



## wwPDB EM Validation Summary Report ⓘ

Nov 20, 2022 – 08:53 PM EST

PDB ID : 7MQA  
EMDB ID : EMD-23938  
Title : Cryo-EM structure of the human SSU processome, state post-A1  
Authors : Vanden Broeck, A.; Singh, S.; Klinge, S.  
Deposited on : 2021-05-05  
Resolution : 2.70 Å(reported)  
Based on initial models : 6ZQD, 6ZOJ, 5FAI, 6G4S, 4JXM, 6O16, 2IPX, 5WLC, 6G18, 2OZB, 6W6C

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

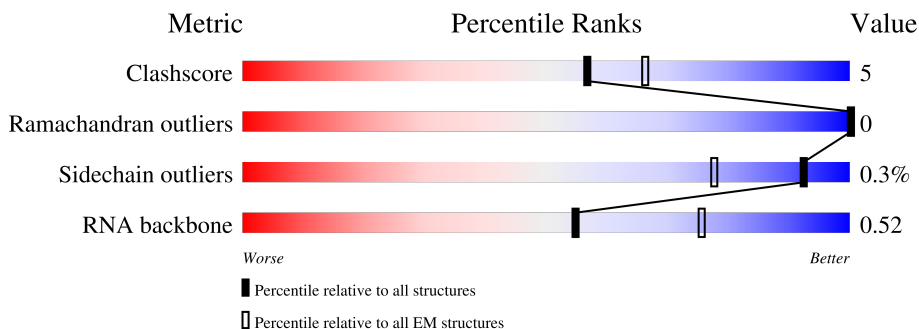
EMDB validation analysis : 0.0.1.dev43  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 2.70 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)
Clashscore	158937	4297
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  $\geq 3$ , 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  $\leq 5\%$ . 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	L0	3617	100%
2	L1	1872	6% (red), 50% (green), 30% (yellow), 6% (orange), 13% (grey)
3	L2	217	6% (red), 48% (green), 29% (yellow), 18% (grey)
4	L3	116	25% (red), 92% (green), 6% (orange), 2% (grey)
5	L4	263	97% (green), 2% (orange), 1% (grey)
6	L5	204	81% (green), 11% (yellow), 7% (grey)
7	L6	249	6% (red), 82% (green), 10% (yellow), 8% (grey)




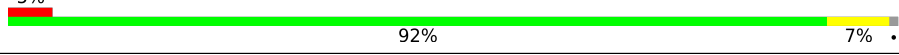



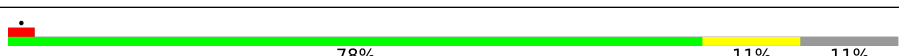
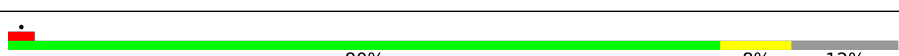

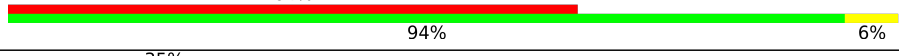


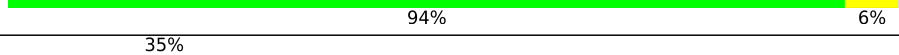



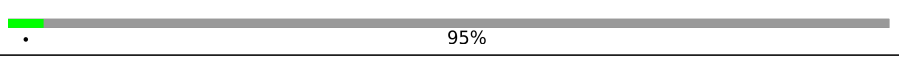


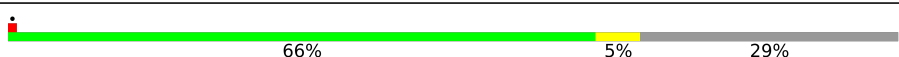

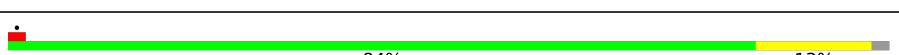
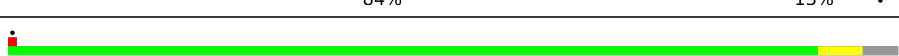
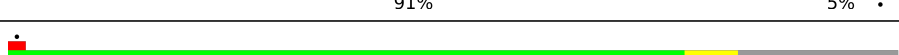
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	L7	194	6% 72% 14% 13%
9	L8	208	78% 9% 13%
10	L9	194	84% 9% 7%
11	LA	132	46% 44% 9%
12	LC	146	84% 9% 7%
13	LD	158	8% 87% 8%
14	LF	133	86% 9% 5%
15	LG	69	78% 12% 10%
16	LH	830	78% 12% 10%
17	LI	678	16% 68% 11% 21%
18	LJ	518	75% 14% 11%
19	LK	677	13% 83%
19	LL	677	5% 62% 13% 25%
20	LM	2144	6% 81% 14% 5%
21	LN	686	85% 13%
22	LO	919	73% 16% 11%
23	LP	597	5% 86% 9% 5%
24	LQ	943	72% 16% 12%
25	LR	808	75% 21%
26	LS	556	71% 11% 19%
27	LT	951	76% 13% 10%
28	LU	445	87% 13%
29	LW	610	65% 10% 26%
30	LY	136	28% 58% 12% 30%
31	LZ	184	79% 9% 12%



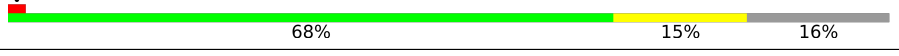



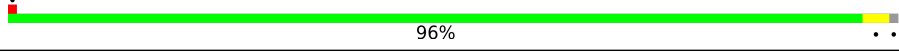
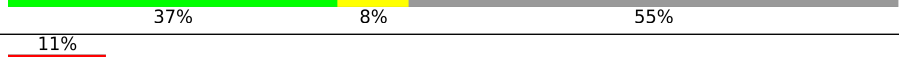

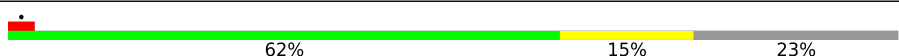
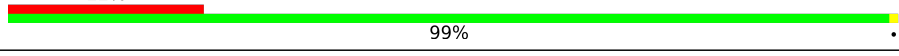
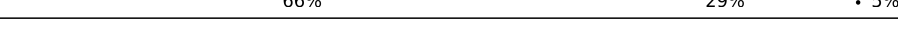


Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
32	NA	681	
33	NB	479	
34	ND	257	
35	NF	151	
36	NG	151	
37	NH	1146	
38	NI	280	
39	NL	313	
40	NM	264	
41	NO	130	
42	NP	145	
43	NQ	84	
44	NR	861	
45	NS	1157	
46	NT	156	
47	NU	135	
48	NV	885	
49	SA	594	
50	SB	529	
51	SC	321	
51	SD	321	
52	SE	128	
52	SF	128	
53	SG	475	
54	SH	373	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
55	SI	1282	
56	SJ	244	
56	SK	244	
57	SL	198	
58	SM	291	
59	SP	2785	
60	SQ	756	
61	SR	143	
62	SS	771	
63	ST	632	
64	SU	472	
65	SW	252	
66	SX	228	
67	SY	253	
68	SZ	304	

## 2 Entry composition [i](#)

There are 73 unique types of molecules in this entry. The entry contains 242409 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 5'ETS rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	L0	16	347	154	64	113	16	0	0

- Molecule 2 is a RNA chain called 18S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	L1	1636	34979	15625	6305	11413	1636	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L1	1842	4AC	C	conflict	GB 151415227

- Molecule 3 is a RNA chain called U3 snoRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	L2	177	3779	1685	666	1251	177	0	0

- Molecule 4 is a protein called 40S ribosomal protein S18.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	L3	115	854	538	164	151	1	0	0

- Molecule 5 is a protein called 40S ribosomal protein S4, X isoform.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	L4	262	2076	1324	386	358	8	0	0

- Molecule 6 is a protein called 40S ribosomal protein S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	L5	190	Total	C	N	O	S	0	0
			1501	939	285	270	7		

- Molecule 7 is a protein called 40S ribosomal protein S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	L6	230	Total	C	N	O	S	0	0
			1862	1164	371	320	7		

- Molecule 8 is a protein called 40S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	L7	168	Total	C	N	O	S	0	0
			1346	862	239	244	1		

- Molecule 9 is a protein called 40S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	L8	180	Total	C	N	O	S	0	0
			1474	925	294	250	5		

- Molecule 10 is a protein called 40S ribosomal protein S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	L9	180	Total	C	N	O	S	0	0
			1499	955	300	242	2		

- Molecule 11 is a protein called 40S ribosomal protein S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	LA	120	Total	C	N	O	S	0	0
			931	584	164	174	9		

- Molecule 12 is a protein called 40S ribosomal protein S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	LC	136	Total	C	N	O	S	0	0
			1068	679	202	184	3		

- Molecule 13 is a protein called 40S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	LD	151	Total	C	N	O	S	0	0
			1229	782	230	211	6		

- Molecule 14 is a protein called 40S ribosomal protein S24.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	LF	126	Total	C	N	O	S	0	0
			1030	650	201	173	6		

- Molecule 15 is a protein called 40S ribosomal protein S28.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	LG	62	Total	C	N	O	S	0	0
			488	297	97	92	2		

- Molecule 16 is a protein called WD repeat-containing protein 75.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	LH	746	Total	C	N	O	S	0	0
			5987	3846	1005	1101	35		

- Molecule 17 is a protein called Nucleolar protein 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	LI	537	Total	C	N	O	S	0	0
			3889	2508	654	706	21		

- Molecule 18 is a protein called U3 small nucleolar RNA-associated protein 15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	LJ	462	Total	C	N	O	S	0	0
			3658	2337	629	678	14		

- Molecule 19 is a protein called WD repeat-containing protein 43.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	LK	118	Total	C	N	O	S	0	0
			943	612	163	163	5		
19	LL	510	Total	C	N	O	S	0	0
			3982	2538	686	731	27		

- Molecule 20 is a protein called HEAT repeat-containing protein 1.



Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	LM	2041	16094	10437	2639	2936	82	0	0

- Molecule 21 is a protein called U3 small nucleolar RNA-associated protein 4 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	LN	671	5299	3394	925	956	24	0	0

- Molecule 22 is a protein called Periodic tryptophan protein 2 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	LO	819	6454	4118	1107	1196	33	0	0

- Molecule 23 is a protein called U3 small nucleolar RNA-associated protein 6 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	LP	566	4696	3017	806	845	28	0	0

- Molecule 24 is a protein called WD repeat-containing protein 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	LQ	830	6497	4140	1122	1202	33	0	0

- Molecule 25 is a protein called Transducin beta-like protein 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	LR	772	6006	3783	1077	1115	31	0	0

- Molecule 26 is a protein called U3 small nucleolar RNA-associated protein 18 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	LS	453	3560	2235	631	671	23	0	0

- Molecule 27 is a protein called WD repeat-containing protein 36.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	LT	855	6645	4257	1133	1223	32	0	0

- Molecule 28 is a protein called DDB1- and CUL4-associated factor 13.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	LU	445	3611	2282	653	651	25	0	0

- Molecule 29 is a protein called WD repeat-containing protein 46.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	LW	452	3519	2217	645	642	15	0	0

- Molecule 30 is a protein called Active regulator of SIRT1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	LY	95	778	480	154	142	2	0	0

- Molecule 31 is a protein called U3 small nucleolar ribonucleoprotein protein IMP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	LZ	162	1344	851	250	239	4	0	0

- Molecule 32 is a protein called U3 small nucleolar ribonucleoprotein protein MPP10.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	NA	295	2404	1517	421	459	7	0	0

- Molecule 33 is a protein called Something about silencing protein 10.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
33	NB	103	849	522	181	146	0	0

- Molecule 34 is a protein called Nucleolar protein 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	ND	83	Total	C	N	O	S	0	0
			685	431	139	114	1		

- Molecule 35 is a protein called 40S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	NF	149	Total	C	N	O	S	0	0
			1202	770	228	203	1		

- Molecule 36 is a protein called 40S ribosomal protein S14.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	NG	127	Total	C	N	O	S	0	0
			957	585	189	177	6		

- Molecule 37 is a protein called Nucleolar protein 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	NH	1066	Total	C	N	O	S	0	0
			8374	5345	1491	1506	32		

- Molecule 38 is a protein called Ribosomal RNA-processing protein 7 homolog A.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	NI	234	Total	C	N	O	S	0	0
			1840	1171	329	336	4		

- Molecule 39 is a protein called Probable dimethyladenosine transferase.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	NL	278	Total	C	N	O	S	0	0
			2213	1428	392	384	9		

- Molecule 40 is a protein called 40S ribosomal protein S3a.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	NM	233	Total	C	N	O	S	0	0
			1873	1186	339	334	14		

- Molecule 41 is a protein called 40S ribosomal protein S15a.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	NO	129	Total	C	N	O	S	0	0
			1034	659	193	176	6		

- Molecule 42 is a protein called 40S ribosomal protein S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	NP	144	Total	C	N	O	S	0	0
			1122	703	217	199	3		

- Molecule 43 is a protein called 40S ribosomal protein S27.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	NQ	82	Total	C	N	O	S	0	0
			640	402	118	113	7		

- Molecule 44 is a protein called RRP12-like protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	NR	861	Total	C	N	O	0	0
			4305	2583	861	861		

- Molecule 45 is a protein called Probable ATP-dependent RNA helicase DHX37.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	NS	914	Total	C	N	O	S	0	0
			5332	3264	1032	1026	10		

- Molecule 46 is a protein called Ubiquitin-40S ribosomal protein S27a.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	NT	58	Total	C	N	O	S	0	0
			470	295	89	79	7		

- Molecule 47 is a protein called 40S ribosomal protein S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
47	NU	60	Total	C	N	O	S	0	0
			495	314	98	81	2		

- Molecule 48 is a protein called Exosome component 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
48	NV	42	Total	C	N	O	S	0	0
			344	220	55	68	1		

- Molecule 49 is a protein called Nucleolar protein 56.

Mol	Chain	Residues	Atoms					AltConf	Trace
49	SA	396	Total	C	N	O	S	0	0
			3077	1948	542	575	12		

- Molecule 50 is a protein called Nucleolar protein 58.

Mol	Chain	Residues	Atoms					AltConf	Trace
50	SB	430	Total	C	N	O	S	0	0
			3371	2138	584	628	21		

- Molecule 51 is a protein called rRNA 2'-O-methyltransferase fibrillar.

Mol	Chain	Residues	Atoms					AltConf	Trace
51	SC	229	Total	C	N	O	S	0	0
			1781	1129	322	323	7		
51	SD	236	Total	C	N	O	S	0	0
			1832	1157	335	333	7		

- Molecule 52 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
52	SE	125	Total	C	N	O	S	0	0
			968	611	172	180	5		
52	SF	123	Total	C	N	O	S	0	0
			955	604	170	176	5		

- Molecule 53 is a protein called U3 small nucleolar RNA-interacting protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
53	SG	389	Total	C	N	O	S	1	0
			3042	1910	557	562	13		

- Molecule 54 is a protein called RNA 3'-terminal phosphate cyclase-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
54	SH	368	Total	C	N	O	S	0	0
			2832	1803	495	518	16		

- Molecule 55 is a protein called Ribosome biogenesis protein BMS1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
55	SI	803	6493	4154	1184	1122	33	0	0

- Molecule 56 is a protein called Ribosomal RNA small subunit methyltransferase NEP1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
56	SJ	204	1579	1012	272	286	9	0	0
56	SK	204	1579	1012	272	286	9	0	0

- Molecule 57 is a protein called rRNA-processing protein FCF1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
57	SL	175	1433	910	259	249	15	0	0

- Molecule 58 is a protein called U3 small nucleolar ribonucleoprotein protein IMP4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
58	SM	273	2253	1415	433	396	9	0	0

- Molecule 59 is a protein called Small subunit processome component 20 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
59	SP	2485	20032	12970	3354	3606	102	0	0

- Molecule 60 is a protein called Deoxynucleotidyltransferase terminal-interacting protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
60	SQ	156	1275	804	236	230	5	0	0

- Molecule 61 is a protein called 40S ribosomal protein S23.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
61	SR	141	1098	693	219	183	3	0	0

- Molecule 62 is a protein called U3 small nucleolar RNA-associated protein 14 homolog A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
62	SS	350	2824	1788	516	516	4	0	0

- Molecule 63 is a protein called Nucleolar protein 14.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
63	ST	570	4188	2646	764	757	21	0	0

- Molecule 64 is a protein called Nucleolar complex protein 4 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
64	SU	442	3154	2030	565	551	8	0	0

- Molecule 65 is a protein called RNA-binding protein PNO1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
65	SW	195	1522	977	273	267	5	0	0

- Molecule 66 is a protein called Unassigned peptides.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
66	SX	228	1140	684	228	228	0	0

- Molecule 67 is a protein called Probable U3 small nucleolar RNA-associated protein 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
67	SY	237	2013	1272	385	350	6	0	0

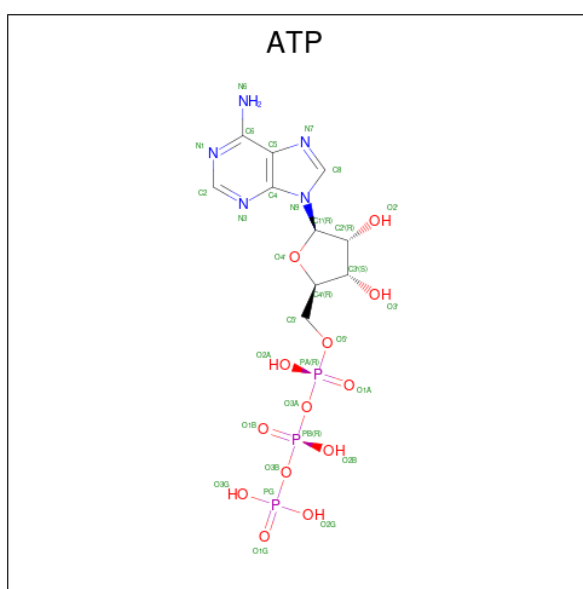
- Molecule 68 is a protein called Bystin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
68	SZ	290	2222	1438	392	383	9	0	0

- Molecule 69 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
69	L1	61	Total Mg 61 61	0
69	NH	1	Total Mg 1 1	0
69	SI	1	Total Mg 1 1	0
69	SL	1	Total Mg 1 1	0

- Molecule 70 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



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

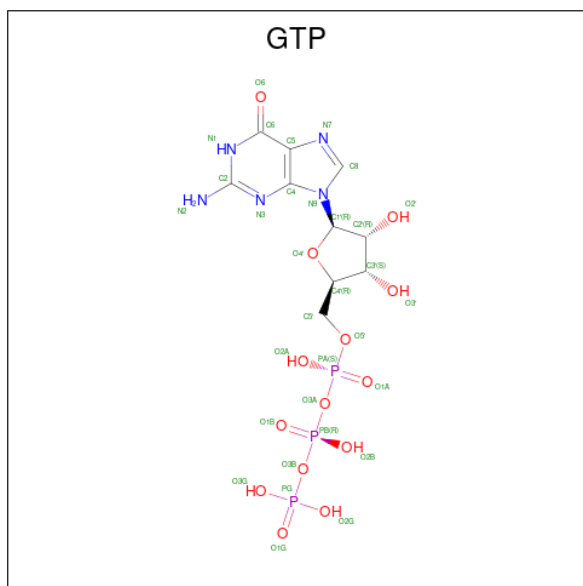
- Molecule 71 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
71	NQ	1	Total Zn 1 1	0
71	NT	1	Total Zn 1 1	0
71	SL	1	Total Zn 1 1	0

- Molecule 72 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:

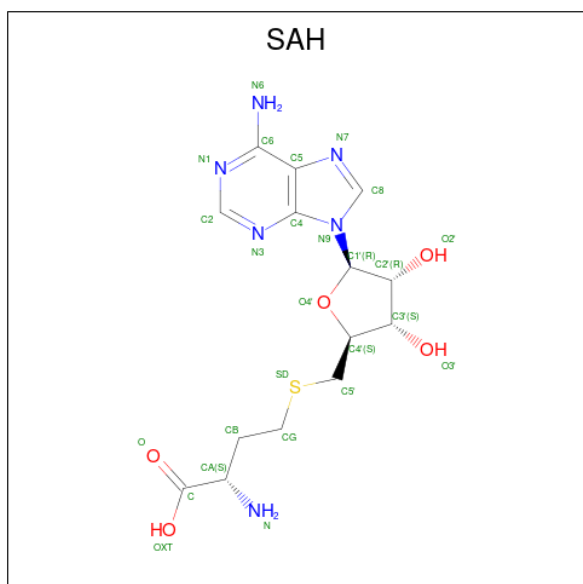


$C_{10}H_{16}N_5O_{14}P_3$ ).



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

- Molecule 73 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ).



