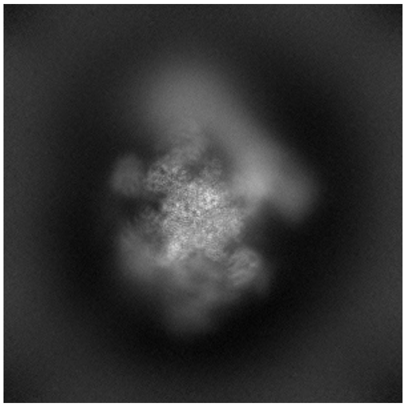


Y

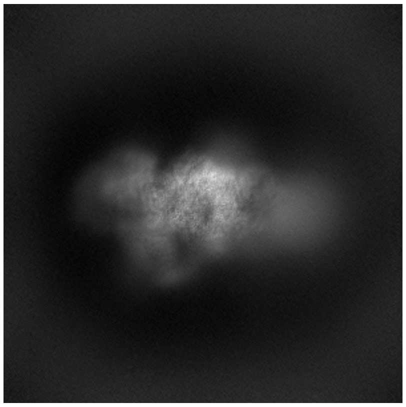


Z

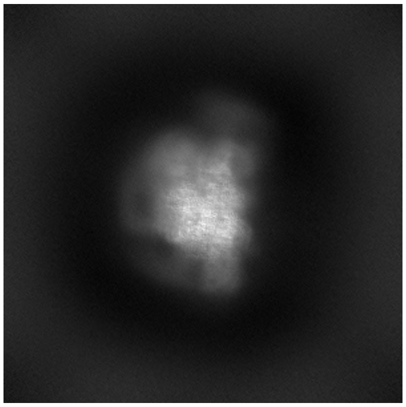
6.1.2 Raw map



X



Y

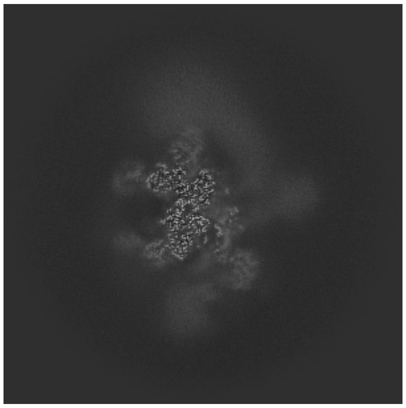


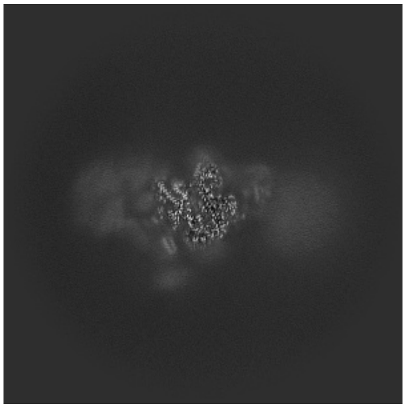
Z

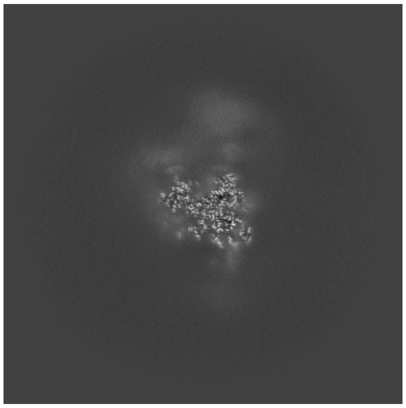
The images above show the map projected in three orthogonal directions.