Mol	Chain	Residues	Atoms				AltConf	
			Total	C	N	O		S
73	SJ	1	26	14	6	5	1	0

*Continued on next page...*

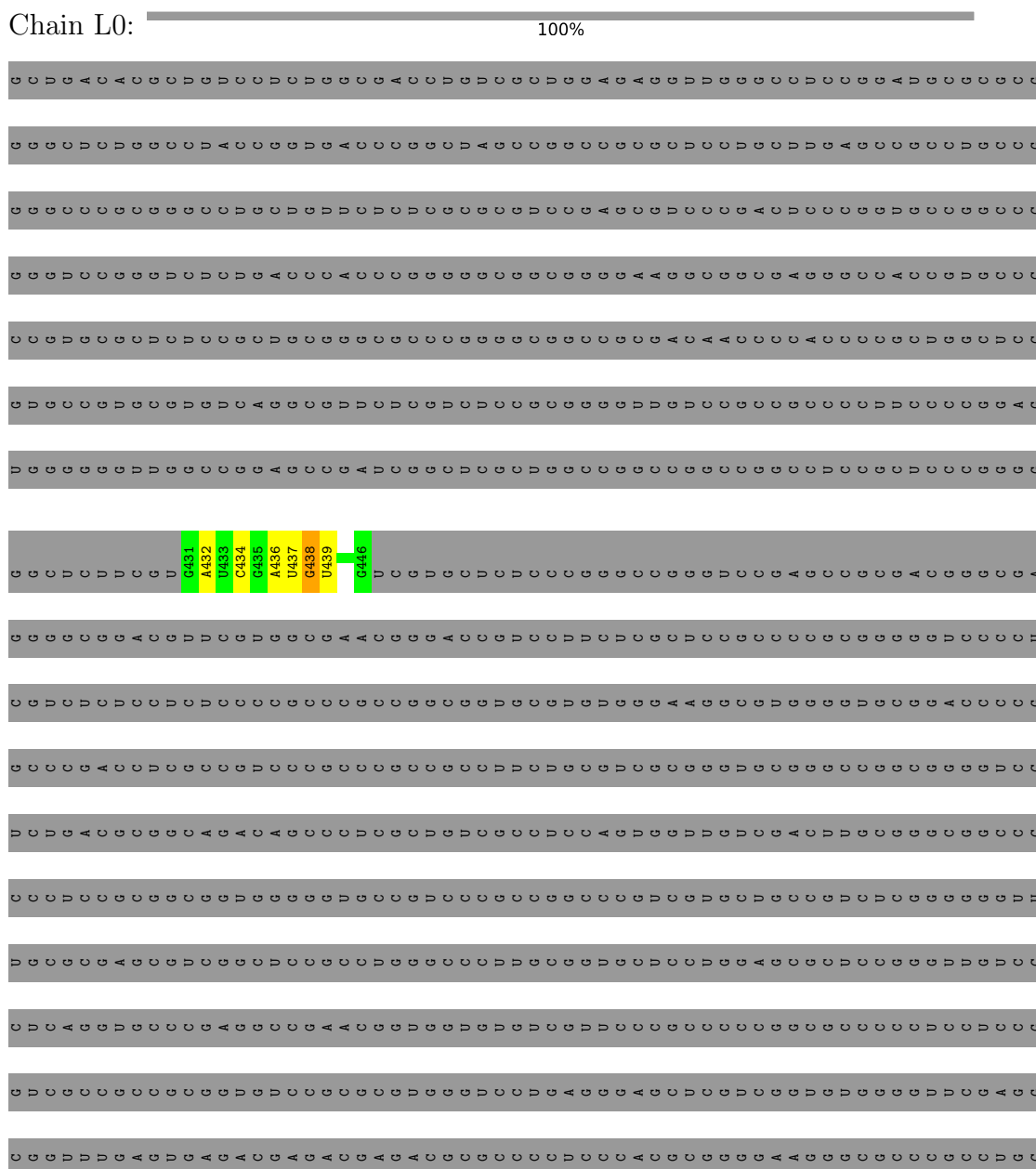
*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	S	
73	SK	1	26	14	6	5	1	0

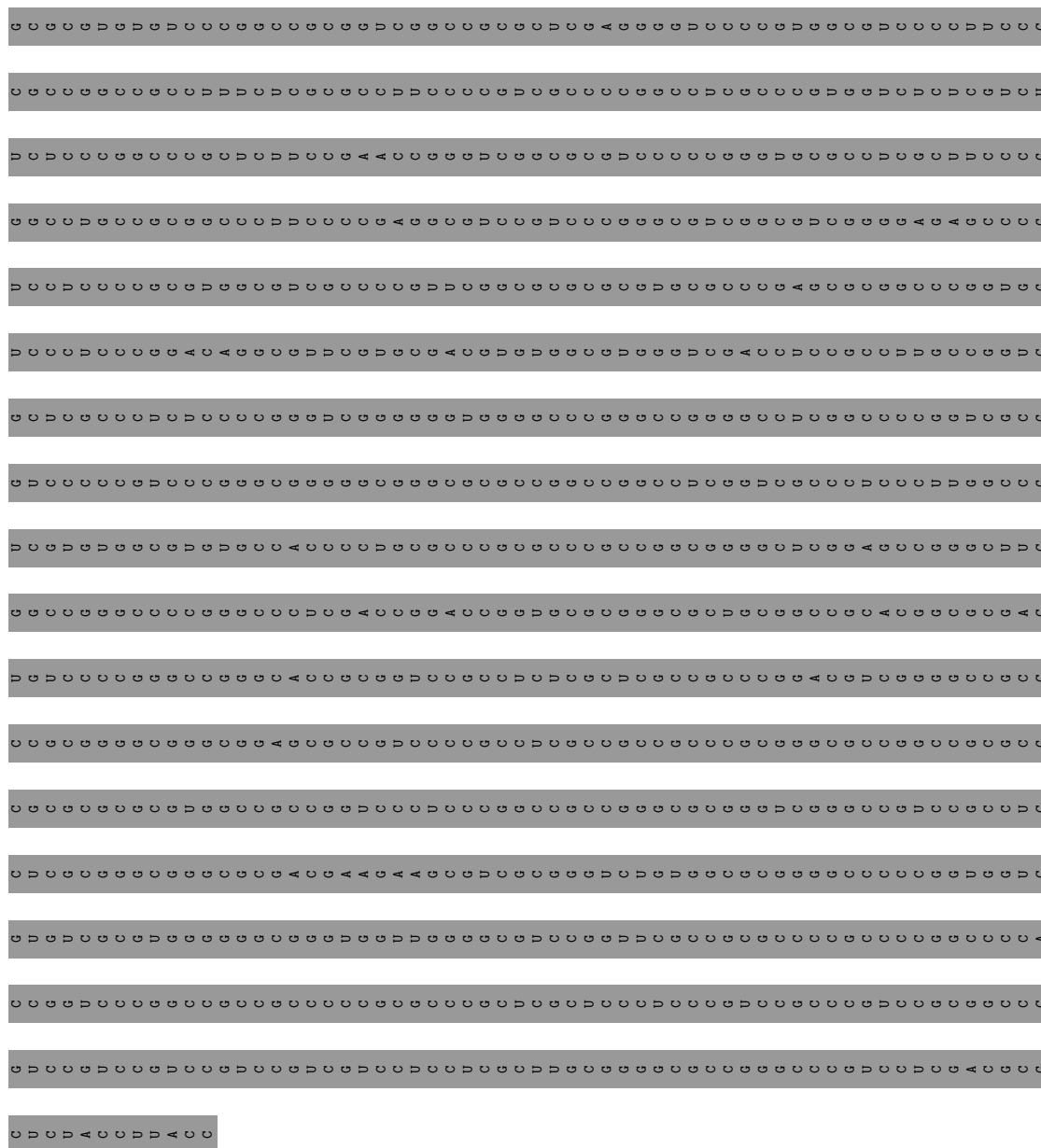
### 3 Residue-property plots

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.

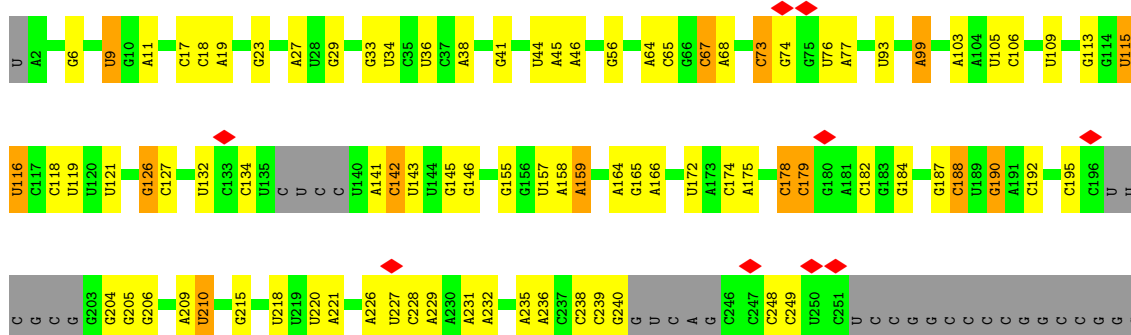
- Molecule 1: 5'ETS rRNA

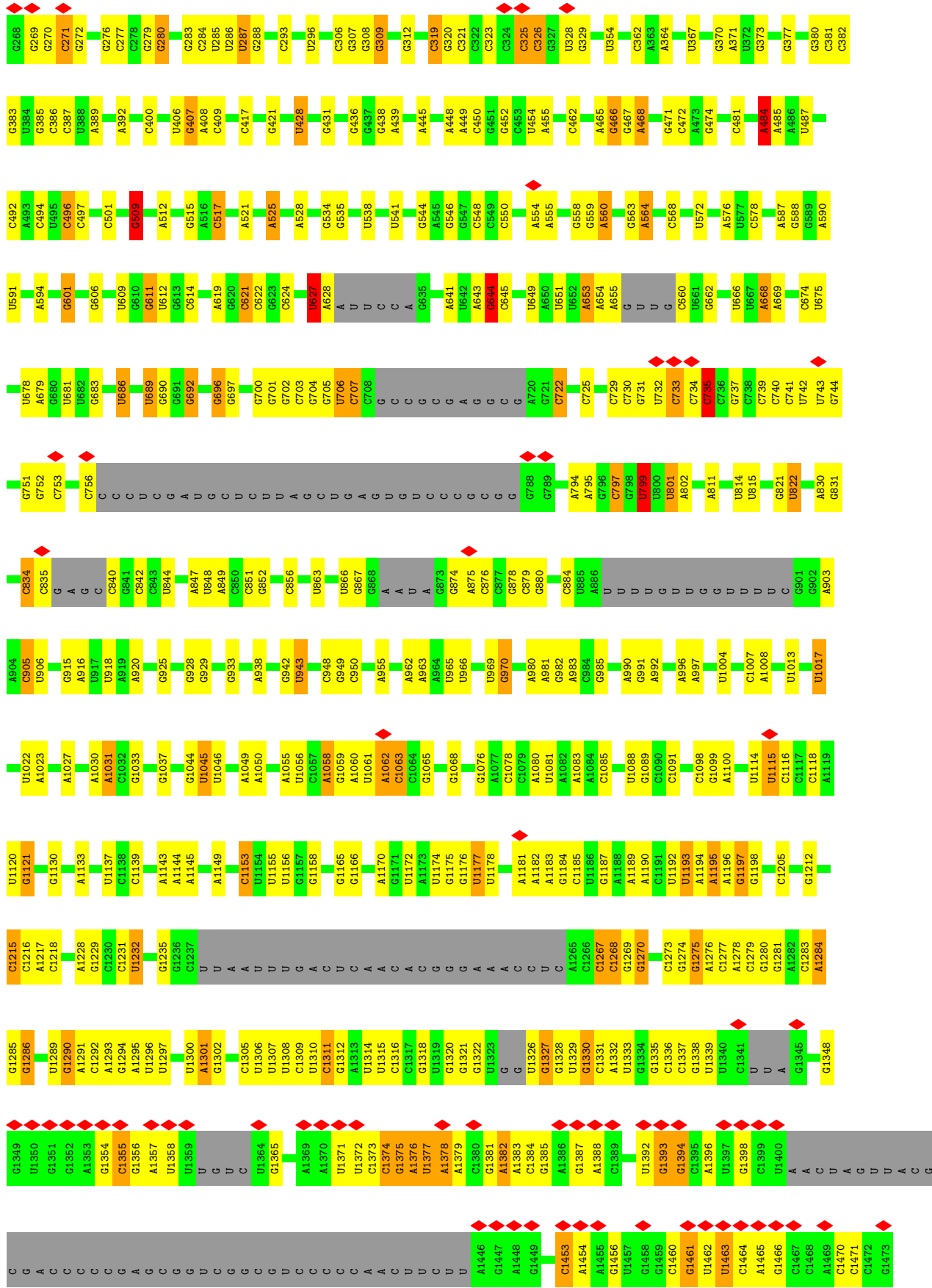


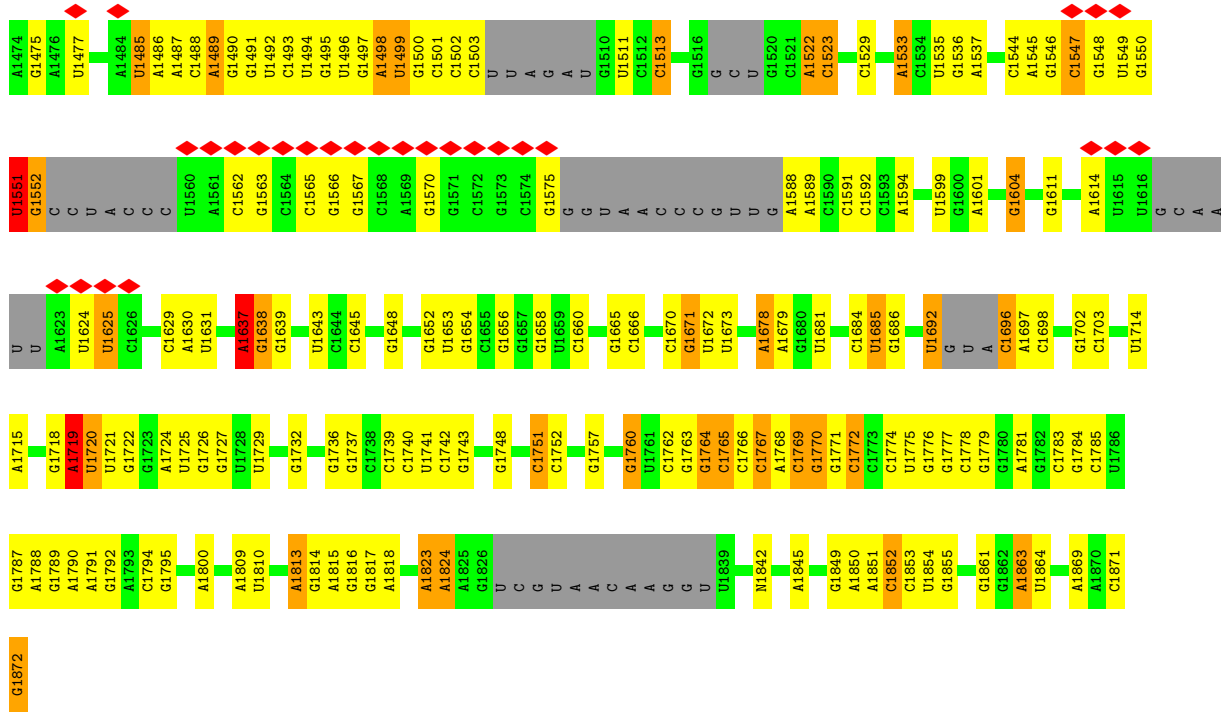
D D D C G G D D G A G C C G A C G G D C C C G D G G C C D D G G C C G G D G C C G G C C G G C C G D D G A G C G G A  
D C G C G G D G G G D D D C G G G C C G G D G D G A C G C G D D G C C G C G C G G C C G G C C G G A G G C D G C C G D D  
C D G C C D C C G A C C G G D C C G D D G D D G C C A C D D C G G A C G C G C D D G C C D C G G A A G C A A G G A  
G D D G G D D G G A C G G G G G G C C D D G G D G G G D D C G C G A C G C G C G A C C C G C C G G C C C G G  
C C C D G A A C G C C A A C G G D C G A G G D G G C C G C G C G C A G D D D C D D G G D A C C G G C A G G C C C C  
G D C C C D D C C C C A A G C C G D C C C D C G G C C D C D G G G C C G A G A A A C C G G C D G C G G G D  
G G G G G A G D G D G A C C C A C C C D D G G G D G A G C C A A A G C C D D C D A G C G A D U G A G A G C G G D G C  
C D D G G G G G D G A C C G G A D U C C C G G G C C C C G C C D D C D D C C C D D A G C C D D A G C C C  
D G C C C G D A G C G A A C C C G D C C G A G A G A C C C D C C C D D C C C C D C G A C G G G D D G G G G  
G G A A A C C G A A C G D D C C G C G G G G C C A C C G C G D G G D G C C G A D D G G G C D C G D C C D A C D  
G D G G C C G G C C C C C C C C C D D C C G A D C C G G G G A G G A D C C G C C G G G C C C G G D D  
C C C A G C G G D D G G A C C G G G G G G G C C G G G G G D D G G G G G G C C C G G G C C D D G D C  
G G C C G C G D G A C C C C D C C G C G A D D C C G C G D C C C C G C C C G G D C C G A G C C G D G D G A  
C C G G D C C C G A C A C C C C C D D D G C C G D D G C C A C G G G D C G C C C C D G G C C A A A G C  
G D C C C A C G G D G G G G C C C C C G G D D C C C G G A C C G G A C C G G G A D D G G G A G A G A A  
A D C A C G A C C G G A C G G D G G C C G D D G D C C G G D D C D G G D G G D C C G G D C C G G C C C  
G G D D G C G G G G G C C C C G G G C C G C G A G G C G G D D D C C G G D D G G C C G G C C G G C  
G D C C C A G G C G G G G C C C G G G A C C G C C D D G D D C D D D G G G C C G G A D C C G G C G G  
D D D D D C C C C C C C G C C G C C D D C C C C G C C A G G C C G D C A C C G C C D D C C D D G C C G G C  
C C G C C C C C C G C C A D C C D D D D D C C C C G A G G C C D C A C C G C D D C A C C D D G G D G  
G C C C C G C D D G G A C C C C G G G C C C D D G G G G A C C G C C G C C G G C C A C D A D C G  
G C C C G C G D C C C C G D C C C C G G G C C C D D G G G G A C C G G G D D G G C C C C G G C C G D G  
G G C C C G G D D C C A G G C G C D D C C C G C G G D C C C D G C C G G C C C C G G C C G A G G A  
G A C C G D D C C G C C G G A C C C C C C G C C C G C C G G D D G D C C G G G A C C C C G G A D C  
G C G A G G C C G D D G G C C C C C C C C C G G G G C C C G G G C C C G G G G A G C C C C



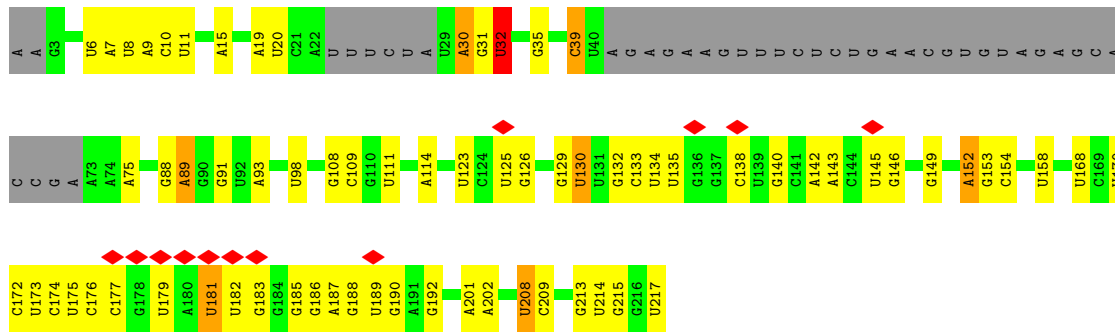
- Molecule 2: 18S rRNA



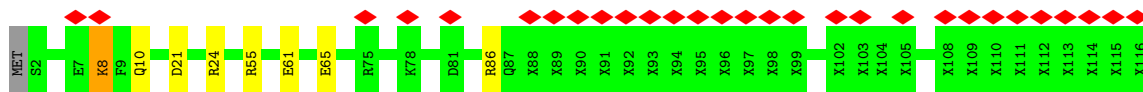
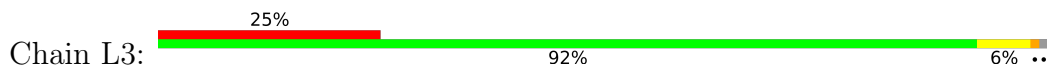




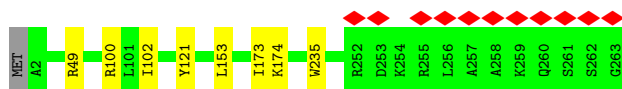
• Molecule 3: U3 snoRNA



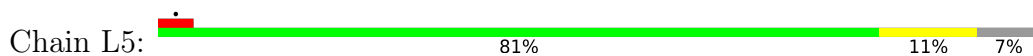
• Molecule 4: 40S ribosomal protein S18



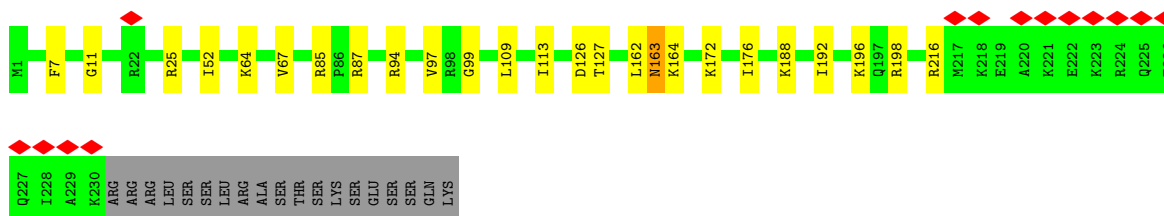
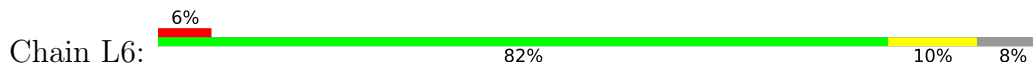
• Molecule 5: 40S ribosomal protein S4, X isoform



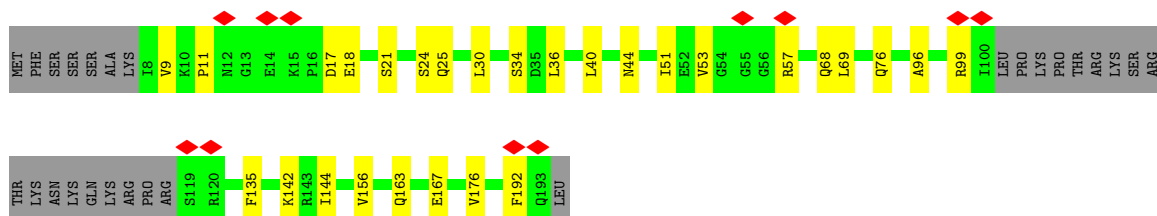
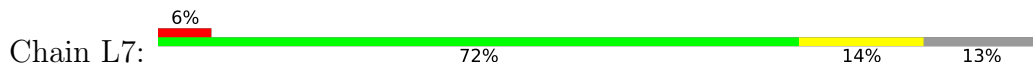
• Molecule 6: 40S ribosomal protein S5



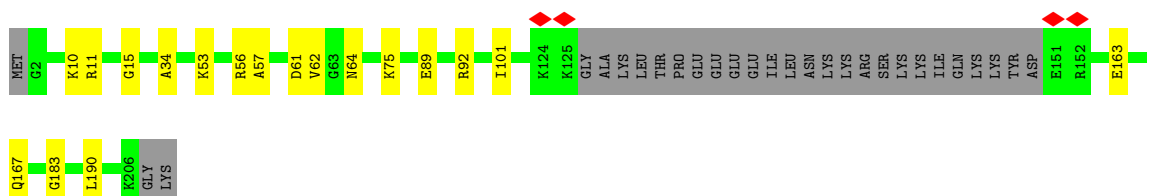
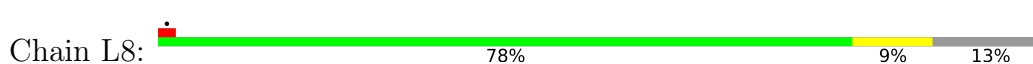
• Molecule 7: 40S ribosomal protein S6



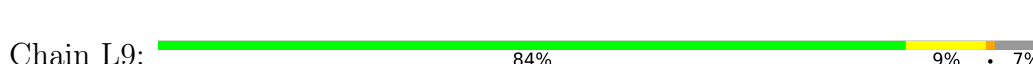
• Molecule 8: 40S ribosomal protein S7



• Molecule 9: 40S ribosomal protein S8

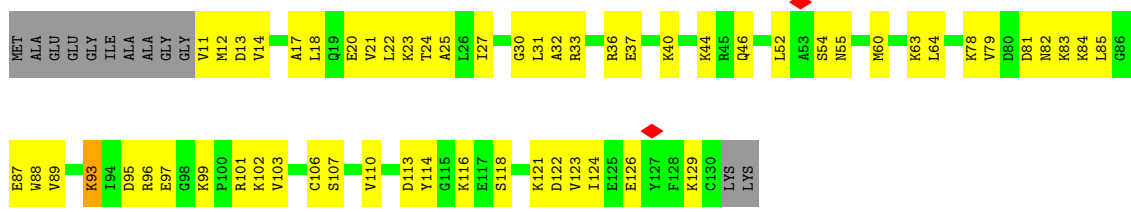


• Molecule 10: 40S ribosomal protein S9

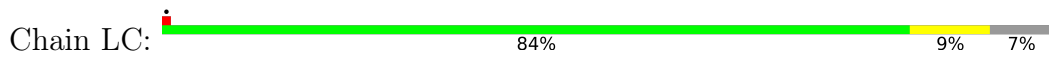


• Molecule 11: 40S ribosomal protein S12

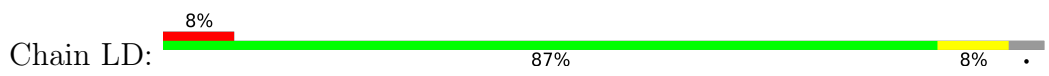




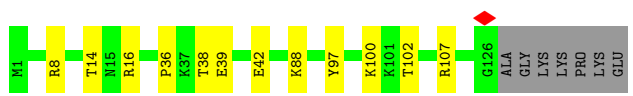
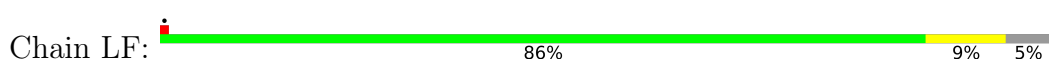
• Molecule 12: 40S ribosomal protein S16



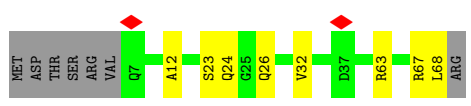
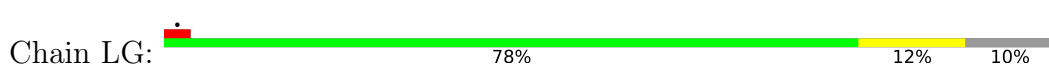
• Molecule 13: 40S ribosomal protein S11



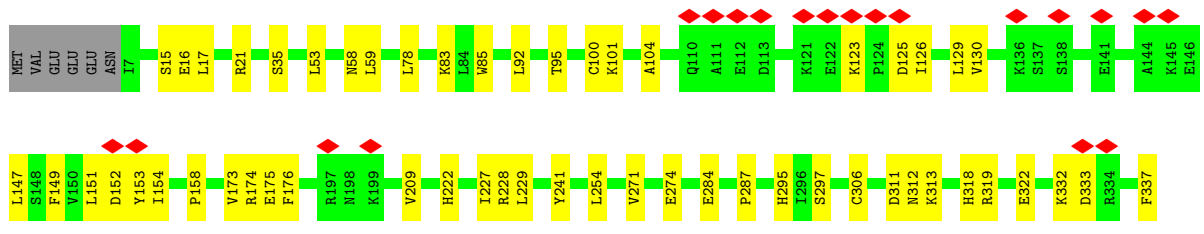
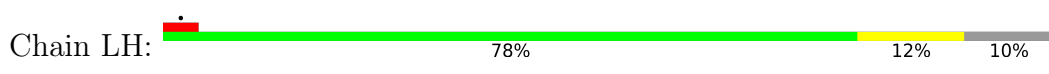
• Molecule 14: 40S ribosomal protein S24