6.2 Central slices

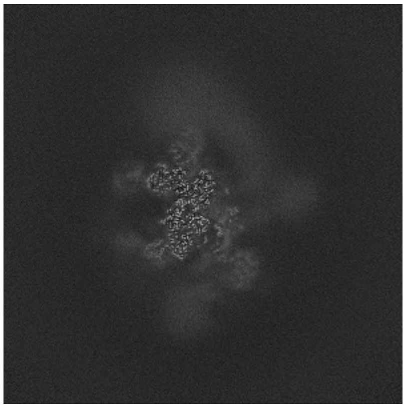
6.2.1 Primary map

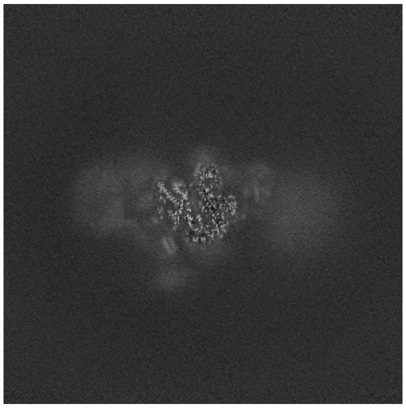

X Index: 290

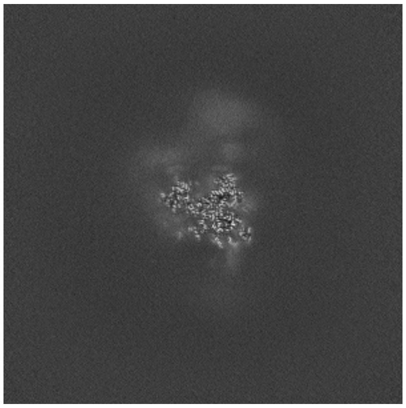

Y Index: 290


Z Index: 290

6.2.2 Raw map


X Index: 290

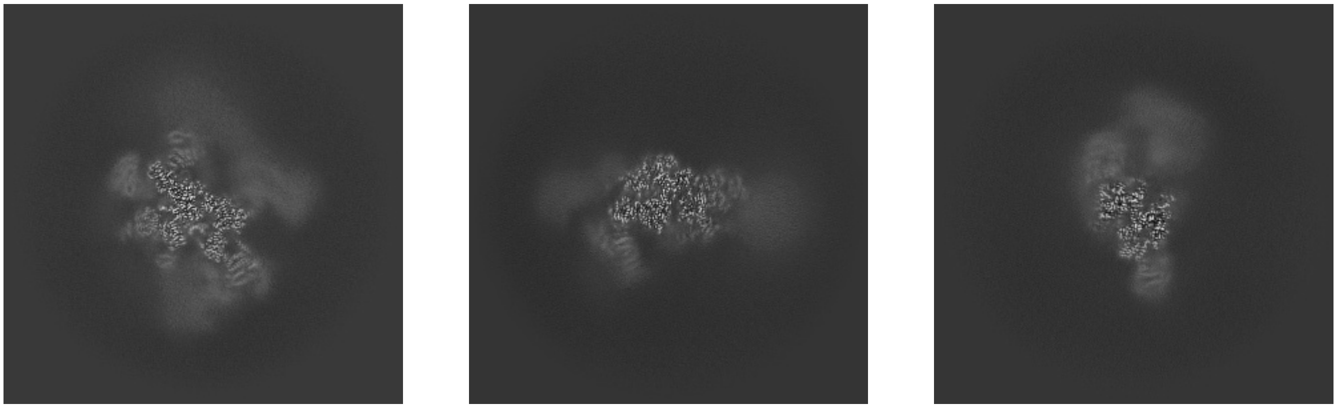

Y Index: 290


Z Index: 290

The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ

6.3.1 Primary map

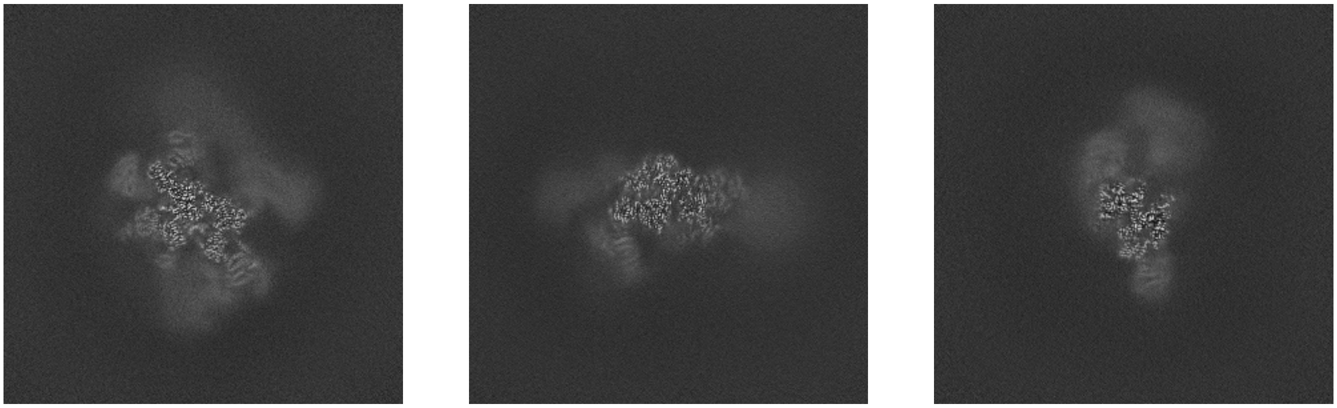


X Index: 321

Y Index: 251

Z Index: 310

6.3.2 Raw map



X Index: 321

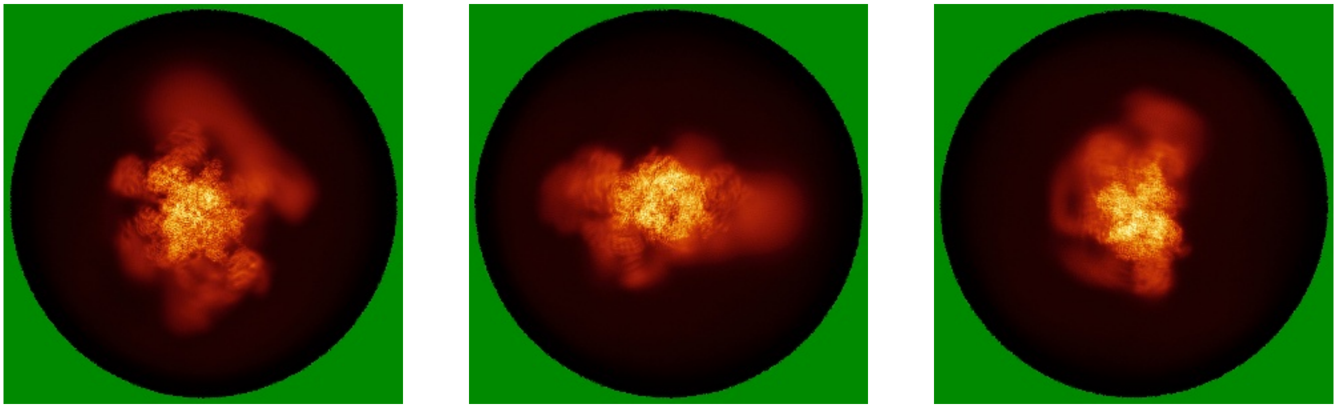
Y Index: 251

Z Index: 310

The images above show the largest variance slices of the map in three orthogonal directions.

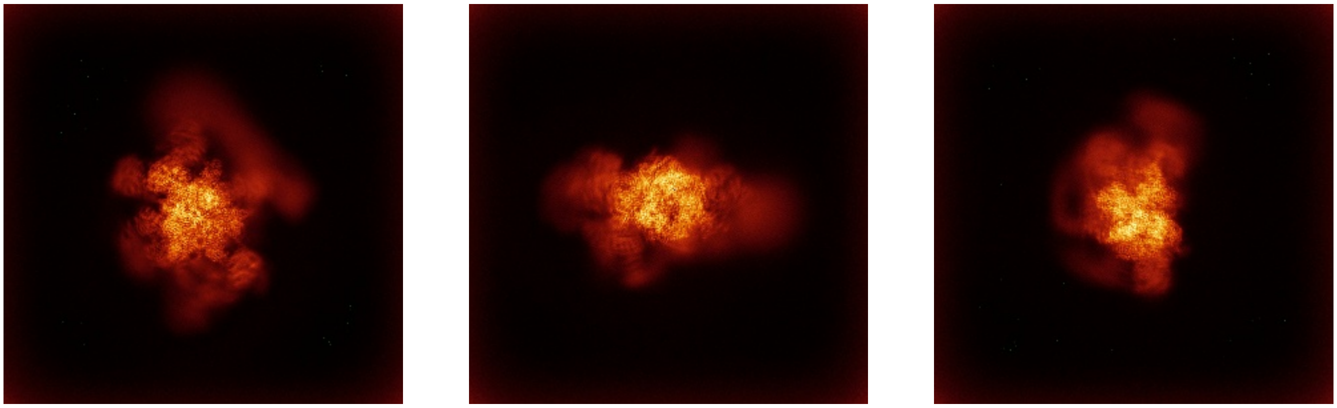
6.4 Orthogonal standard-deviation projections (False-color) ⓘ

6.4.1 Primary map



X Y Z

6.4.2 Raw map



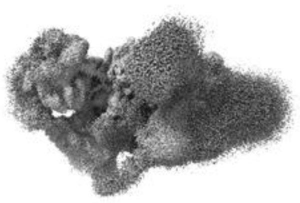
X Y Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

X



Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.115. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X