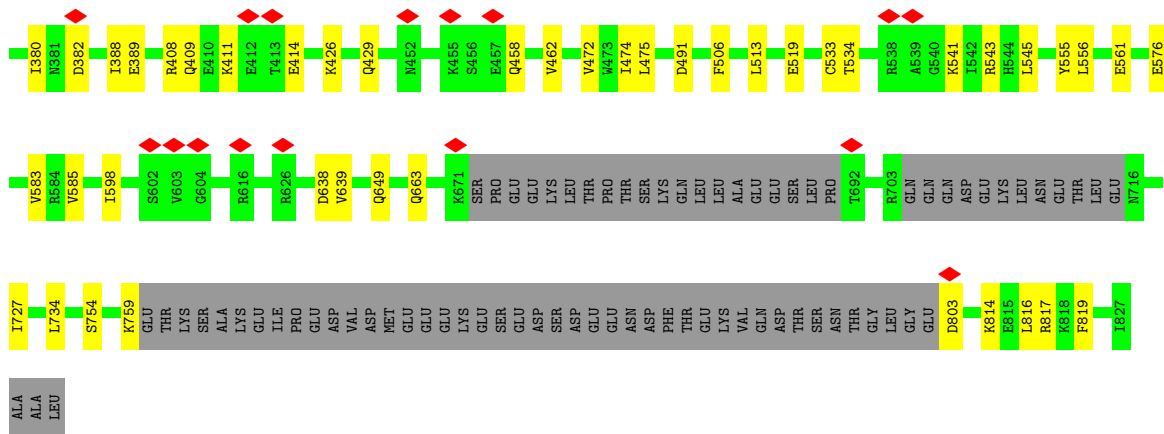


• Molecule 15: 40S ribosomal protein S28

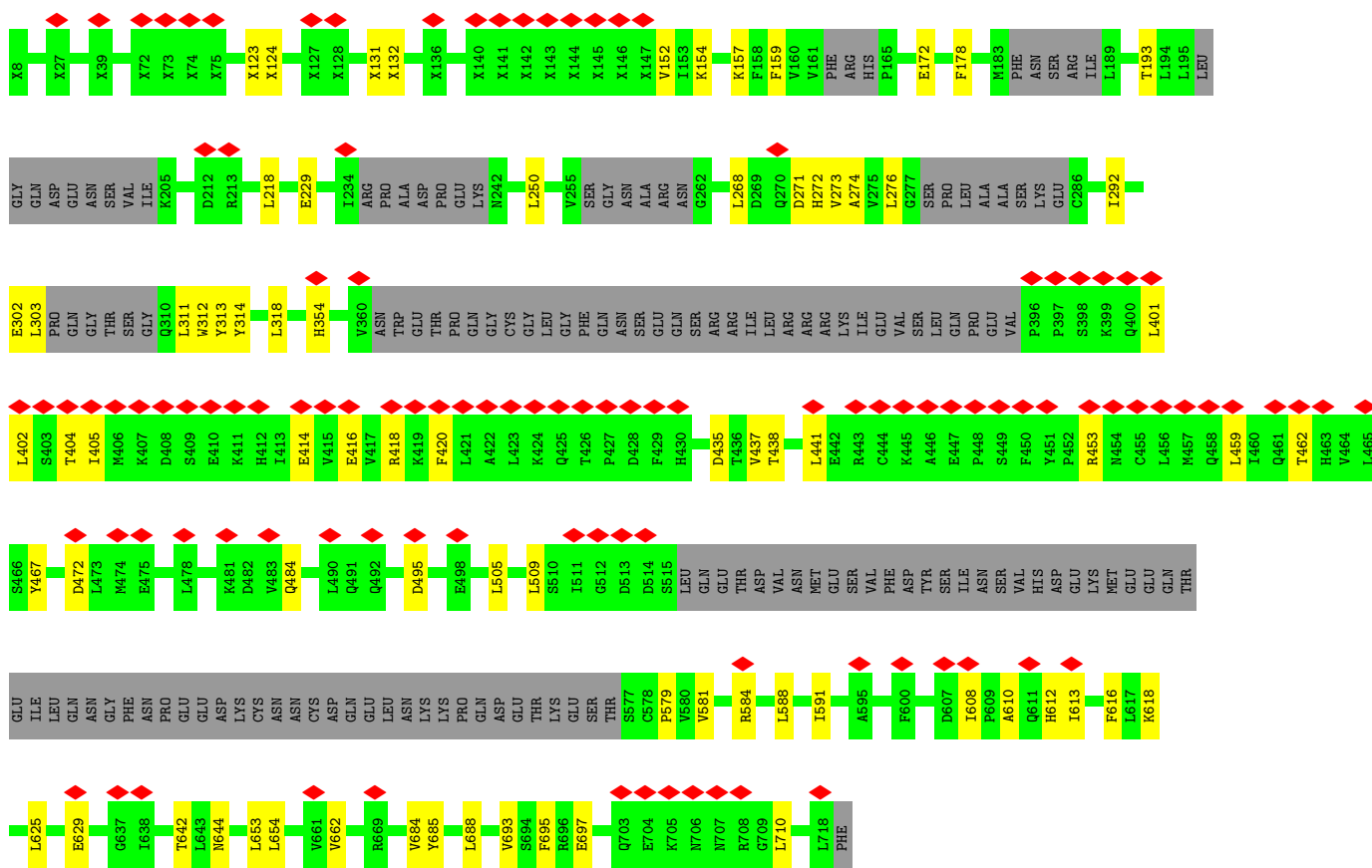


• Molecule 16: WD repeat-containing protein 75

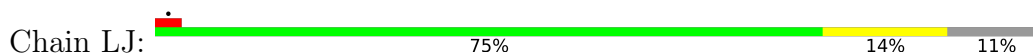


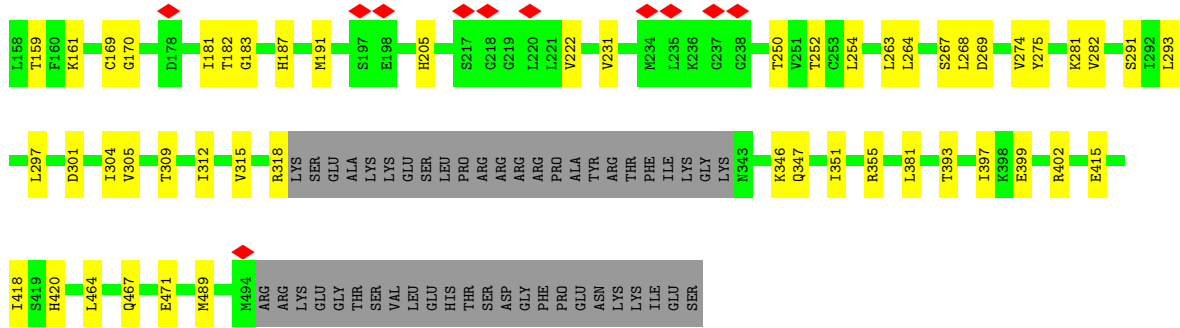


• Molecule 17: Nucleolar protein 11

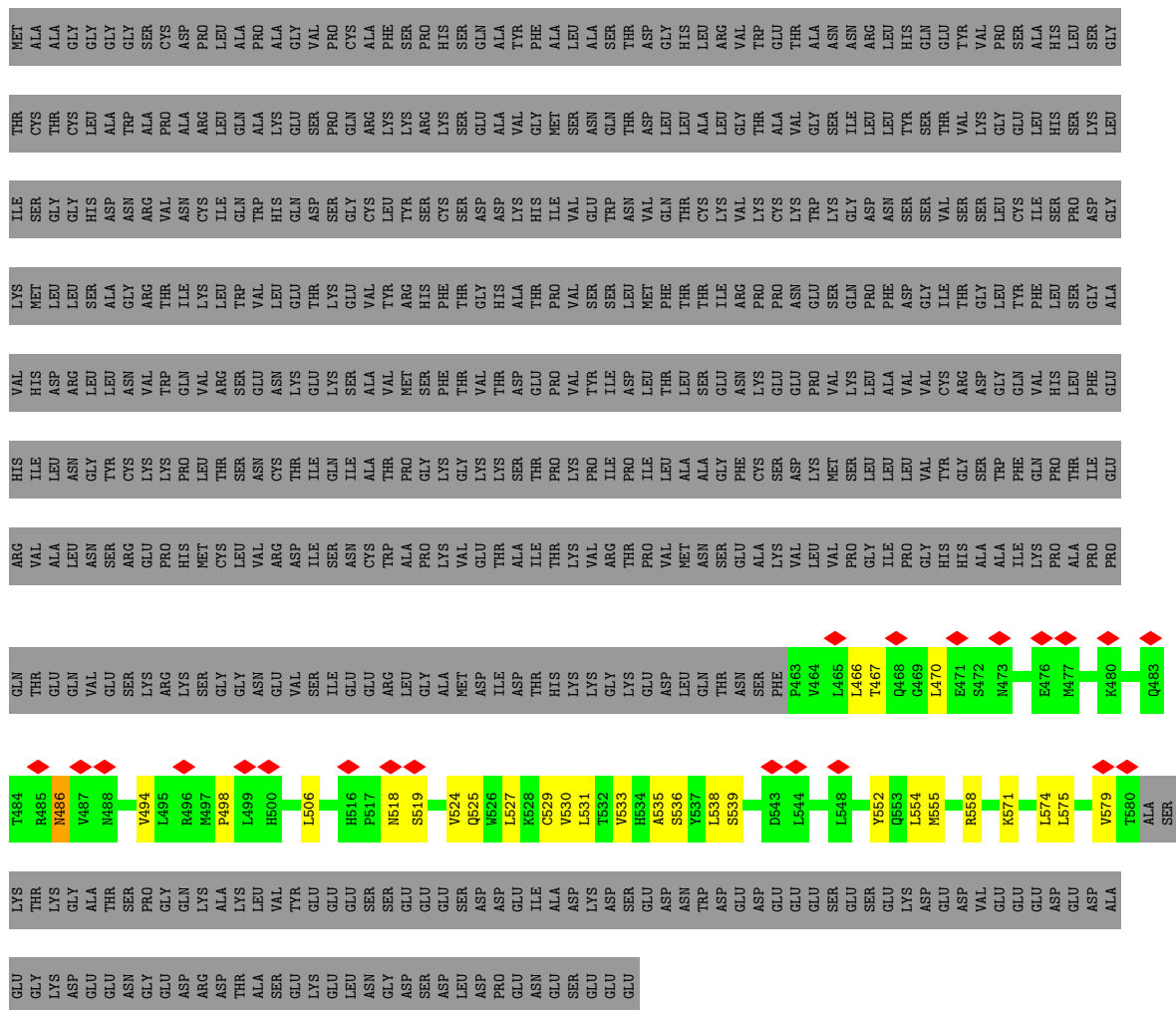


• Molecule 18: U3 small nucleolar RNA-associated protein 15 homolog

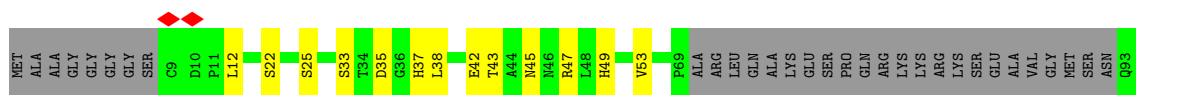


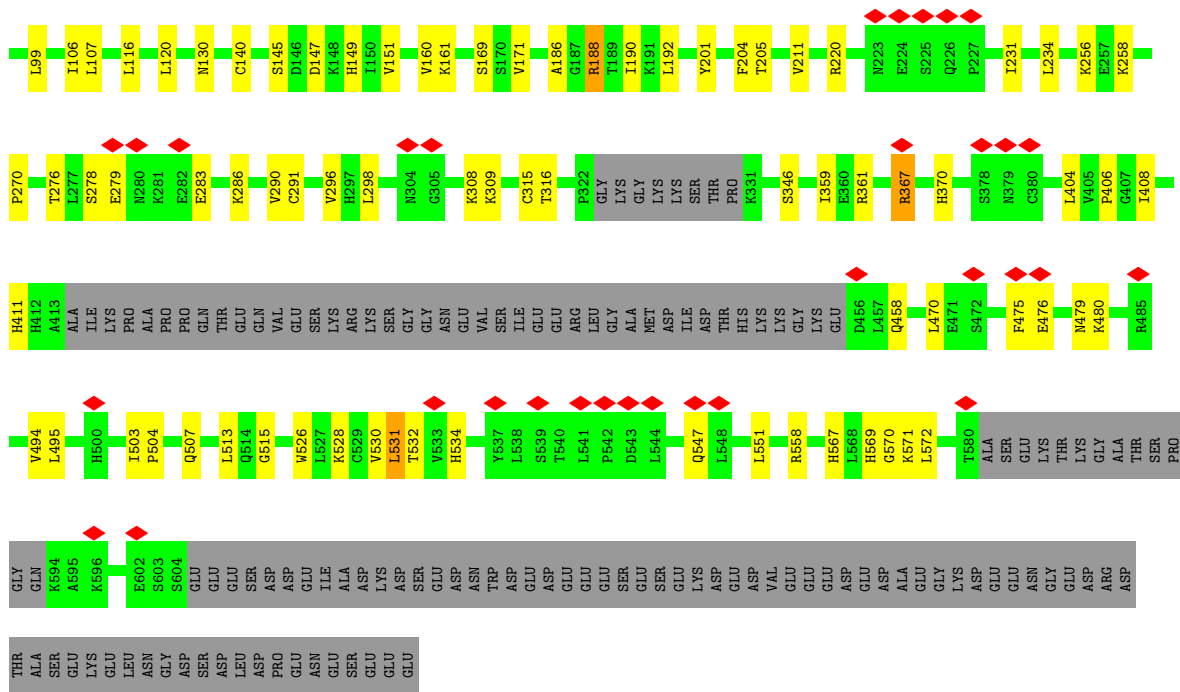


• Molecule 19: WD repeat-containing protein 43

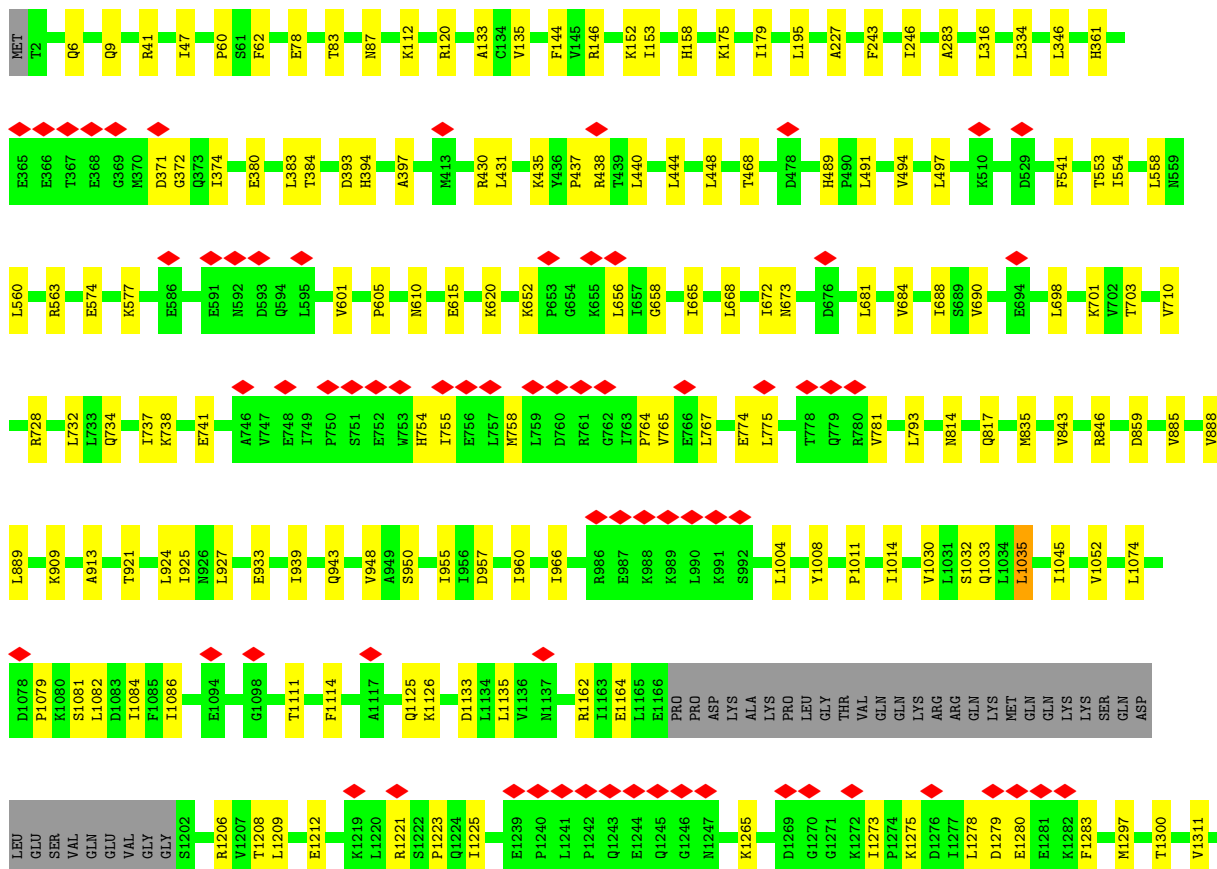
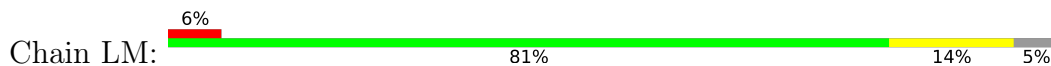


• Molecule 19: WD repeat-containing protein 43



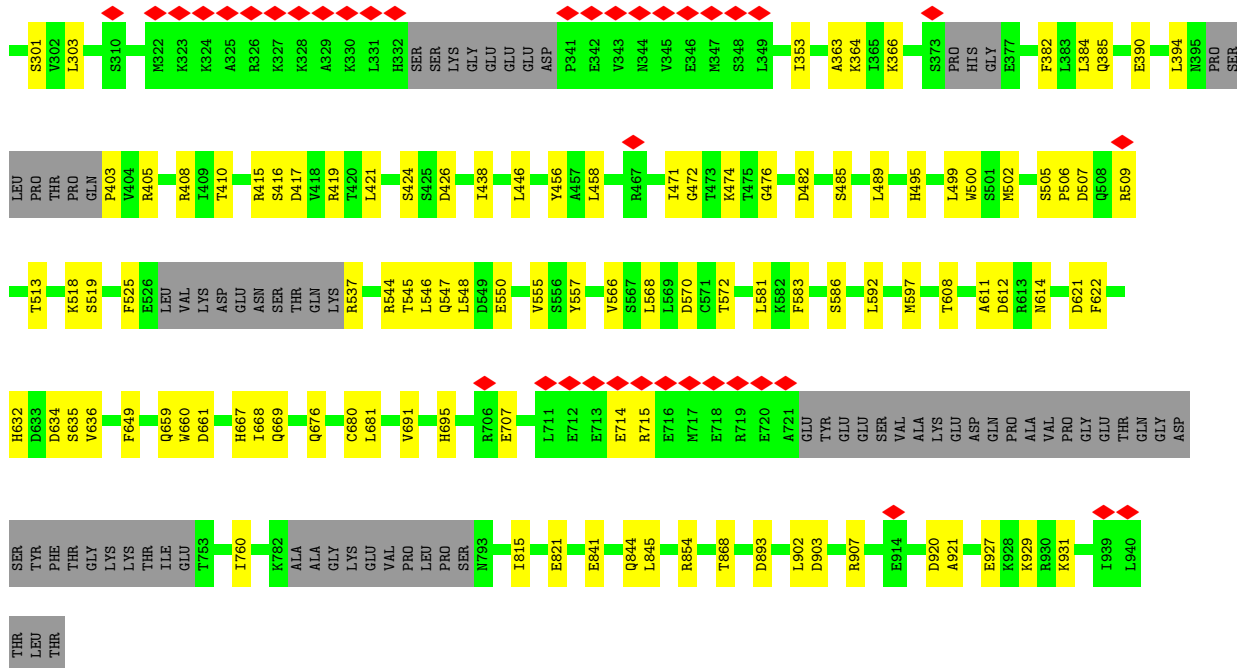


● Molecule 20: HEAT repeat-containing protein 1

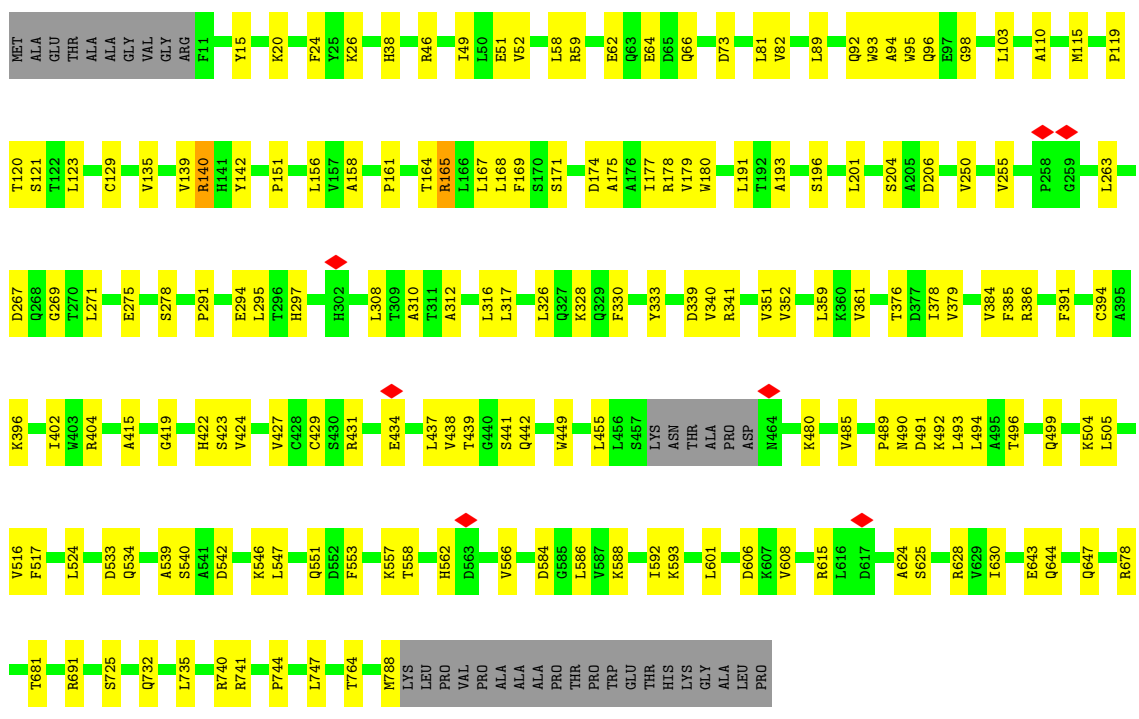
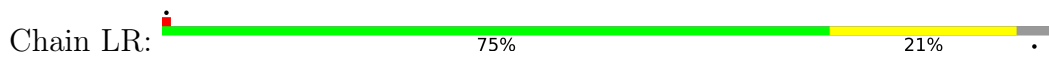




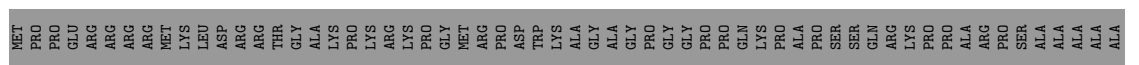


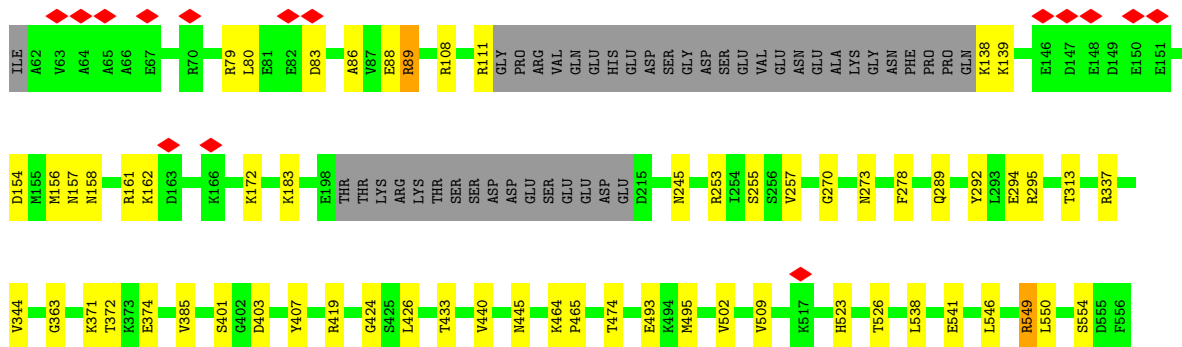


• Molecule 25: Transducin beta-like protein 3

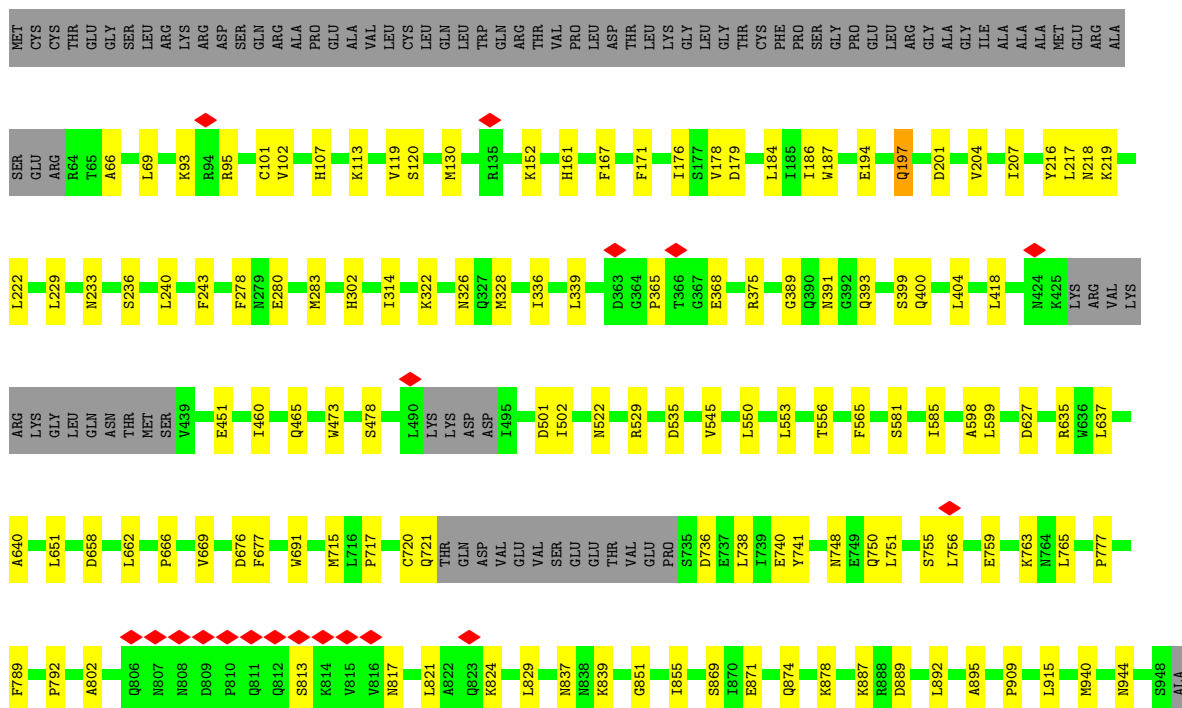
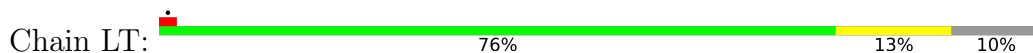


• Molecule 26: U3 small nucleolar RNA-associated protein 18 homolog

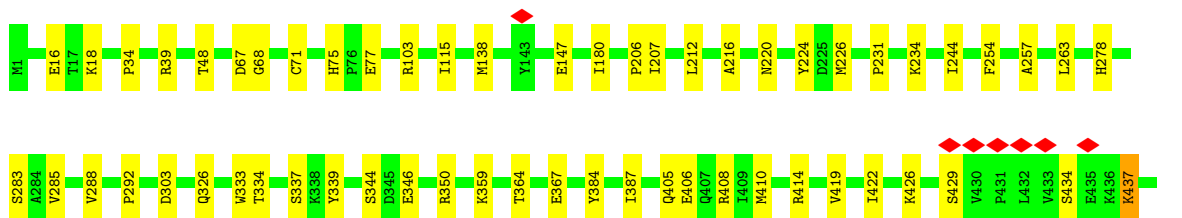




• Molecule 27: WD repeat-containing protein 36

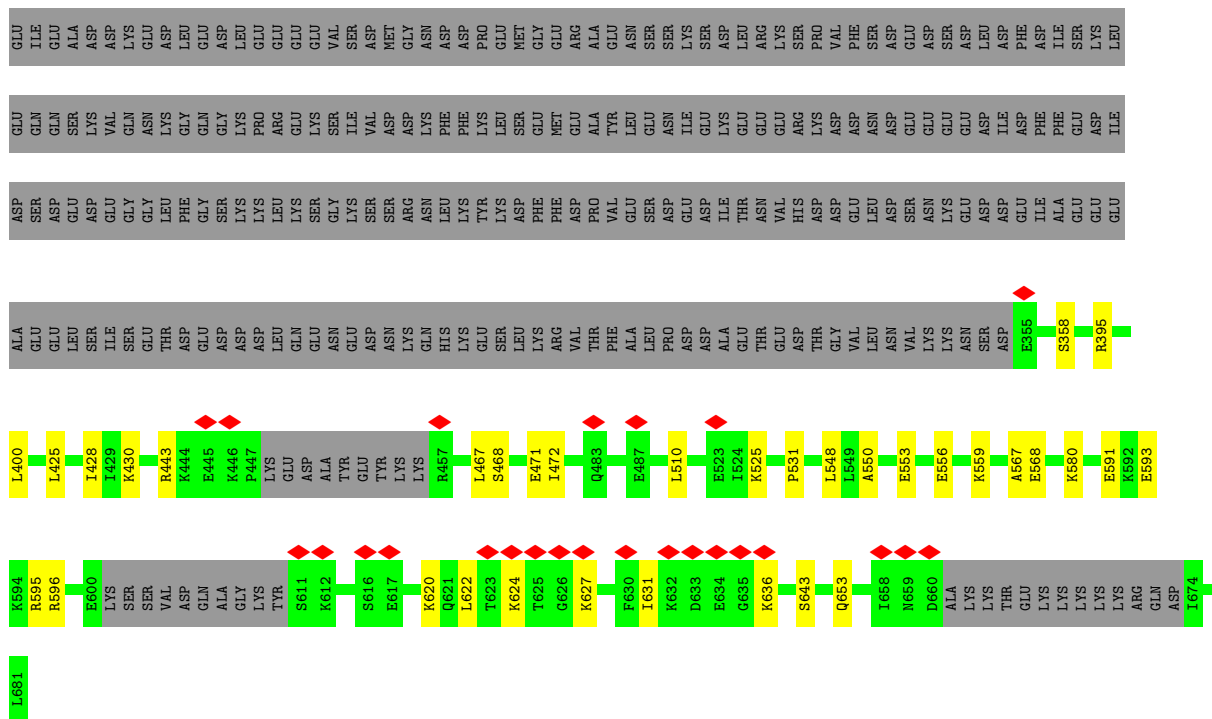


• Molecule 28: DDB1- and CUL4-associated factor 13

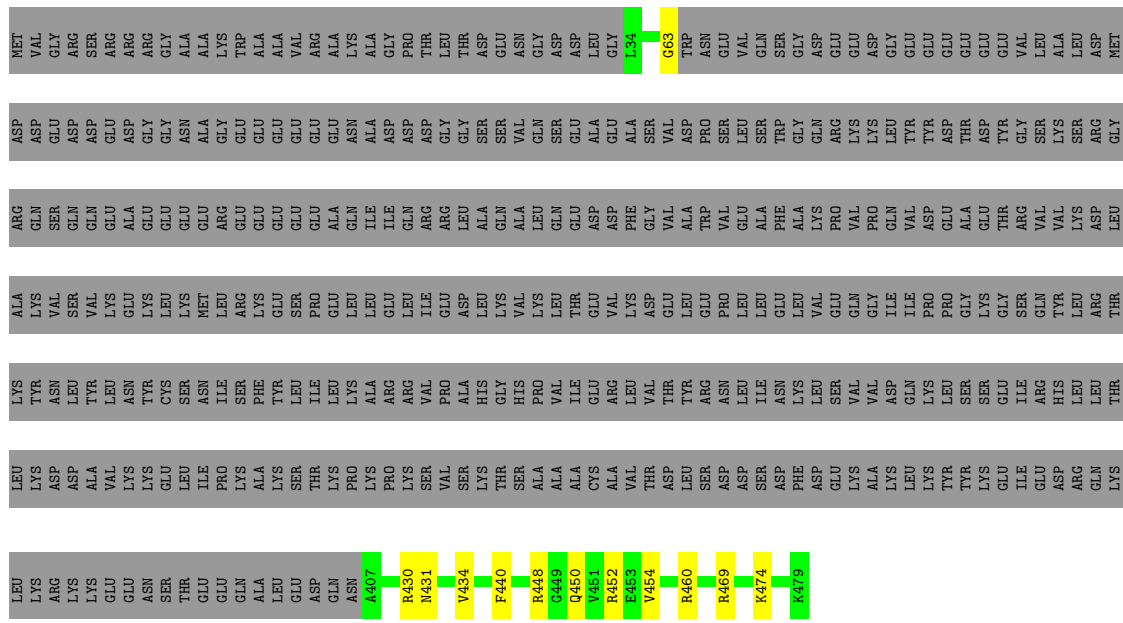




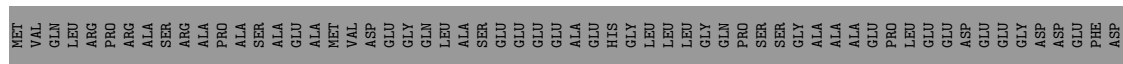




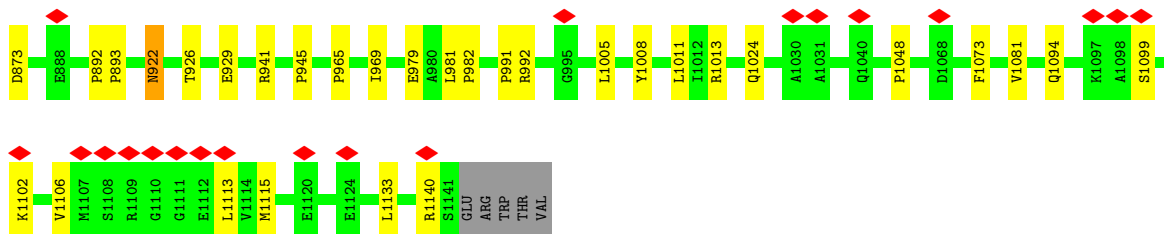
• Molecule 33: Something about silencing protein 10



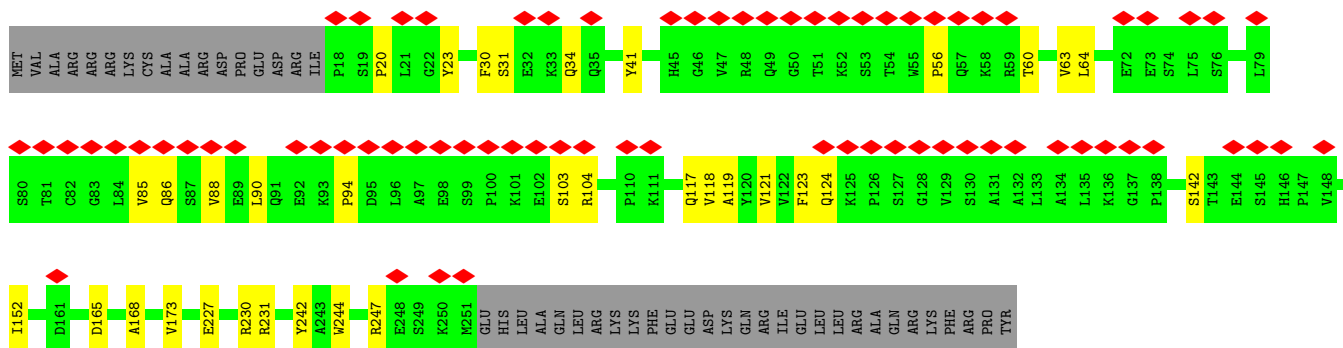
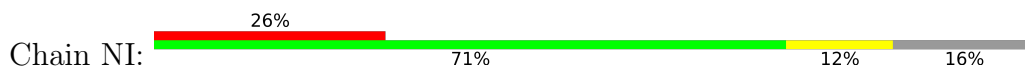
• Molecule 34: Nucleolar protein 7



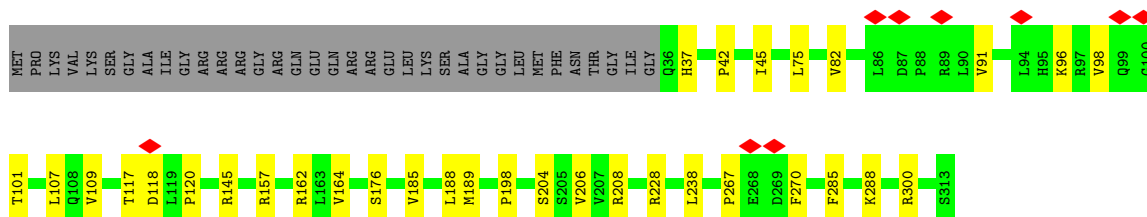
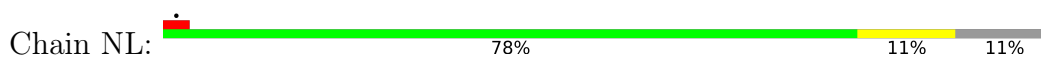




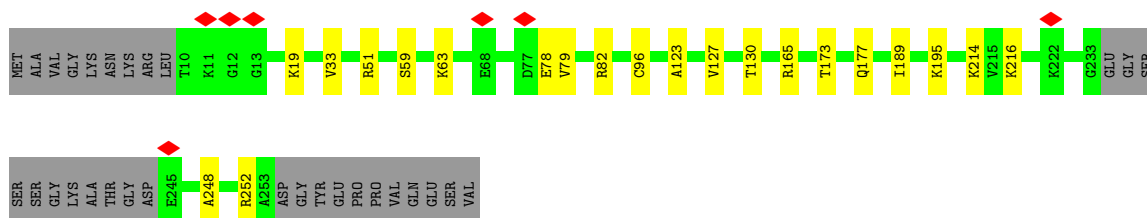
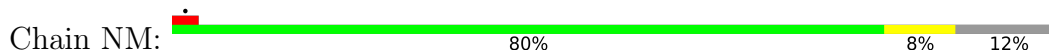
• Molecule 38: Ribosomal RNA-processing protein 7 homolog A



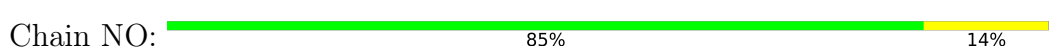
• Molecule 39: Probable dimethyladenosine transferase



• Molecule 40: 40S ribosomal protein S3a

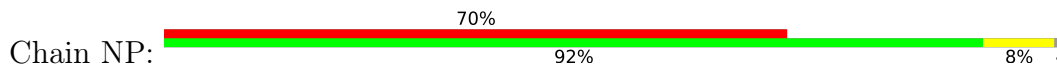


• Molecule 41: 40S ribosomal protein S15a

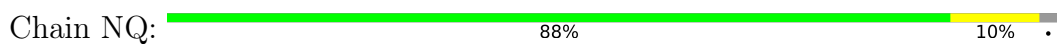




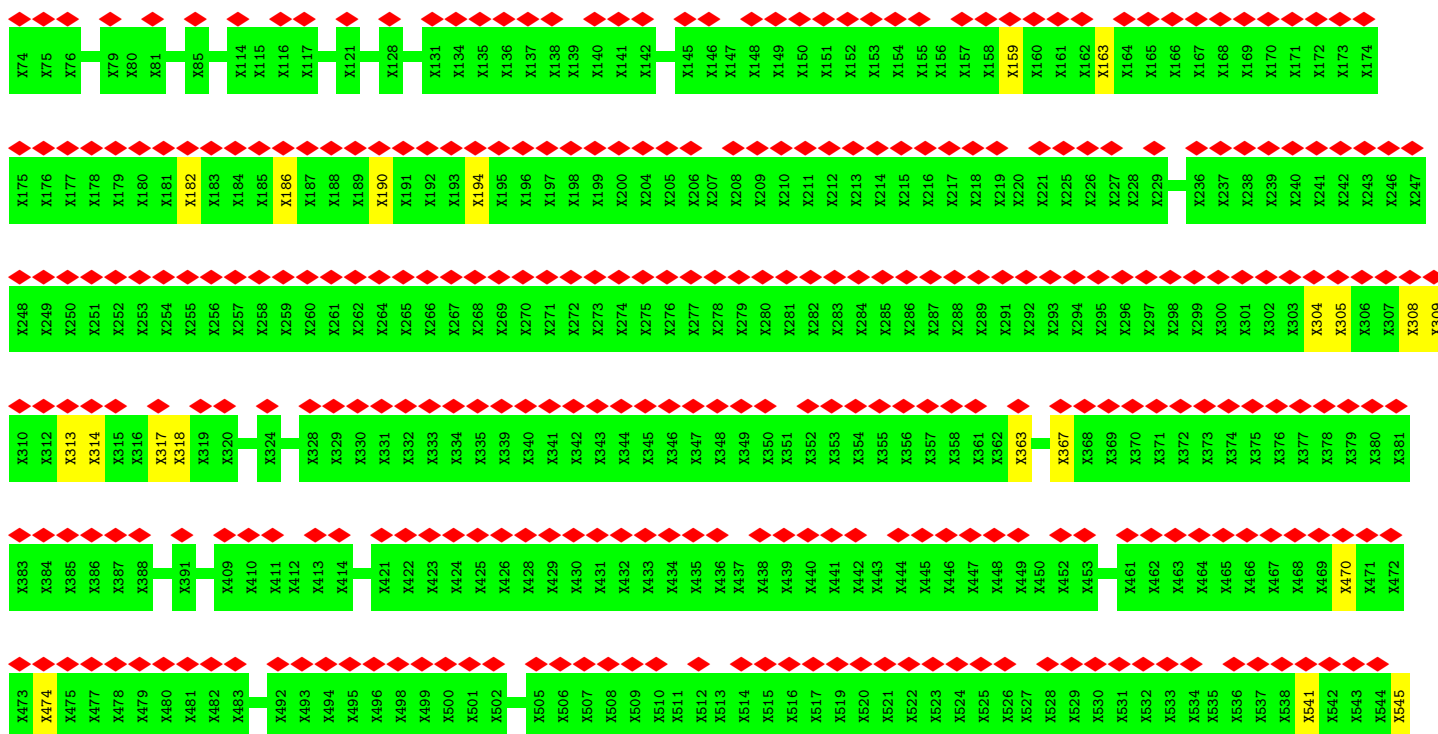
• Molecule 42: 40S ribosomal protein S19

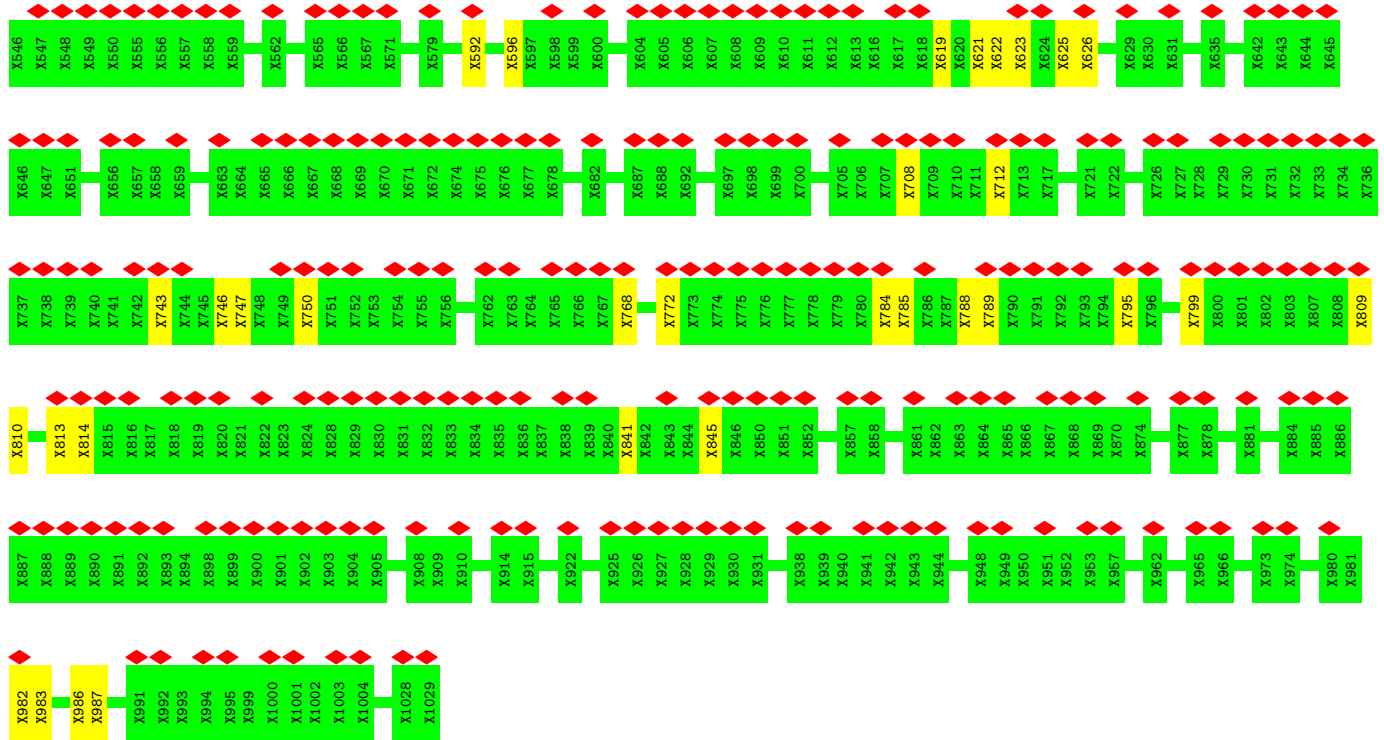


• Molecule 43: 40S ribosomal protein S27

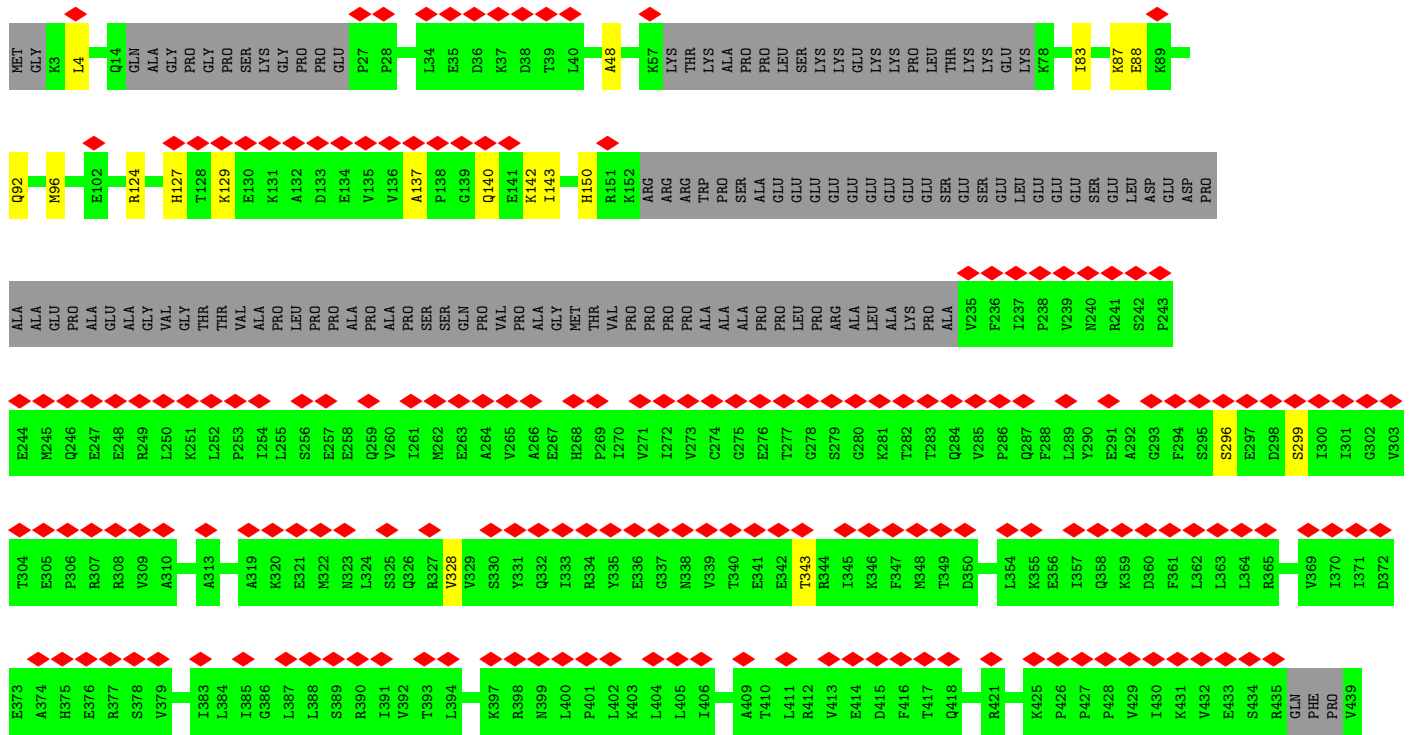
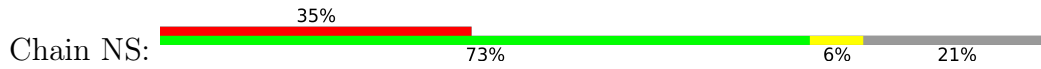


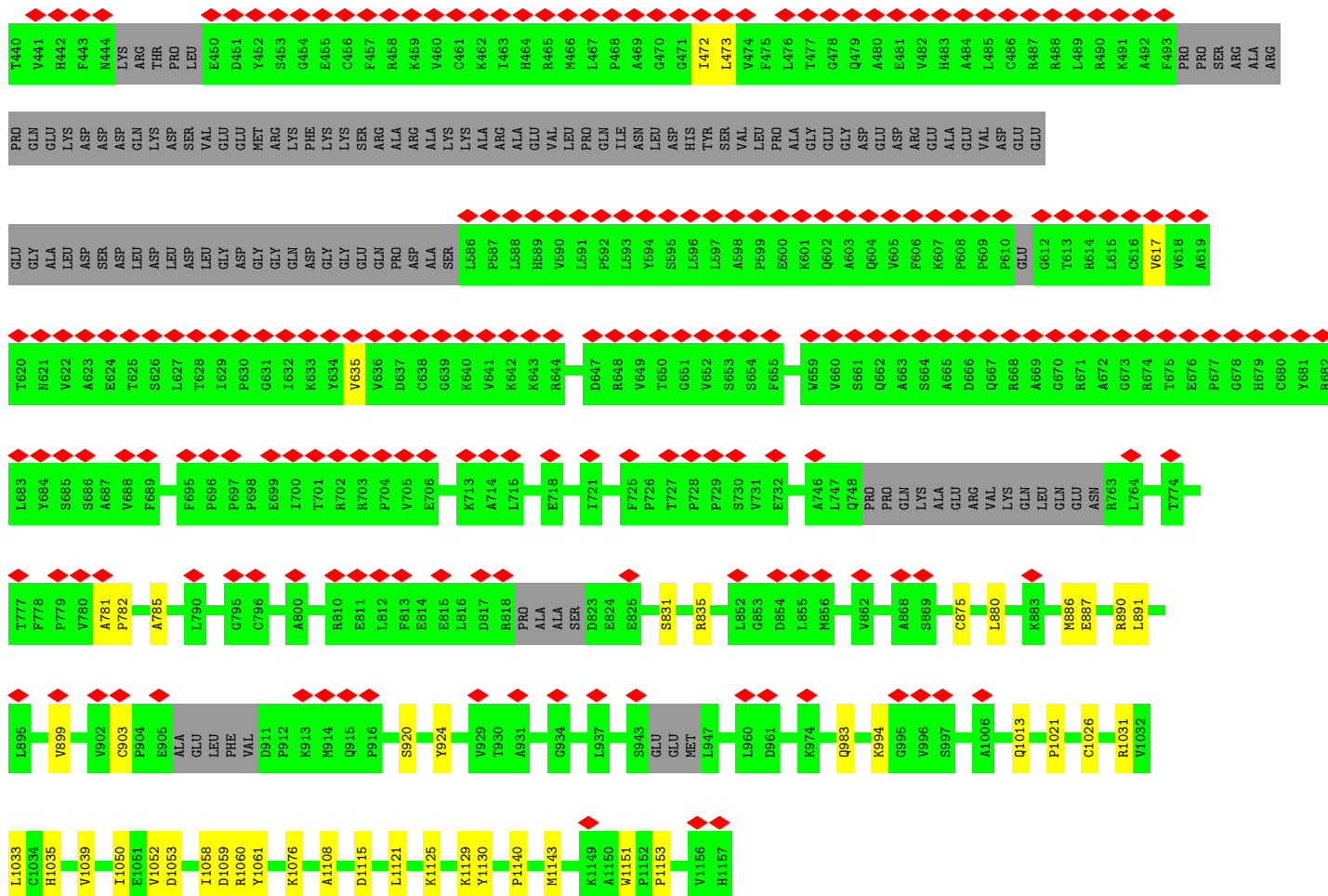
• Molecule 44: RRP12-like protein



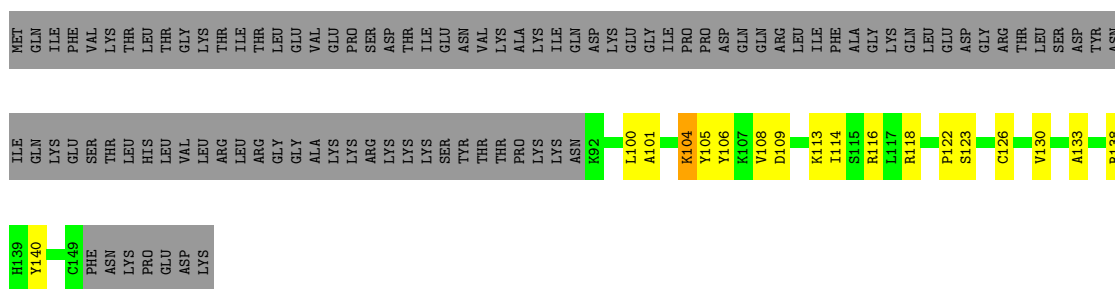


● Molecule 45: Probable ATP-dependent RNA helicase DHX37

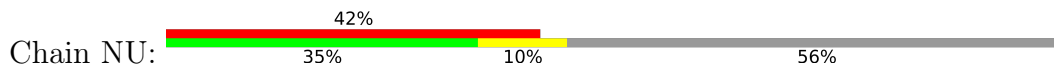




• Molecule 46: Ubiquitin-40S ribosomal protein S27a



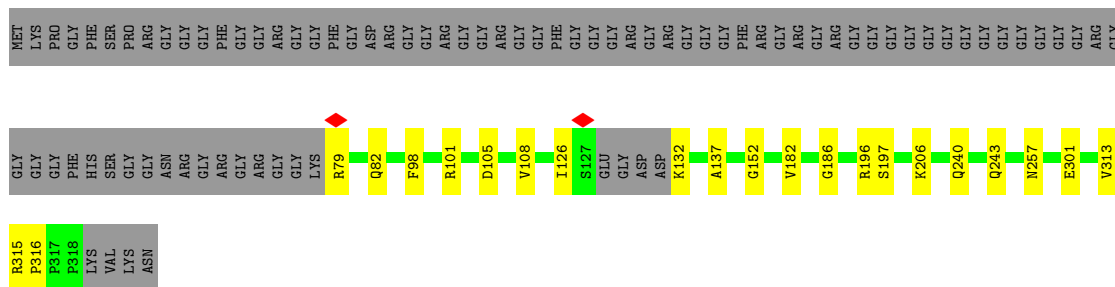
• Molecule 47: 40S ribosomal protein S17