Y

Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

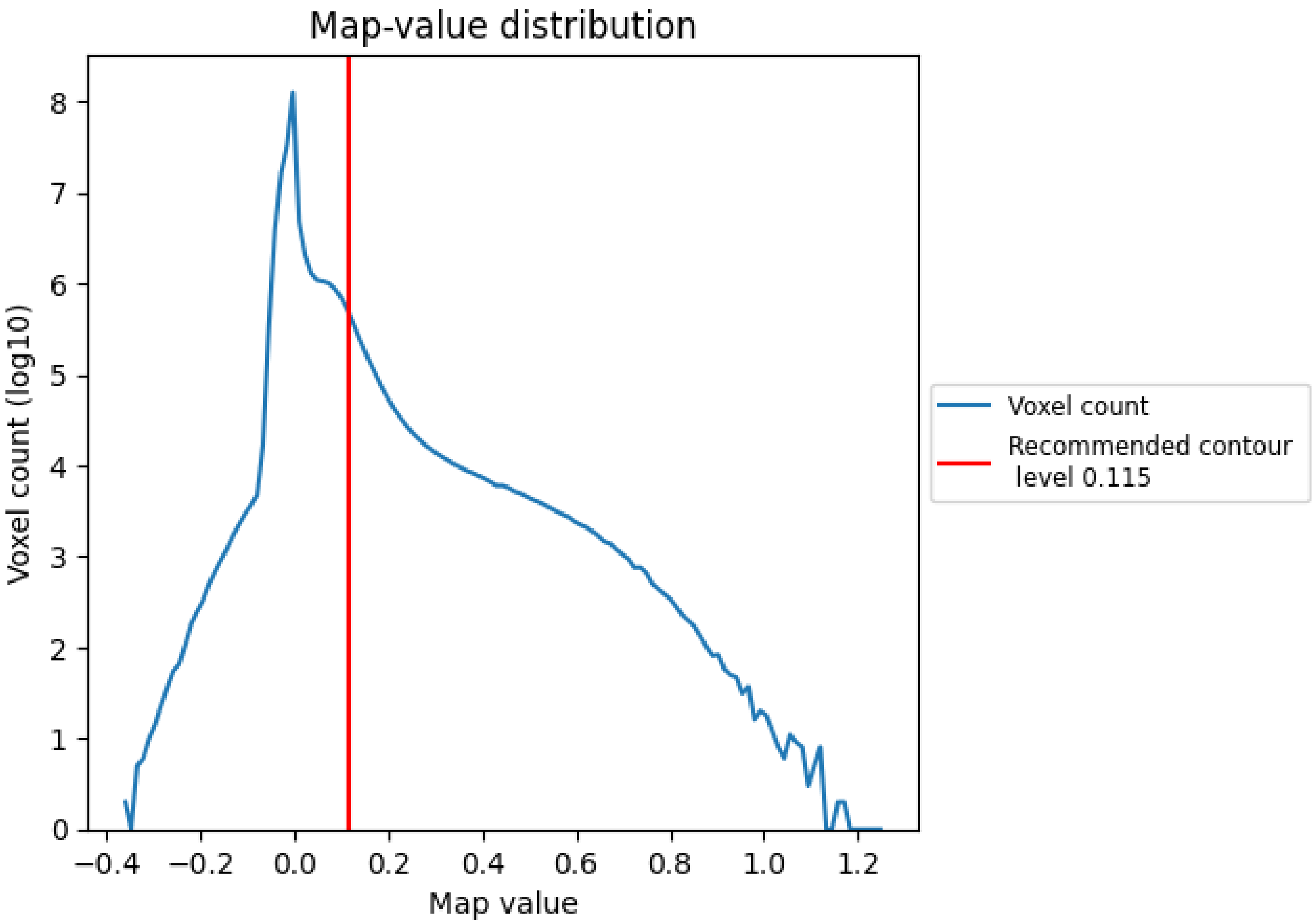
6.6 Mask visualisation ⓘ

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

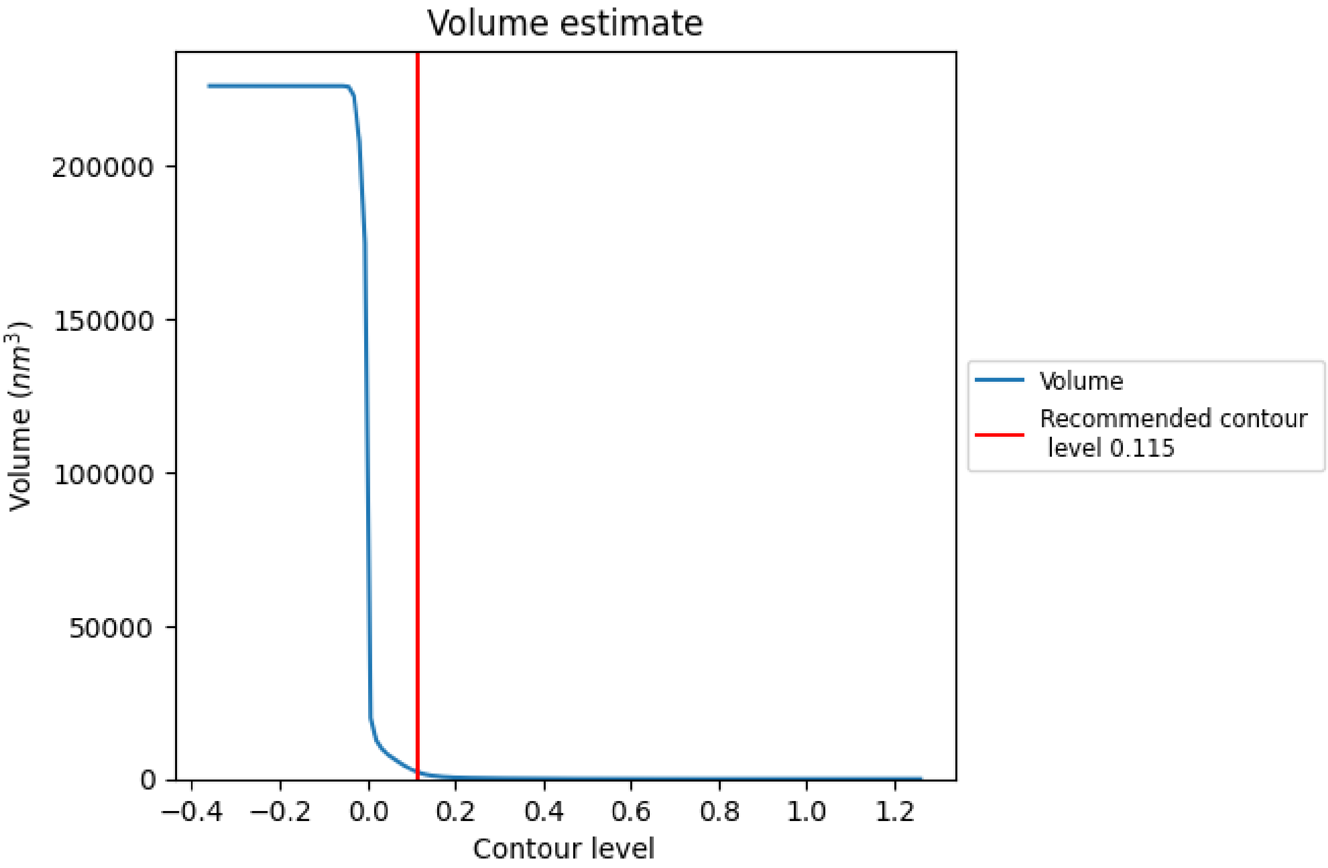
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