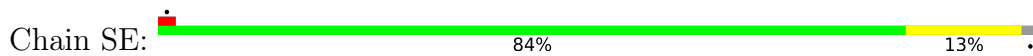




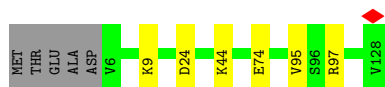




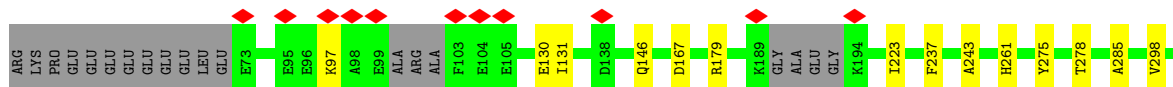
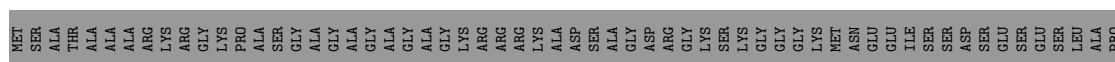
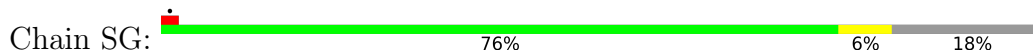
• Molecule 52: NHP2-like protein 1



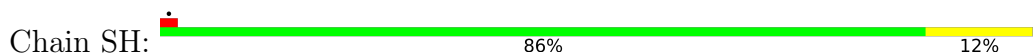
• Molecule 52: NHP2-like protein 1



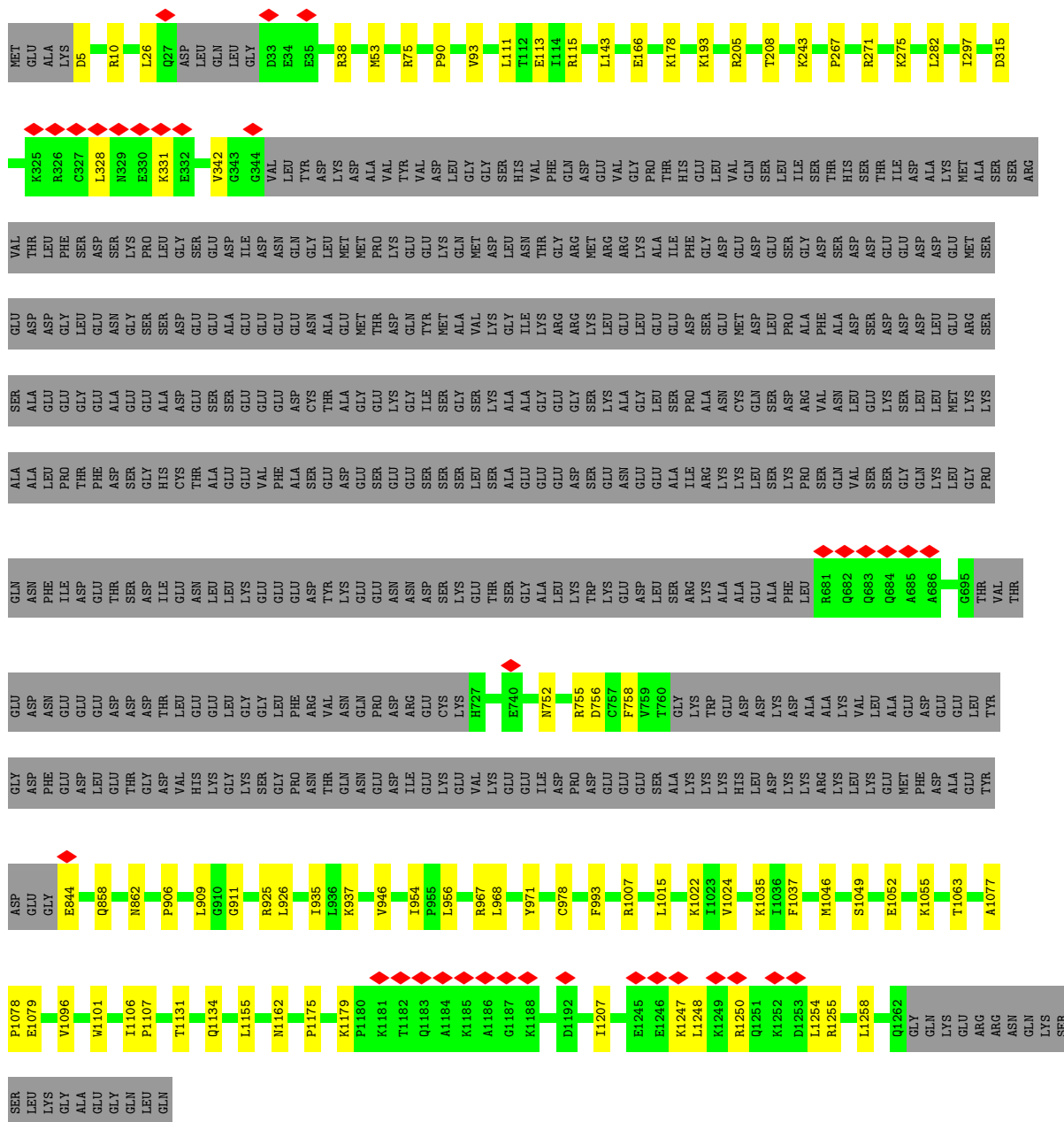
• Molecule 53: U3 small nucleolar RNA-interacting protein 2



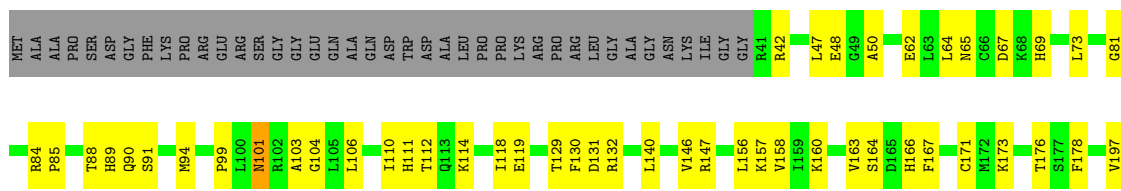
• Molecule 54: RNA 3'-terminal phosphate cyclase-like protein



• Molecule 55: Ribosome biogenesis protein BMS1 homolog

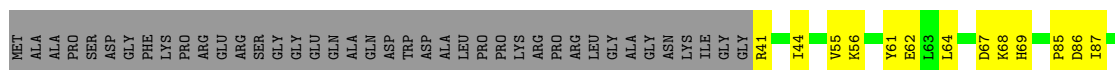


● Molecule 56: Ribosomal RNA small subunit methyltransferase NEP1

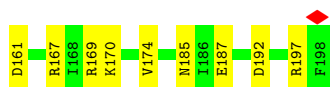
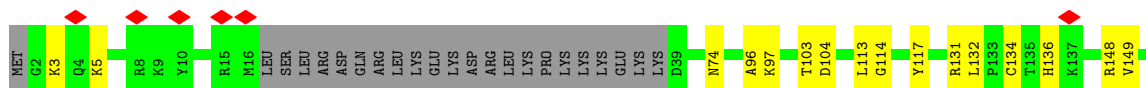
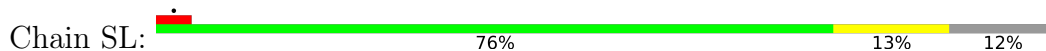




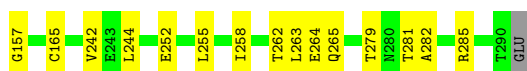
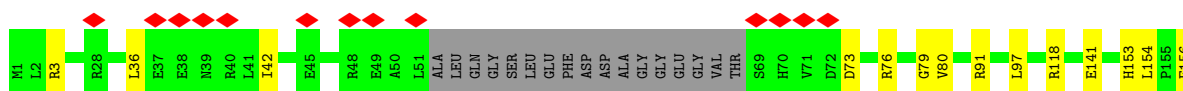
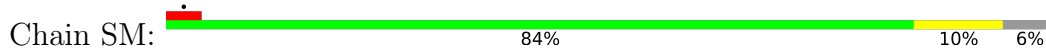
- Molecule 56: Ribosomal RNA small subunit methyltransferase NEP1



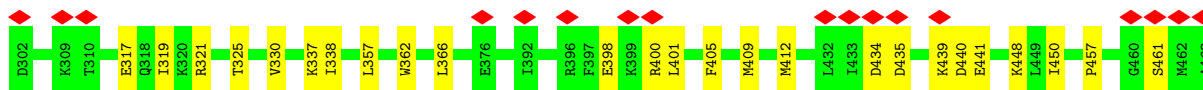
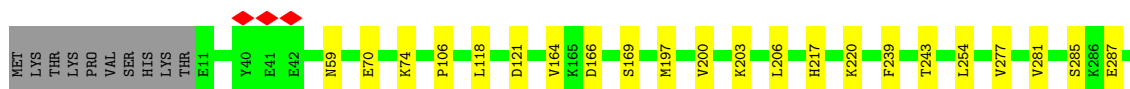
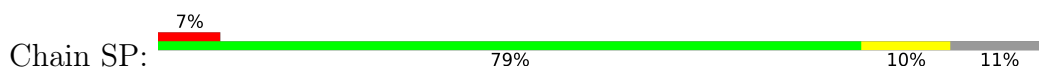
- Molecule 57: rRNA-processing protein FCF1 homolog

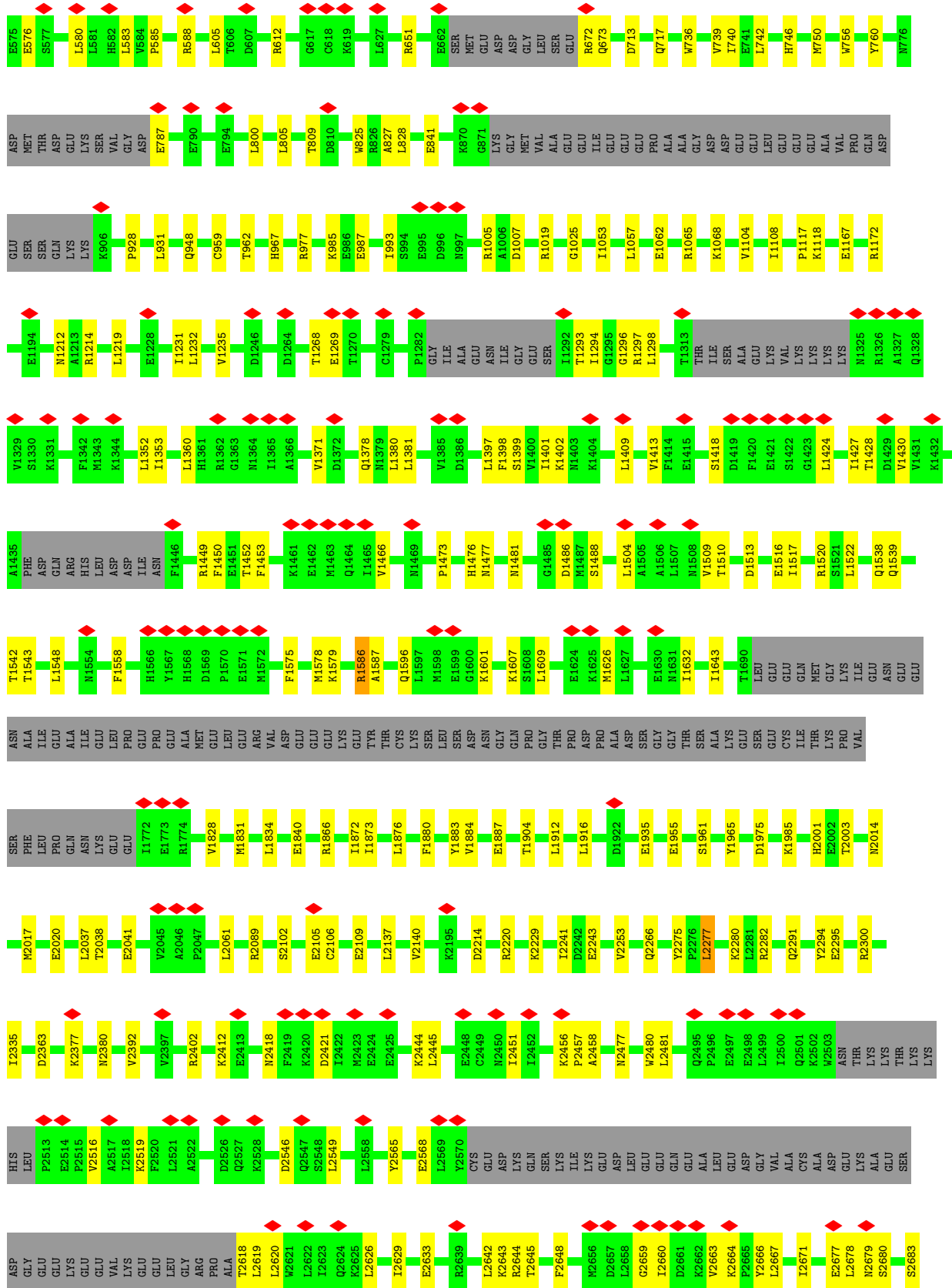


- Molecule 58: U3 small nucleolar ribonucleoprotein protein IMP4

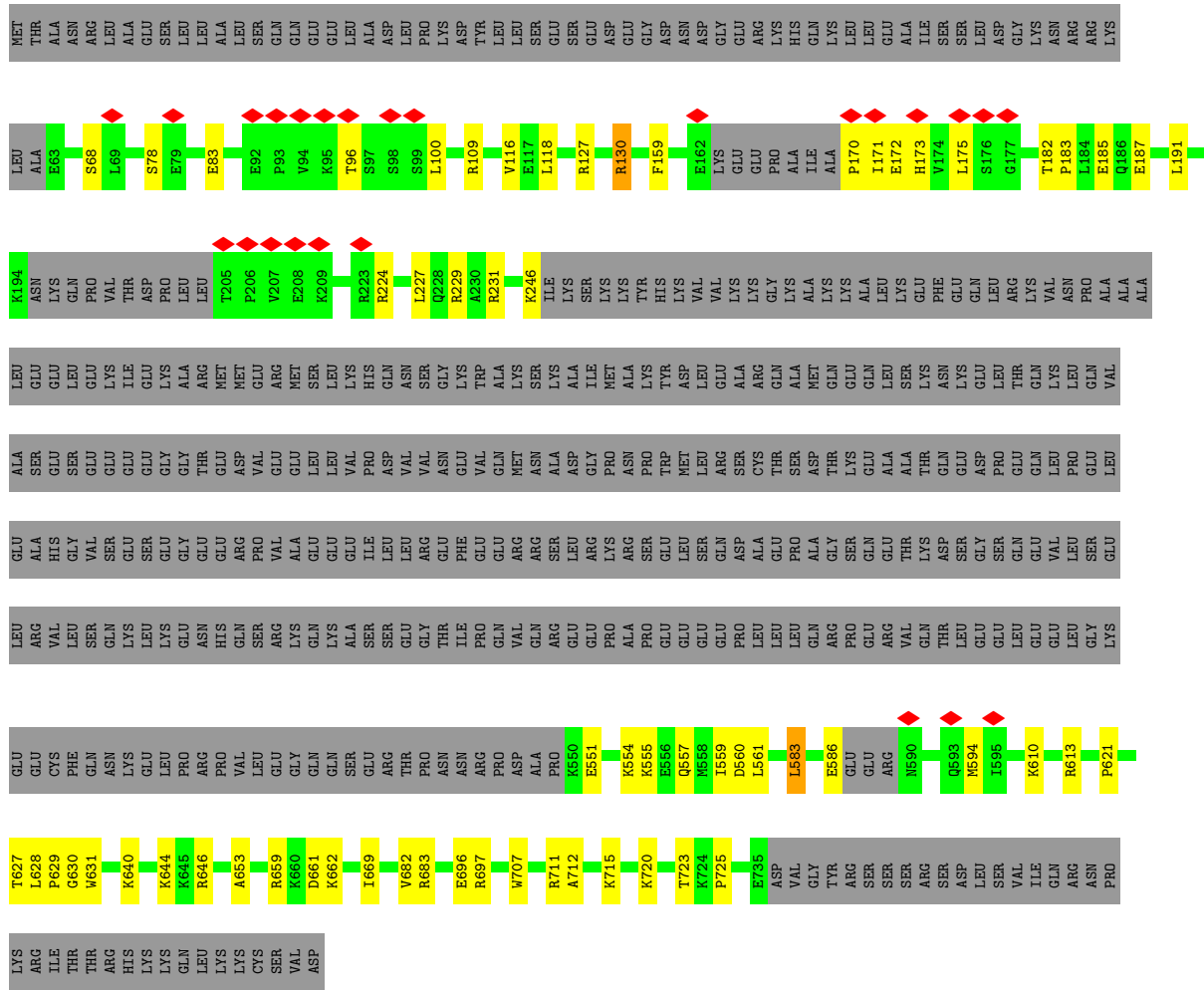


- Molecule 59: Small subunit processome component 20 homolog

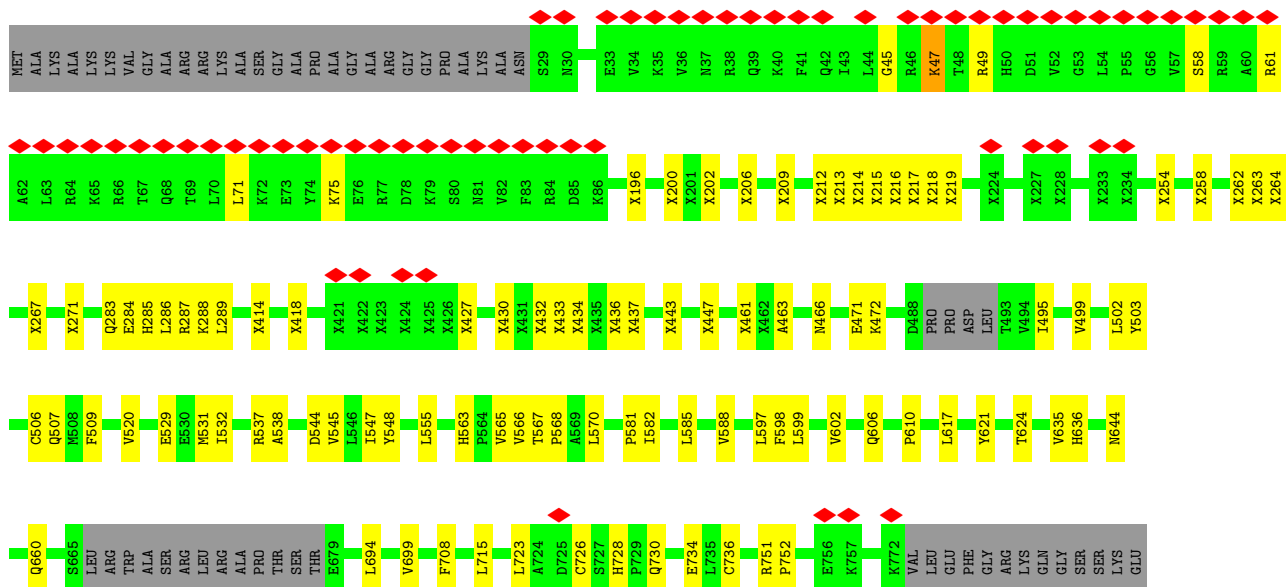
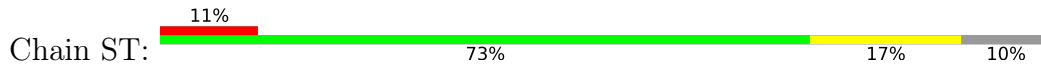


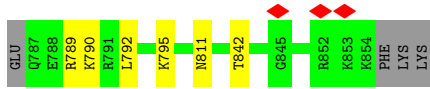




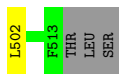
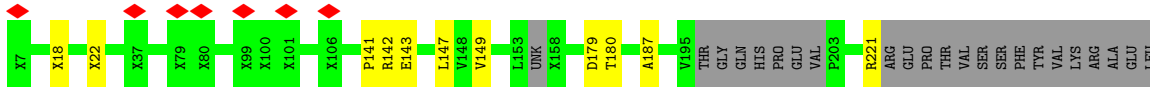
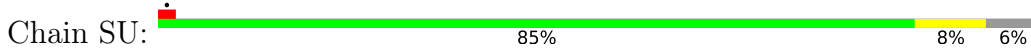