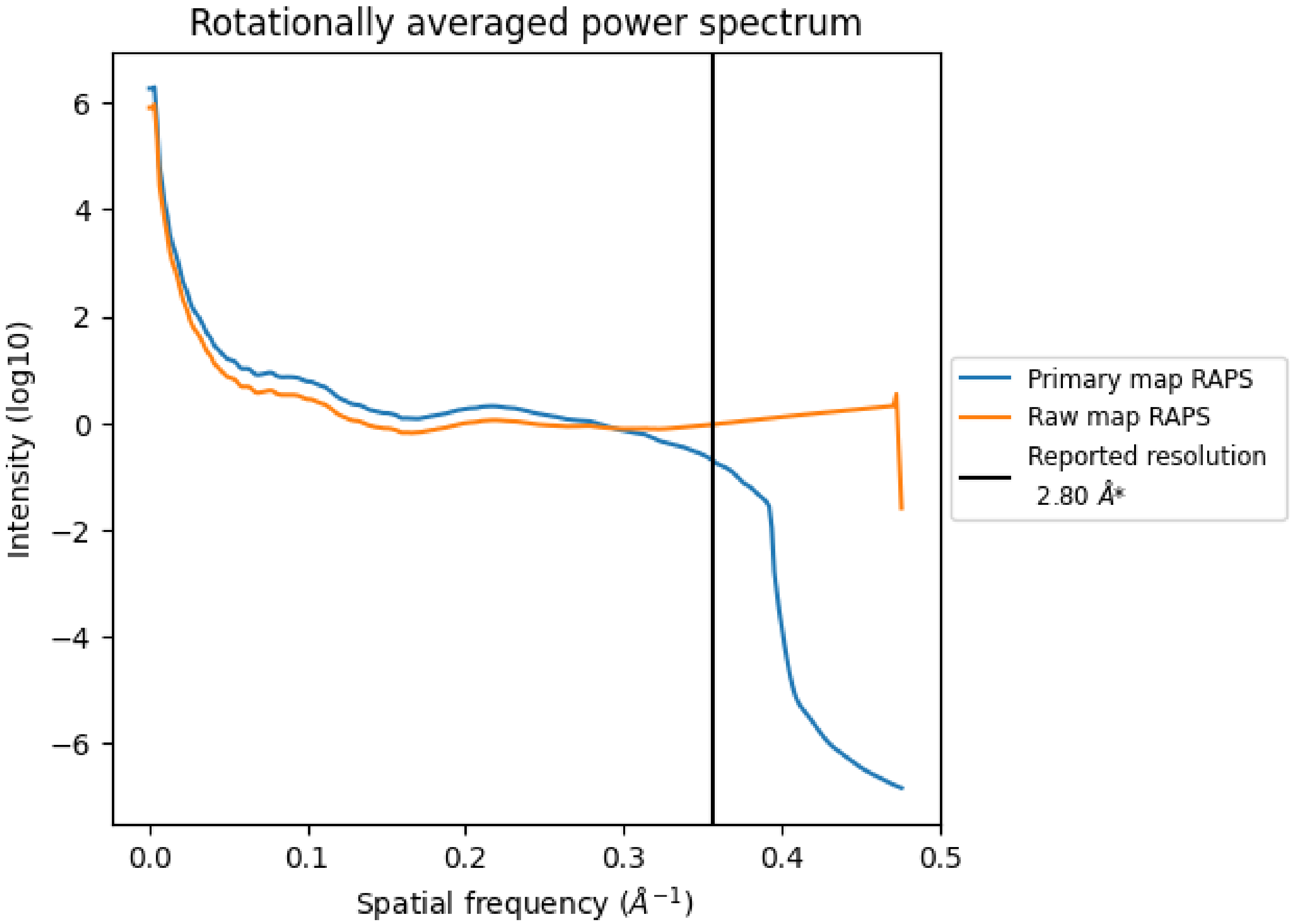
7.2 Volume estimate ⓘ



The volume at the recommended contour level is 2235 nm^3; this corresponds to an approximate mass of 2019 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum ⓘ

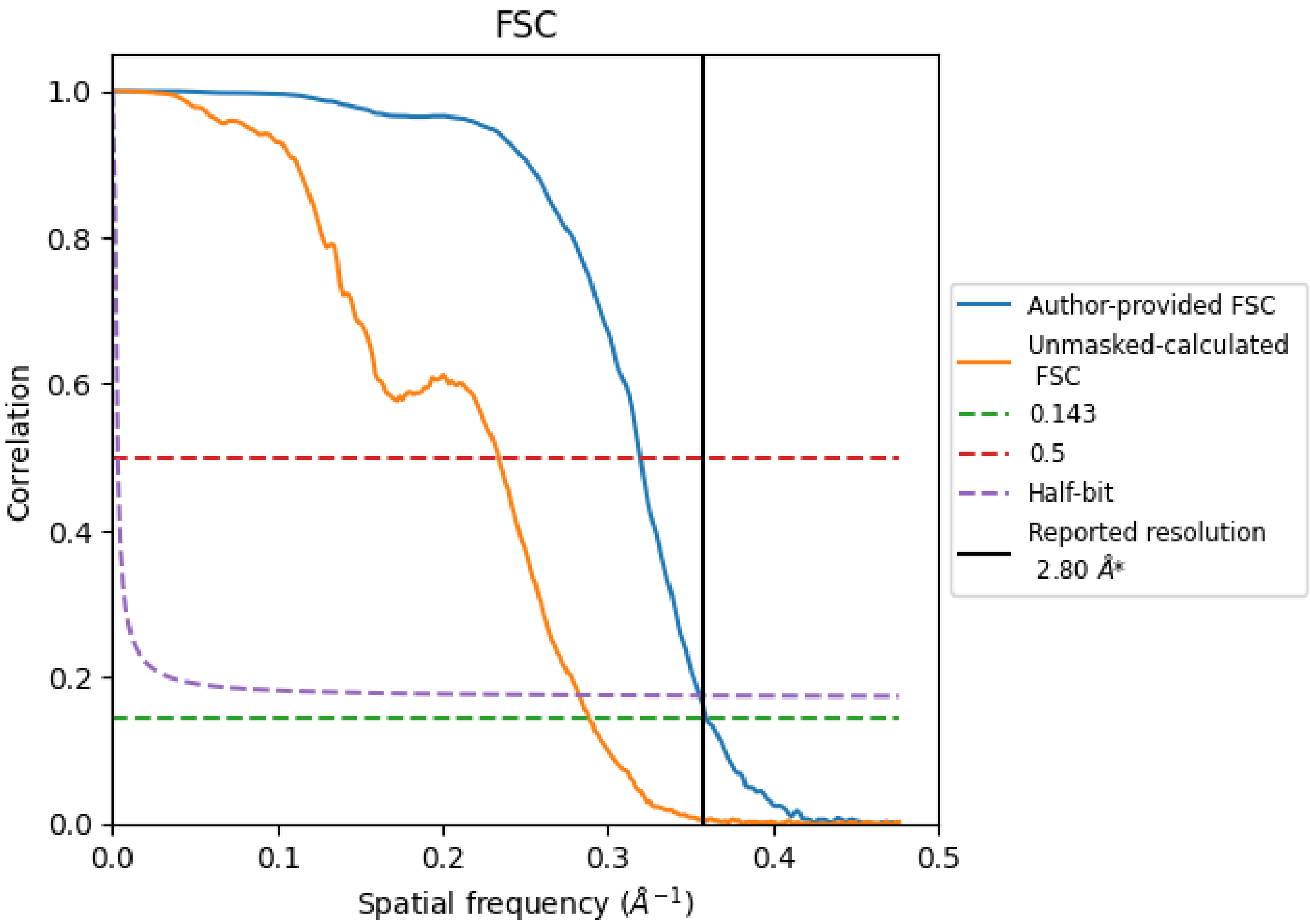


*Reported resolution corresponds to spatial frequency of 0.357 Å^{-1}

8 Fourier-Shell correlation ⓘ

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC ⓘ



*Reported resolution corresponds to spatial frequency of 0.357 Å^{-1}

8.2 Resolution estimates ⓘ

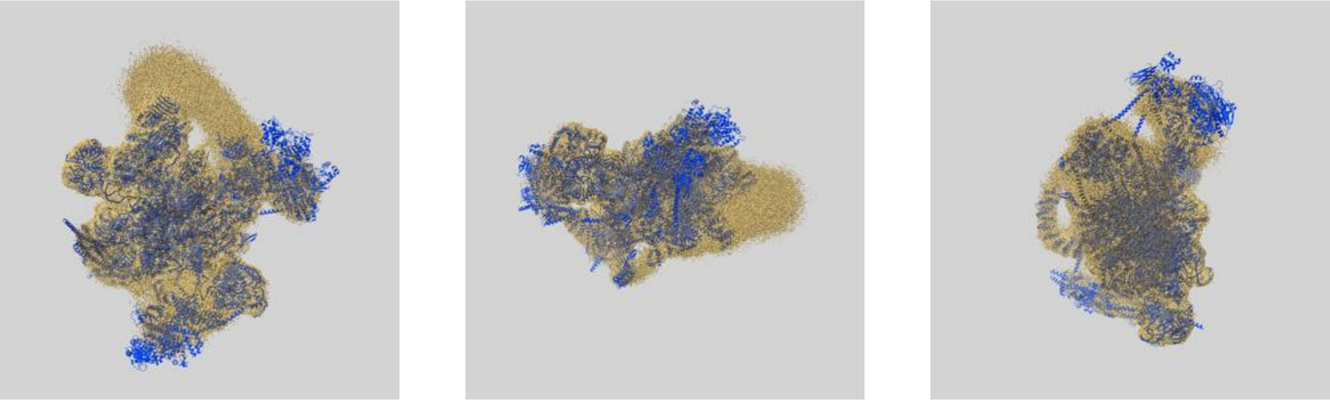
Resolution estimate (Å)	**Estimation criterion (FSC cut-off)**		
	0.143	0.5	Half-bit
Reported by author	2.80	-	-
Author-provided FSC curve	2.78	3.13	2.81