● Molecule 63: Nucleolar protein 14

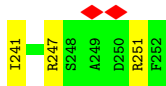
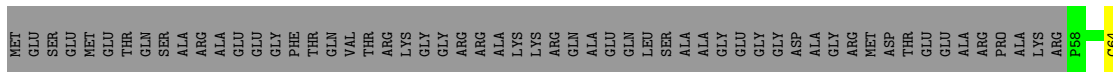




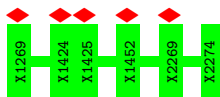
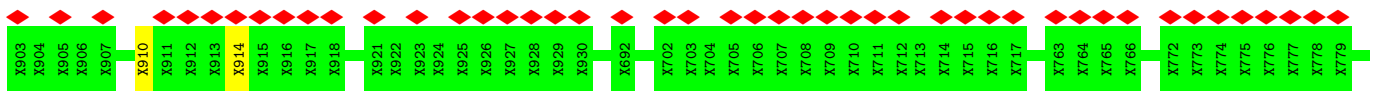
• Molecule 64: Nucleolar complex protein 4 homolog



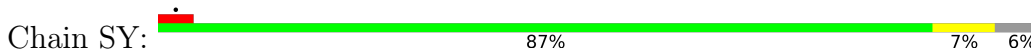
• Molecule 65: RNA-binding protein PNO1



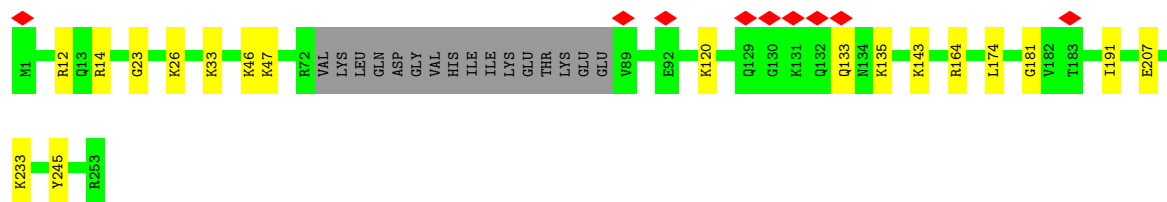
• Molecule 66: Unassigned peptides



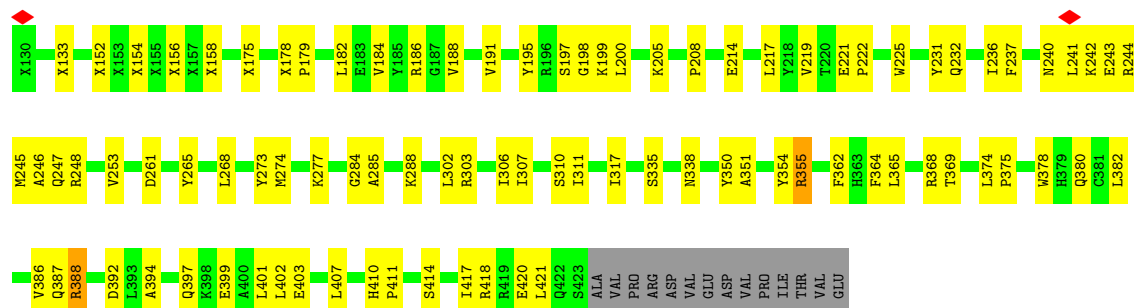
• Molecule 67: Probable U3 small nucleolar RNA-associated protein 11







• Molecule 68: Bystin



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	459775	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 ( $e^-/\text{\AA}^2$ )	58	Depositor
Minimum defocus (nm)	700	Depositor
Maximum defocus (nm)	2700	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	29.079	Depositor
Minimum map value	-20.930	Depositor
Average map value	0.036	Depositor
Map value standard deviation	0.793	Depositor
Recommended contour level	3.28	Depositor
Map size ( $\text{\AA}$ )	604.80005, 604.80005, 604.80005	wwPDB
Map dimensions	560, 560, 560	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.08, 1.08, 1.08	Depositor

## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: SAH, ZN, MG, OMG, OMU, 4AC, OMC, A2M, ATP, PSU, GTP

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	L0	0.25	0/388	0.70	0/604
2	L1	0.31	0/37548	0.84	72/58496 (0.1%)
3	L2	0.26	0/4222	0.81	6/6575 (0.1%)
4	L3	0.25	0/720	0.58	0/970
5	L4	0.26	0/2118	0.54	0/2849
6	L5	0.26	0/1523	0.53	0/2048
7	L6	0.26	0/1885	0.56	0/2510
8	L7	0.25	0/1365	0.49	0/1830
9	L8	0.26	0/1500	0.57	0/2002
10	L9	0.26	0/1524	0.56	0/2035
11	LA	0.37	0/941	0.72	0/1264
12	LC	0.29	0/1084	0.57	0/1453
13	LD	0.27	0/1250	0.54	0/1673
14	LF	0.27	0/1047	0.55	0/1391
15	LG	0.24	0/490	0.60	0/656
16	LH	0.26	0/6127	0.51	0/8292
17	LI	0.27	0/3245	0.49	0/4386
18	LJ	0.27	0/3734	0.54	0/5054
19	LK	0.27	0/959	0.54	0/1302
19	LL	0.25	0/4072	0.51	1/5539 (0.0%)
20	LM	0.26	0/16395	0.48	2/22212 (0.0%)
21	LN	0.26	0/5438	0.51	0/7377
22	LO	0.32	0/6607	0.54	0/8948
23	LP	0.25	0/4797	0.46	0/6443
24	LQ	0.25	0/6610	0.52	0/8915
25	LR	0.25	0/6132	0.53	0/8337
26	LS	0.26	0/3621	0.51	0/4876
27	LT	0.26	0/6795	0.49	0/9213
28	LU	0.26	0/3695	0.53	0/4986
29	LW	0.27	0/3592	0.53	0/4856
30	LY	0.25	0/785	0.55	0/1043
31	LZ	0.26	0/1368	0.57	0/1844

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
32	NA	0.25	0/2433	0.46	0/3248
33	NB	0.26	0/860	0.60	0/1139
34	ND	0.25	0/697	0.53	0/934
35	NF	0.24	0/1226	0.48	0/1649
36	NG	0.26	0/969	0.59	0/1298
37	NH	0.25	0/8580	0.52	1/11685 (0.0%)
38	NI	0.25	0/1887	0.50	0/2558
39	NL	0.25	0/2261	0.46	0/3061
40	NM	0.24	0/1899	0.49	0/2533
41	NO	0.26	0/1051	0.53	0/1406
42	NP	0.23	0/1142	0.48	0/1530
43	NQ	0.26	0/653	0.53	0/876
45	NS	0.26	0/5371	0.49	0/7371
46	NT	0.37	0/479	0.69	2/635 (0.3%)
47	NU	0.24	0/501	0.52	0/667
48	NV	0.27	0/354	0.42	0/474
49	SA	0.25	0/3122	0.46	0/4208
50	SB	0.25	0/3422	0.46	0/4602
51	SC	0.26	0/1818	0.51	0/2463
51	SD	0.26	0/1869	0.51	0/2529
52	SE	0.25	0/980	0.48	0/1323
52	SF	0.25	0/967	0.49	0/1305
53	SG	0.25	0/3101	0.53	0/4186
54	SH	0.25	0/2882	0.51	1/3887 (0.0%)
55	SI	0.26	0/6638	0.50	1/8919 (0.0%)
56	SJ	0.30	0/1609	0.57	0/2181
56	SK	0.28	0/1609	0.53	0/2181
57	SL	0.27	0/1464	0.54	0/1969
58	SM	0.25	0/2302	0.56	0/3103
59	SP	0.25	0/20427	0.46	2/27617 (0.0%)
60	SQ	0.27	0/1295	0.51	0/1723
61	SR	0.26	0/1116	0.52	0/1490
62	SS	0.28	0/2875	0.54	1/3867 (0.0%)
63	ST	0.28	0/3615	0.52	0/4858
64	SU	0.27	0/2634	0.50	1/3581 (0.0%)
65	SW	0.29	0/1549	0.54	0/2088
67	SY	0.25	0/2041	0.50	0/2709
68	SZ	0.35	0/2043	0.63	1/2763 (0.0%)
All	All	0.27	0/241318	0.59	91/334595 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a

sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
6	L5	0	1
37	NH	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 91 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L1	501	C	C2-N1-C1'	10.69	130.56	118.80
2	L1	501	C	N1-C2-O2	9.84	124.80	118.90
2	L1	1453	C	N1-C2-O2	9.67	124.70	118.90
2	L1	733	C	N1-C2-O2	9.46	124.57	118.90
2	L1	1453	C	C2-N1-C1'	9.41	129.15	118.80

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	L5	135	ARG	Peptide
37	NH	922	ASN	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	L0	347	0	173	1	0
2	L1	34979	0	17743	278	0
3	L2	3779	0	1901	21	0
4	L3	854	0	773	5	0
5	L4	2076	0	2177	5	0
6	L5	1501	0	1557	21	0
7	L6	1862	0	2018	18	0
8	L7	1346	0	1410	20	0
9	L8	1474	0	1542	11	0
10	L9	1499	0	1618	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	LA	931	0	961	51	0
12	LC	1068	0	1138	9	0
13	LD	1229	0	1302	7	0
14	LF	1030	0	1100	7	0
15	LG	488	0	514	6	0
16	LH	5987	0	5953	62	0
17	LI	3889	0	3466	54	0
18	LJ	3658	0	3695	50	0
19	LK	943	0	1023	21	0
19	LL	3982	0	4031	54	0
20	LM	16094	0	16586	174	0
21	LN	5299	0	5269	56	0
22	LO	6454	0	6342	96	0
23	LP	4696	0	4712	35	0
24	LQ	6497	0	6497	88	0
25	LR	6006	0	5968	111	0
26	LS	3560	0	3570	41	0
27	LT	6645	0	6653	82	0
28	LU	3611	0	3618	41	0
29	LW	3519	0	3532	38	0
30	LY	778	0	794	14	0
31	LZ	1344	0	1360	11	0
32	NA	2404	0	2518	26	0
33	NB	849	0	895	11	0
34	ND	685	0	711	7	0
35	NF	1202	0	1289	7	0
36	NG	957	0	982	15	0
37	NH	8374	0	8456	73	0
38	NI	1840	0	1812	22	0
39	NL	2213	0	2312	22	0
40	NM	1873	0	1968	16	0
41	NO	1034	0	1080	12	0
42	NP	1122	0	1153	8	0
43	NQ	640	0	661	4	0
44	NR	4305	0	945	27	0
45	NS	5332	0	3593	38	0
46	NT	470	0	473	16	0
47	NU	495	0	541	8	0
48	NV	344	0	301	2	0
49	SA	3077	0	3139	32	0
50	SB	3371	0	3483	18	0
51	SC	1781	0	1803	12	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
51	SD	1832	0	1854	17	0
52	SE	968	0	1017	10	0
52	SF	955	0	1008	5	0
53	SG	3042	0	3030	20	0
54	SH	2832	0	2937	28	0
55	SI	6493	0	6679	55	0
56	SJ	1579	0	1646	42	0
56	SK	1579	0	1646	24	0
57	SL	1433	0	1463	17	0
58	SM	2253	0	2273	25	0
59	SP	20032	0	20757	173	0
60	SQ	1275	0	1316	10	0
61	SR	1098	0	1167	4	0
62	SS	2824	0	2932	45	0
63	ST	4188	0	3831	66	0
64	SU	3154	0	2736	26	0
65	SW	1522	0	1602	32	0
66	SX	1140	0	249	1	0
67	SY	2013	0	2136	18	0
68	SZ	2222	0	2142	61	0
69	L1	61	0	0	0	0
69	NH	1	0	0	0	0
69	SI	1	0	0	0	0
69	SL	1	0	0	0	0
70	NH	31	0	12	1	0
71	NQ	1	0	0	0	0
71	NT	1	0	0	0	0
71	SL	1	0	0	0	0
72	SI	32	0	12	0	0
73	SJ	26	0	19	5	0
73	SK	26	0	19	5	0
All	All	242409	0	219594	2260	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 2260 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:L1:1396:A:O2'	2:L1:1398:G:N7	2.14	0.80
17:LI:608:ILE:HG23	17:LI:613:ILE:HD11	1.63	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
25:LR:49:ILE:HG12	25:LR:95:TRP:HE1	1.49	0.78
37:NH:716:GLU:HG2	37:NH:717:ARG:HD2	1.65	0.78
49:SA:154:ARG:HH21	51:SD:206:LYS:HE2	1.48	0.78

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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	
4	L3	85/116 (73%)	83 (98%)	2 (2%)	0	100	100
5	L4	260/263 (99%)	256 (98%)	4 (2%)	0	100	100
6	L5	188/204 (92%)	182 (97%)	6 (3%)	0	100	100
7	L6	228/249 (92%)	224 (98%)	4 (2%)	0	100	100
8	L7	164/194 (84%)	162 (99%)	2 (1%)	0	100	100
9	L8	176/208 (85%)	173 (98%)	3 (2%)	0	100	100
10	L9	178/194 (92%)	176 (99%)	2 (1%)	0	100	100
11	LA	118/132 (89%)	111 (94%)	7 (6%)	0	100	100
12	LC	134/146 (92%)	133 (99%)	1 (1%)	0	100	100
13	LD	149/158 (94%)	145 (97%)	4 (3%)	0	100	100
14	LF	124/133 (93%)	124 (100%)	0	0	100	100
15	LG	60/69 (87%)	58 (97%)	2 (3%)	0	100	100
16	LH	738/830 (89%)	720 (98%)	18 (2%)	0	100	100
17	LI	374/678 (55%)	370 (99%)	4 (1%)	0	100	100
18	LJ	456/518 (88%)	440 (96%)	16 (4%)	0	100	100
19	LK	116/677 (17%)	113 (97%)	3 (3%)	0	100	100
19	LL	500/677 (74%)	480 (96%)	20 (4%)	0	100	100

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	LM	2021/2144 (94%)	1980 (98%)	41 (2%)	0	100	100
21	LN	667/686 (97%)	646 (97%)	21 (3%)	0	100	100
22	LO	813/919 (88%)	798 (98%)	15 (2%)	0	100	100
23	LP	558/597 (94%)	553 (99%)	5 (1%)	0	100	100
24	LQ	814/943 (86%)	797 (98%)	17 (2%)	0	100	100
25	LR	768/808 (95%)	751 (98%)	17 (2%)	0	100	100
26	LS	447/556 (80%)	435 (97%)	12 (3%)	0	100	100
27	LT	847/951 (89%)	831 (98%)	16 (2%)	0	100	100
28	LU	443/445 (100%)	432 (98%)	11 (2%)	0	100	100
29	LW	448/610 (73%)	436 (97%)	12 (3%)	0	100	100
30	LY	89/136 (65%)	89 (100%)	0	0	100	100
31	LZ	158/184 (86%)	155 (98%)	3 (2%)	0	100	100
32	NA	287/681 (42%)	285 (99%)	2 (1%)	0	100	100
33	NB	99/479 (21%)	97 (98%)	2 (2%)	0	100	100
34	ND	81/257 (32%)	80 (99%)	1 (1%)	0	100	100
35	NF	147/151 (97%)	145 (99%)	2 (1%)	0	100	100
36	NG	125/151 (83%)	122 (98%)	3 (2%)	0	100	100
37	NH	1064/1146 (93%)	1038 (98%)	26 (2%)	0	100	100
38	NI	232/280 (83%)	229 (99%)	3 (1%)	0	100	100
39	NL	276/313 (88%)	274 (99%)	2 (1%)	0	100	100
40	NM	229/264 (87%)	225 (98%)	4 (2%)	0	100	100
41	NO	127/130 (98%)	125 (98%)	2 (2%)	0	100	100
42	NP	142/145 (98%)	141 (99%)	1 (1%)	0	100	100
43	NQ	80/84 (95%)	76 (95%)	4 (5%)	0	100	100
45	NS	890/1157 (77%)	860 (97%)	30 (3%)	0	100	100
46	NT	56/156 (36%)	54 (96%)	2 (4%)	0	100	100
47	NU	58/135 (43%)	58 (100%)	0	0	100	100
48	NV	34/885 (4%)	34 (100%)	0	0	100	100
49	SA	390/594 (66%)	386 (99%)	4 (1%)	0	100	100
50	SB	426/529 (80%)	424 (100%)	2 (0%)	0	100	100
51	SC	225/321 (70%)	220 (98%)	5 (2%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
51	SD	232/321 (72%)	226 (97%)	6 (3%)	0	100	100
52	SE	123/128 (96%)	122 (99%)	1 (1%)	0	100	100
52	SF	121/128 (94%)	120 (99%)	1 (1%)	0	100	100
53	SG	383/475 (81%)	372 (97%)	11 (3%)	0	100	100
54	SH	366/373 (98%)	357 (98%)	9 (2%)	0	100	100
55	SI	793/1282 (62%)	779 (98%)	14 (2%)	0	100	100
56	SJ	202/244 (83%)	194 (96%)	8 (4%)	0	100	100
56	SK	202/244 (83%)	198 (98%)	4 (2%)	0	100	100
57	SL	171/198 (86%)	167 (98%)	4 (2%)	0	100	100
58	SM	269/291 (92%)	263 (98%)	6 (2%)	0	100	100
59	SP	2463/2785 (88%)	2426 (98%)	37 (2%)	0	100	100
60	SQ	150/756 (20%)	148 (99%)	2 (1%)	0	100	100
61	SR	139/143 (97%)	135 (97%)	4 (3%)	0	100	100
62	SS	340/771 (44%)	335 (98%)	5 (2%)	0	100	100
63	ST	434/632 (69%)	424 (98%)	10 (2%)	0	100	100
64	SU	316/472 (67%)	307 (97%)	9 (3%)	0	100	100
65	SW	193/252 (77%)	192 (100%)	1 (0%)	0	100	100
67	SY	233/253 (92%)	231 (99%)	2 (1%)	0	100	100
68	SZ	244/304 (80%)	234 (96%)	10 (4%)	0	100	100
All	All	24463/32335 (76%)	23956 (98%)	507 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	
4	L3	76/77 (99%)	75 (99%)	1 (1%)	69	87
5	L4	224/225 (100%)	224 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	L5	160/170 (94%)	158 (99%)	2 (1%)	69	87
7	L6	200/218 (92%)	199 (100%)	1 (0%)	88	96
8	L7	149/174 (86%)	149 (100%)	0	100	100
9	L8	155/180 (86%)	155 (100%)	0	100	100
10	L9	160/168 (95%)	158 (99%)	2 (1%)	69	87
11	LA	102/108 (94%)	101 (99%)	1 (1%)	76	91
12	LC	111/121 (92%)	111 (100%)	0	100	100
13	LD	135/142 (95%)	133 (98%)	2 (2%)	65	86
14	LF	110/115 (96%)	110 (100%)	0	100	100
15	LG	55/62 (89%)	55 (100%)	0	100	100
16	LH	670/748 (90%)	668 (100%)	2 (0%)	92	98
17	LI	368/497 (74%)	368 (100%)	0	100	100
18	LJ	406/456 (89%)	405 (100%)	1 (0%)	93	98
19	LK	112/594 (19%)	111 (99%)	1 (1%)	78	92
19	LL	456/594 (77%)	454 (100%)	2 (0%)	91	97
20	LM	1813/1943 (93%)	1810 (100%)	3 (0%)	93	98
21	LN	582/597 (98%)	581 (100%)	1 (0%)	93	98
22	LO	705/783 (90%)	701 (99%)	4 (1%)	86	95
23	LP	498/527 (94%)	496 (100%)	2 (0%)	91	97
24	LQ	701/828 (85%)	700 (100%)	1 (0%)	93	98
25	LR	647/672 (96%)	645 (100%)	2 (0%)	92	98
26	LS	393/476 (83%)	391 (100%)	2 (0%)	88	96
27	LT	733/823 (89%)	732 (100%)	1 (0%)	93	98
28	LU	399/399 (100%)	398 (100%)	1 (0%)	92	98
29	LW	372/512 (73%)	372 (100%)	0	100	100
30	LY	83/114 (73%)	83 (100%)	0	100	100
31	LZ	146/167 (87%)	146 (100%)	0	100	100
32	NA	270/626 (43%)	269 (100%)	1 (0%)	91	97
33	NB	87/413 (21%)	87 (100%)	0	100	100
34	ND	71/222 (32%)	71 (100%)	0	100	100
35	NF	130/131 (99%)	130 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
36	NG	100/119 (84%)	100 (100%)	0	100	100
37	NH	917/984 (93%)	916 (100%)	1 (0%)	93	98
38	NI	193/246 (78%)	193 (100%)	0	100	100
39	NL	250/276 (91%)	249 (100%)	1 (0%)	91	97
40	NM	207/231 (90%)	207 (100%)	0	100	100
41	NO	112/113 (99%)	112 (100%)	0	100	100
42	NP	114/115 (99%)	114 (100%)	0	100	100
43	NQ	74/76 (97%)	73 (99%)	1 (1%)	67	86
45	NS	224/992 (23%)	223 (100%)	1 (0%)	91	97
46	NT	51/140 (36%)	50 (98%)	1 (2%)	55	81
47	NU	53/122 (43%)	52 (98%)	1 (2%)	57	82
48	NV	38/784 (5%)	38 (100%)	0	100	100
49	SA	334/511 (65%)	334 (100%)	0	100	100
50	SB	365/455 (80%)	365 (100%)	0	100	100
51	SC	192/234 (82%)	192 (100%)	0	100	100
51	SD	197/234 (84%)	197 (100%)	0	100	100
52	SE	108/111 (97%)	108 (100%)	0	100	100
52	SF	107/111 (96%)	107 (100%)	0	100	100
53	SG	327/382 (86%)	327 (100%)	0	100	100
54	SH	315/318 (99%)	315 (100%)	0	100	100
55	SI	702/1119 (63%)	702 (100%)	0	100	100
56	SJ	181/209 (87%)	178 (98%)	3 (2%)	60	84
56	SK	181/209 (87%)	181 (100%)	0	100	100
57	SL	159/182 (87%)	158 (99%)	1 (1%)	86	95
58	SM	242/254 (95%)	242 (100%)	0	100	100
59	SP	2260/2522 (90%)	2258 (100%)	2 (0%)	93	98
60	SQ	135/676 (20%)	135 (100%)	0	100	100
61	SR	113/115 (98%)	113 (100%)	0	100	100
62	SS	311/686 (45%)	306 (98%)	5 (2%)	62	85
63	ST	391/439 (89%)	388 (99%)	3 (1%)	81	93
64	SU	269/303 (89%)	269 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
65	SW	165/208 (79%)	164 (99%)	1 (1%)	86	95
67	SY	217/232 (94%)	217 (100%)	0	100	100
68	SZ	214/227 (94%)	212 (99%)	2 (1%)	78	92
All	All	21097/27817 (76%)	21041 (100%)	56 (0%)	92	98