Unmasked-calculated*	3.46	4.28	3.54

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.46 differs from the reported value 2.8 by more than 10 %

9 Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-16452 and PDB model 8C6J. Per-residue inclusion information can be found in section 3 on page 17.

9.1 Map-model overlay ⓘ



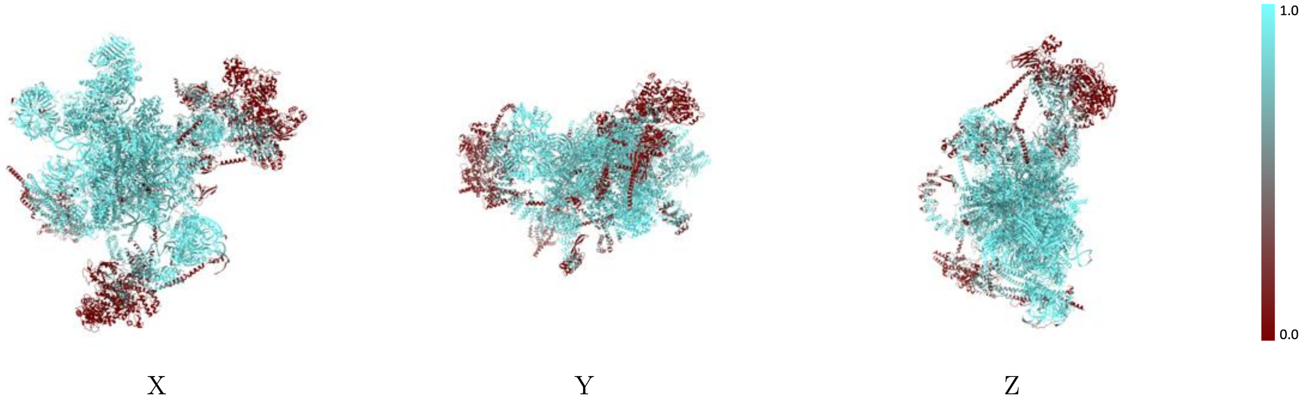
X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.115 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model ⓘ

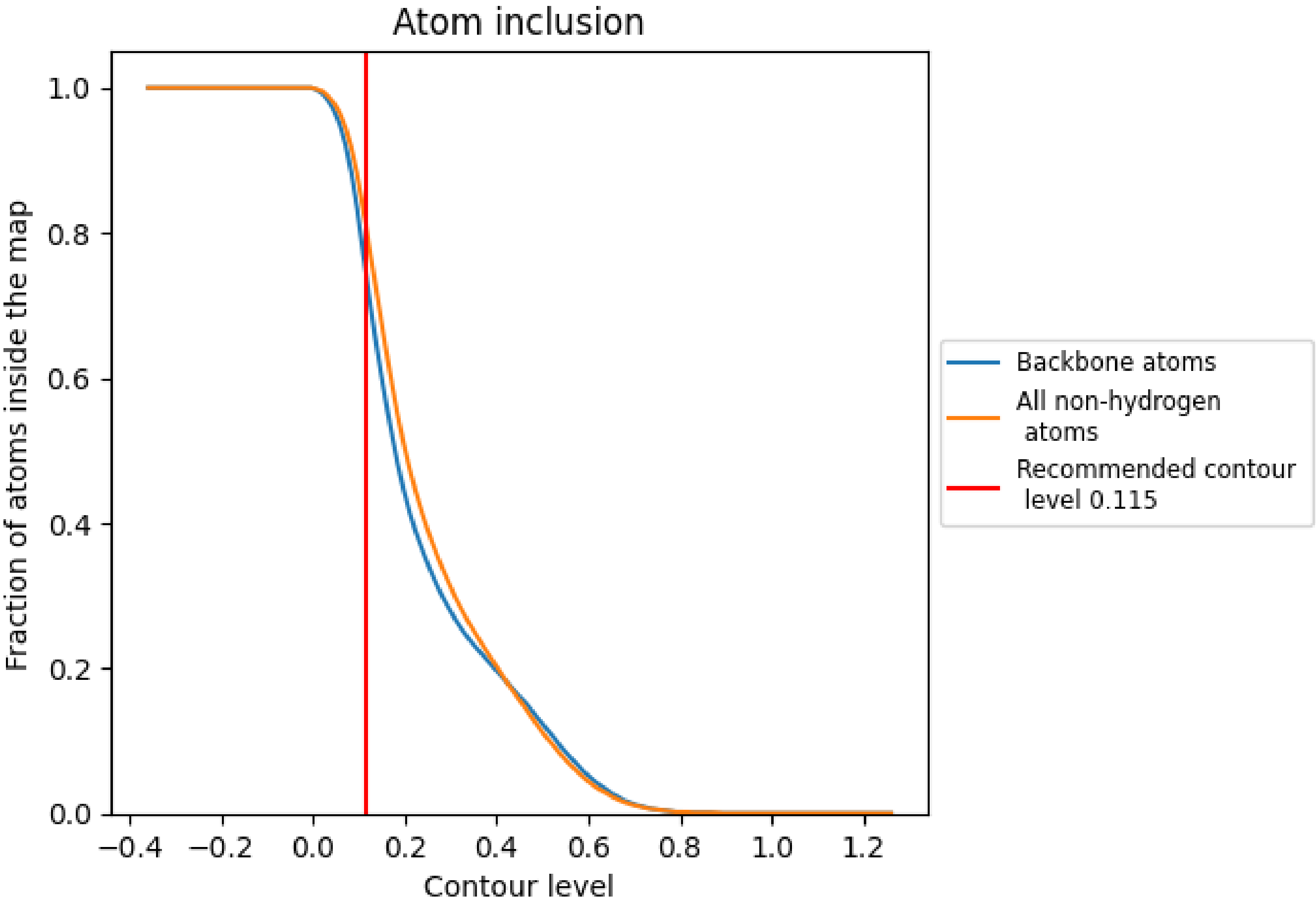
This section was not generated.

9.3 Atom inclusion mapped to coordinate model ⓘ



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.115).

9.4 Atom inclusion ⓘ



At the recommended contour level, 75% of all backbone atoms, 81% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.115) and Q-score for the entire model and for each chain.

Chain	Atom inclusion
All	0.8120
2	0.9200
4	0.7600
5	0.9690
6	0.9820
7	0.9160
8	0.9360
9	0.9880
A	0.9740
B	0.2780
C	0.9810
C3	0.3120
CD	0.5880
CE	0.7300
CF	0.1770
CI	0.3850
CN	0.7090
CT	0.6160
D	0.9090
E	0.9470
F	0.9870
G	0.9750
H	0.9670
I	0.9750
J	0.9410
K	0.9720
L	0.9920
M	0.9720
N	0.9770
O	0.7810
P	0.9770
R	0.9800
S	0.8460
T	0.8510
U	0.0940

Continued on next page...

Continued from previous page...

Chain	Atom inclusion
V	0.8880
W	0.9100
Y	0.9970
Z	0.6340
a	0.8040
b	0.8810
c	0.9730
d	0.9930
e	0.9360
f	0.8720
g	0.9530
h	0.9650
i	0.9810
j	0.1960
k	0.9680
l	0.9760
m	0.8960
n	0.9970
o	0.9640
p	0.9850
q	0.9870
r	0.9870
s	0.4790
t	0.3100
u	0.2360
v	0.4390
w	0.1990
y	0.8900
z	0.4650