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
26	LS	89	ARG
68	SZ	388	ARG
43	NQ	16	LYS
68	SZ	355	ARG
62	SS	644	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
65	SW	110	ASN
65	SW	115	ASN
22	LO	879	GLN
22	LO	648	ASN
65	SW	232	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	L0	15/3617 (0%)	5 (33%)	0
2	L1	1611/1872 (86%)	393 (24%)	22 (1%)
3	L2	174/217 (80%)	50 (28%)	3 (1%)
All	All	1800/5706 (31%)	448 (24%)	25 (1%)

5 of 448 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	L0	432	A
1	L0	434	C
1	L0	436	A
1	L0	438	G
1	L0	439	U

5 of 25 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	L1	1461	G
2	L1	1637	A
3	L2	188	G
2	L1	1551	U
2	L1	1684	C

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

66 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
2	OMU	L1	121	2	19,22,23	3.30	7 (36%)	26,31,34	1.87	5 (19%)
2	PSU	L1	1056	2	18,21,22	3.47	4 (22%)	22,30,33	2.25	6 (27%)
2	PSU	L1	686	2	18,21,22	3.42	5 (27%)	22,30,33	2.17	6 (27%)
2	PSU	L1	572	2	18,21,22	3.37	4 (22%)	22,30,33	2.22	7 (31%)
2	4AC	L1	1842	2	21,24,25	2.00	6 (28%)	29,34,37	2.02	6 (20%)
2	PSU	L1	866	2	18,21,22	3.41	4 (22%)	22,30,33	2.24	6 (27%)
2	A2M	L1	576	2	18,25,26	4.26	7 (38%)	18,36,39	1.39	1 (5%)
2	PSU	L1	814	2	18,21,22	3.39	4 (22%)	22,30,33	2.23	6 (27%)
2	OMG	L1	867	2	18,26,27	3.32	5 (27%)	19,38,41	1.26	3 (15%)
2	OMG	L1	683	2	18,26,27	3.34	5 (27%)	19,38,41	1.23	3 (15%)
2	PSU	L1	1232	2	18,21,22	3.43	4 (22%)	22,30,33	2.26	6 (27%)
2	PSU	L1	406	2	18,21,22	3.37	5 (27%)	22,30,33	2.25	6 (27%)
2	PSU	L1	966	2	18,21,22	3.42	4 (22%)	22,30,33	2.19	6 (27%)
2	PSU	L1	681	2	18,21,22	3.41	5 (27%)	22,30,33	2.24	6 (27%)
2	PSU	L1	1045	2	18,21,22	3.42	4 (22%)	22,30,33	2.23	6 (27%)
2	A2M	L1	484	2	18,25,26	4.25	7 (38%)	18,36,39	1.73	3 (16%)
2	PSU	L1	1625	2	18,21,22	3.48	4 (22%)	22,30,33	2.23	6 (27%)
2	PSU	L1	1081	2	18,21,22	3.41	4 (22%)	22,30,33	2.23	6 (27%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	A2M	L1	166	2	18,25,26	4.33	7 (38%)	18,36,39	1.41	1 (5%)
2	OMU	L1	354	2	19,22,23	3.30	7 (36%)	26,31,34	1.92	5 (19%)
2	PSU	L1	119	2	18,21,22	3.40	4 (22%)	22,30,33	2.24	5 (22%)
2	A2M	L1	27	2	18,25,26	4.30	7 (38%)	18,36,39	1.42	2 (11%)
2	OMC	L1	621	2	19,22,23	1.00	2 (10%)	26,31,34	0.97	2 (7%)
2	PSU	L1	34	2	18,21,22	3.44	4 (22%)	22,30,33	2.25	6 (27%)
2	OMU	L1	428	2	19,22,23	3.27	7 (36%)	26,31,34	2.05	7 (26%)
2	OMG	L1	509	69,2	18,26,27	3.30	5 (27%)	19,38,41	1.26	3 (15%)
2	PSU	L1	815	2	18,21,22	3.39	4 (22%)	22,30,33	2.25	6 (27%)
2	A2M	L1	668	2,3	18,25,26	4.28	7 (38%)	18,36,39	1.72	2 (11%)
2	OMG	L1	436	2	18,26,27	3.31	5 (27%)	19,38,41	1.30	3 (15%)
2	A2M	L1	468	2	18,25,26	4.28	7 (38%)	18,36,39	1.42	2 (11%)
2	OMU	L1	627	2	19,22,23	3.28	7 (36%)	26,31,34	1.97	5 (19%)
2	PSU	L1	1174	2,3	18,21,22	3.46	4 (22%)	22,30,33	2.22	6 (27%)
2	PSU	L1	609	2	18,21,22	3.38	5 (27%)	22,30,33	2.22	6 (27%)
2	OMG	L1	601	2	18,26,27	3.33	5 (27%)	19,38,41	1.33	3 (15%)
2	PSU	L1	649	2	18,21,22	3.40	5 (27%)	22,30,33	2.21	6 (27%)
2	PSU	L1	1177	2,3	18,21,22	3.48	5 (27%)	22,30,33	2.28	6 (27%)
2	OMC	L1	174	69,2	19,22,23	0.91	1 (5%)	26,31,34	0.68	0
2	A2M	L1	1678	2	18,25,26	4.30	7 (38%)	18,36,39	1.44	2 (11%)
2	OMC	L1	797	2	19,22,23	1.01	2 (10%)	26,31,34	0.70	0
2	OMC	L1	1703	2	19,22,23	0.95	2 (10%)	26,31,34	0.66	0
2	A2M	L1	512	2	18,25,26	4.27	6 (33%)	18,36,39	1.47	2 (11%)
2	PSU	L1	210	2	18,21,22	3.48	4 (22%)	22,30,33	2.20	6 (27%)
2	A2M	L1	99	69,2	18,25,26	4.32	6 (33%)	18,36,39	1.39	1 (5%)
2	OMC	L1	462	2	19,22,23	0.91	1 (5%)	26,31,34	0.67	0
2	OMC	L1	517	69,2	19,22,23	0.99	2 (10%)	26,31,34	0.78	0
2	PSU	L1	218	2	18,21,22	3.38	4 (22%)	22,30,33	2.23	7 (31%)
2	A2M	L1	590	2	18,25,26	4.25	7 (38%)	18,36,39	1.43	1 (5%)
2	PSU	L1	651	2	18,21,22	3.42	4 (22%)	22,30,33	2.24	6 (27%)
2	PSU	L1	863	2	18,21,22	3.36	4 (22%)	22,30,33	2.24	6 (27%)
2	PSU	L1	1046	2	18,21,22	3.41	4 (22%)	22,30,33	2.26	6 (27%)
2	PSU	L1	918	2	18,21,22	3.42	5 (27%)	22,30,33	2.35	7 (31%)
2	PSU	L1	1643	2	18,21,22	3.43	4 (22%)	22,30,33	2.22	6 (27%)
2	PSU	L1	36	2	18,21,22	3.37	5 (27%)	22,30,33	2.22	6 (27%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	PSU	L1	105	2	18,21,22	3.41	4 (22%)	22,30,33	2.25	6 (27%)
2	OMG	L1	644	2	18,26,27	3.31	5 (27%)	19,38,41	1.31	3 (15%)
2	PSU	L1	1004	2	18,21,22	3.42	5 (27%)	22,30,33	2.27	6 (27%)
2	OMU	L1	116	2	19,22,23	3.30	7 (36%)	26,31,34	1.89	5 (19%)
2	PSU	L1	296	2	18,21,22	3.42	4 (22%)	22,30,33	2.26	6 (27%)
2	PSU	L1	93	2	18,21,22	3.39	4 (22%)	22,30,33	2.23	6 (27%)
2	PSU	L1	801	2	18,21,22	3.44	5 (27%)	22,30,33	2.20	6 (27%)
2	A2M	L1	159	2	18,25,26	4.26	7 (38%)	18,36,39	1.46	2 (11%)
2	OMU	L1	172	2	19,22,23	3.32	7 (36%)	26,31,34	1.98	5 (19%)
2	A2M	L1	1031	2	18,25,26	4.28	7 (38%)	18,36,39	1.39	2 (11%)
2	PSU	L1	109	2	18,21,22	3.42	4 (22%)	22,30,33	2.28	6 (27%)
2	PSU	L1	822	2	18,21,22	3.43	5 (27%)	22,30,33	2.26	6 (27%)
2	OMU	L1	799	2	19,22,23	3.29	7 (36%)	26,31,34	1.87	6 (23%)

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
2	OMU	L1	121	2	-	0/9/27/28	0/2/2/2
2	PSU	L1	1056	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	686	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	572	2	-	0/7/25/26	0/2/2/2
2	4AC	L1	1842	2	-	2/11/29/30	0/2/2/2
2	PSU	L1	866	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	576	2	-	2/5/27/28	0/3/3/3
2	PSU	L1	814	2	-	0/7/25/26	0/2/2/2
2	OMG	L1	867	2	-	0/5/27/28	0/3/3/3
2	OMG	L1	683	2	-	0/5/27/28	0/3/3/3
2	PSU	L1	1232	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	406	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	966	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	681	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	1045	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	484	2	-	1/5/27/28	0/3/3/3
2	PSU	L1	1625	2	-	2/7/25/26	0/2/2/2
2	PSU	L1	1081	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	166	2	-	0/5/27/28	0/3/3/3

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	OMU	L1	354	2	-	0/9/27/28	0/2/2/2
2	PSU	L1	119	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	27	2	-	1/5/27/28	0/3/3/3
2	OMC	L1	621	2	-	3/9/27/28	0/2/2/2
2	PSU	L1	34	2	-	0/7/25/26	0/2/2/2
2	OMU	L1	428	2	-	6/9/27/28	0/2/2/2
2	OMG	L1	509	69,2	-	2/5/27/28	0/3/3/3
2	PSU	L1	815	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	668	2,3	-	2/5/27/28	0/3/3/3
2	OMG	L1	436	2	-	0/5/27/28	0/3/3/3
2	A2M	L1	468	2	-	1/5/27/28	0/3/3/3
2	OMU	L1	627	2	-	4/9/27/28	0/2/2/2
2	PSU	L1	1174	2,3	-	0/7/25/26	0/2/2/2
2	PSU	L1	609	2	-	0/7/25/26	0/2/2/2
2	OMG	L1	601	2	-	2/5/27/28	0/3/3/3
2	PSU	L1	649	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	1177	2,3	-	3/7/25/26	0/2/2/2
2	OMC	L1	174	69,2	-	0/9/27/28	0/2/2/2
2	A2M	L1	1678	2	-	1/5/27/28	0/3/3/3
2	OMC	L1	797	2	-	0/9/27/28	0/2/2/2
2	OMC	L1	1703	2	-	5/9/27/28	0/2/2/2
2	A2M	L1	512	2	-	0/5/27/28	0/3/3/3
2	PSU	L1	210	2	-	2/7/25/26	0/2/2/2
2	A2M	L1	99	69,2	-	2/5/27/28	0/3/3/3
2	OMC	L1	462	2	-	0/9/27/28	0/2/2/2
2	OMC	L1	517	69,2	-	2/9/27/28	0/2/2/2
2	PSU	L1	218	2	-	0/7/25/26	0/2/2/2
2	A2M	L1	590	2	-	2/5/27/28	0/3/3/3
2	PSU	L1	651	2	-	1/7/25/26	0/2/2/2
2	PSU	L1	863	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	1046	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	918	2	-	1/7/25/26	0/2/2/2
2	PSU	L1	1643	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	36	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	105	2	-	0/7/25/26	0/2/2/2
2	OMG	L1	644	2	-	4/5/27/28	0/3/3/3
2	PSU	L1	1004	2	-	0/7/25/26	0/2/2/2
2	OMU	L1	116	2	-	0/9/27/28	0/2/2/2
2	PSU	L1	296	2	-	0/7/25/26	0/2/2/2

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PSU	L1	93	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	801	2	-	1/7/25/26	0/2/2/2
2	A2M	L1	159	2	-	1/5/27/28	0/3/3/3
2	OMU	L1	172	2	-	0/9/27/28	0/2/2/2
2	A2M	L1	1031	2	-	0/5/27/28	0/3/3/3
2	PSU	L1	109	2	-	0/7/25/26	0/2/2/2
2	PSU	L1	822	2	-	2/7/25/26	0/2/2/2
2	OMU	L1	799	2	-	4/9/27/28	0/2/2/2

The worst 5 of 324 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L1	99	A2M	C3'-C2'	-13.17	1.23	1.52
2	L1	166	A2M	C3'-C2'	-13.17	1.23	1.52
2	L1	512	A2M	C3'-C2'	-13.09	1.23	1.52
2	L1	27	A2M	C3'-C2'	-13.08	1.23	1.52
2	L1	576	A2M	C3'-C2'	-13.05	1.23	1.52

The worst 5 of 291 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L1	1842	4AC	CM7-C7-N4	7.69	128.59	115.29
2	L1	918	PSU	O2-C2-N1	-5.79	116.42	122.79
2	L1	627	OMU	N3-C2-N1	5.77	122.55	114.89
2	L1	109	PSU	O2-C2-N1	-5.70	116.52	122.79
2	L1	172	OMU	N3-C2-N1	5.68	122.43	114.89

There are no chirality outliers.

5 of 59 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	L1	27	A2M	C1'-C2'-O2'-CM'
2	L1	159	A2M	C1'-C2'-O2'-CM'
2	L1	468	A2M	C1'-C2'-O2'-CM'
2	L1	484	A2M	C1'-C2'-O2'-CM'
2	L1	509	OMG	C3'-C4'-C5'-O5'

There are no ring outliers.

18 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L1	686	PSU	1	0
2	L1	1232	PSU	1	0
2	L1	484	A2M	1	0
2	L1	1625	PSU	1	0
2	L1	621	OMC	2	0
2	L1	509	OMG	1	0
2	L1	668	A2M	1	0
2	L1	468	A2M	1	0
2	L1	627	OMU	1	0
2	L1	601	OMG	1	0
2	L1	1678	A2M	1	0
2	L1	797	OMC	1	0
2	L1	644	OMG	1	0
2	L1	116	OMU	2	0
2	L1	801	PSU	1	0
2	L1	159	A2M	1	0
2	L1	1031	A2M	1	0
2	L1	799	OMU	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 71 ligands modelled in this entry, 67 are monoatomic - leaving 4 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	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
72	GTP	SI	2001	69	26,34,34	1.15	2 (7%)	32,54,54	1.56	7 (21%)
73	SAH	SJ	301	-	24,28,28	1.20	3 (12%)	25,40,40	1.66	5 (20%)
70	ATP	NH	3000	69	26,33,33	0.61	0	31,52,52	0.74	2 (6%)
73	SAH	SK	301	-	24,28,28	1.22	3 (12%)	25,40,40	1.68	5 (20%)

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
72	GTP	SI	2001	69	-	8/18/38/38	0/3/3/3
73	SAH	SJ	301	-	-	4/11/31/31	0/3/3/3
70	ATP	NH	3000	69	-	2/18/38/38	0/3/3/3
73	SAH	SK	301	-	-	8/11/31/31	0/3/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
72	SI	2001	GTP	C5-C6	-4.08	1.39	1.47
73	SK	301	SAH	C2-N3	3.97	1.38	1.32
73	SJ	301	SAH	C2-N3	3.91	1.38	1.32
73	SK	301	SAH	C2-N1	2.37	1.38	1.33
73	SJ	301	SAH	C2-N1	2.36	1.38	1.33

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
73	SK	301	SAH	N3-C2-N1	-5.36	120.30	128.68
73	SJ	301	SAH	N3-C2-N1	-5.24	120.49	128.68
72	SI	2001	GTP	PB-O3B-PG	-3.56	120.61	132.83
73	SJ	301	SAH	C5'-SD-CG	-3.37	92.14	102.27
72	SI	2001	GTP	C5-C6-N1	3.30	119.78	113.95

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

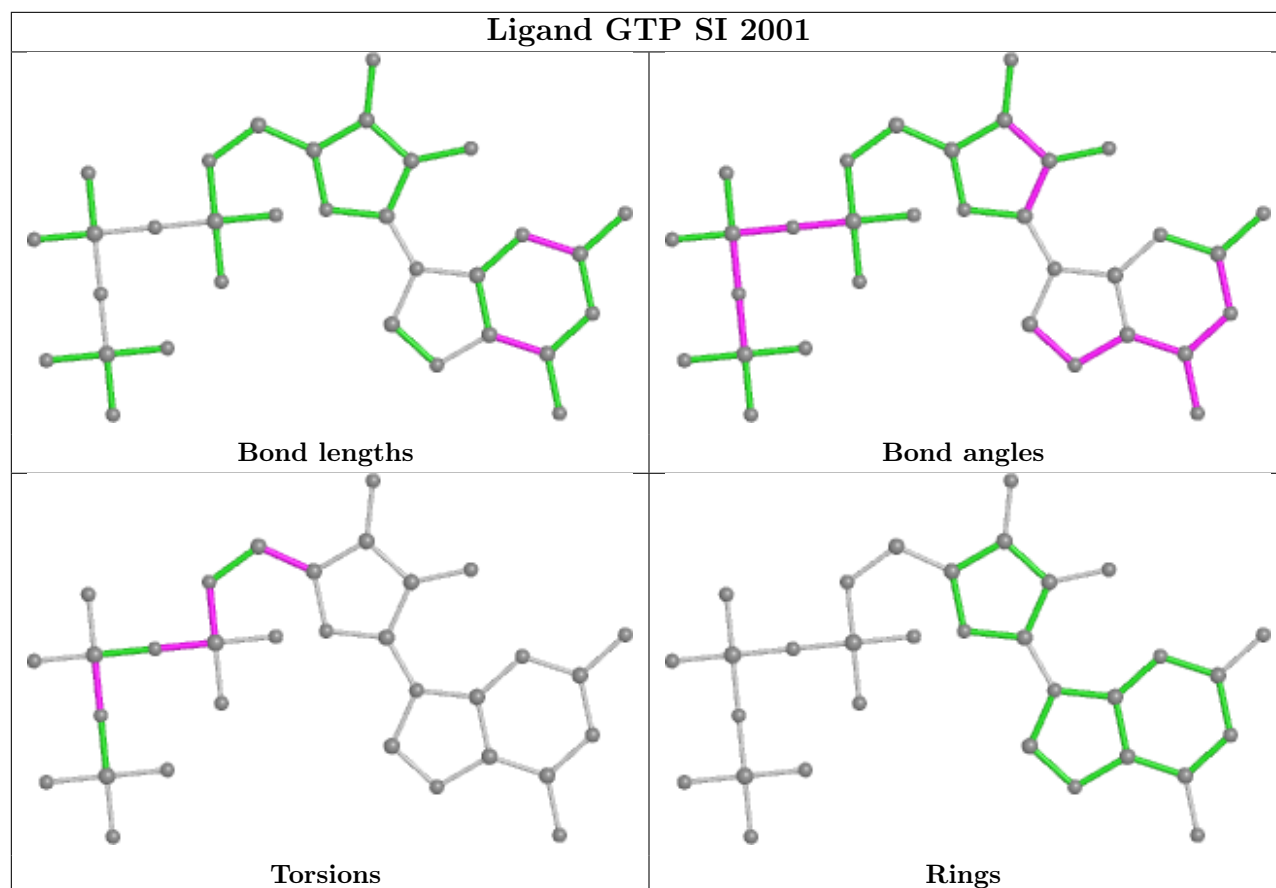
Mol	Chain	Res	Type	Atoms
72	SI	2001	GTP	C5'-O5'-PA-O1A
72	SI	2001	GTP	C5'-O5'-PA-O2A
73	SJ	301	SAH	N-CA-CB-CG
73	SJ	301	SAH	C-CA-CB-CG
73	SK	301	SAH	OXT-C-CA-CB

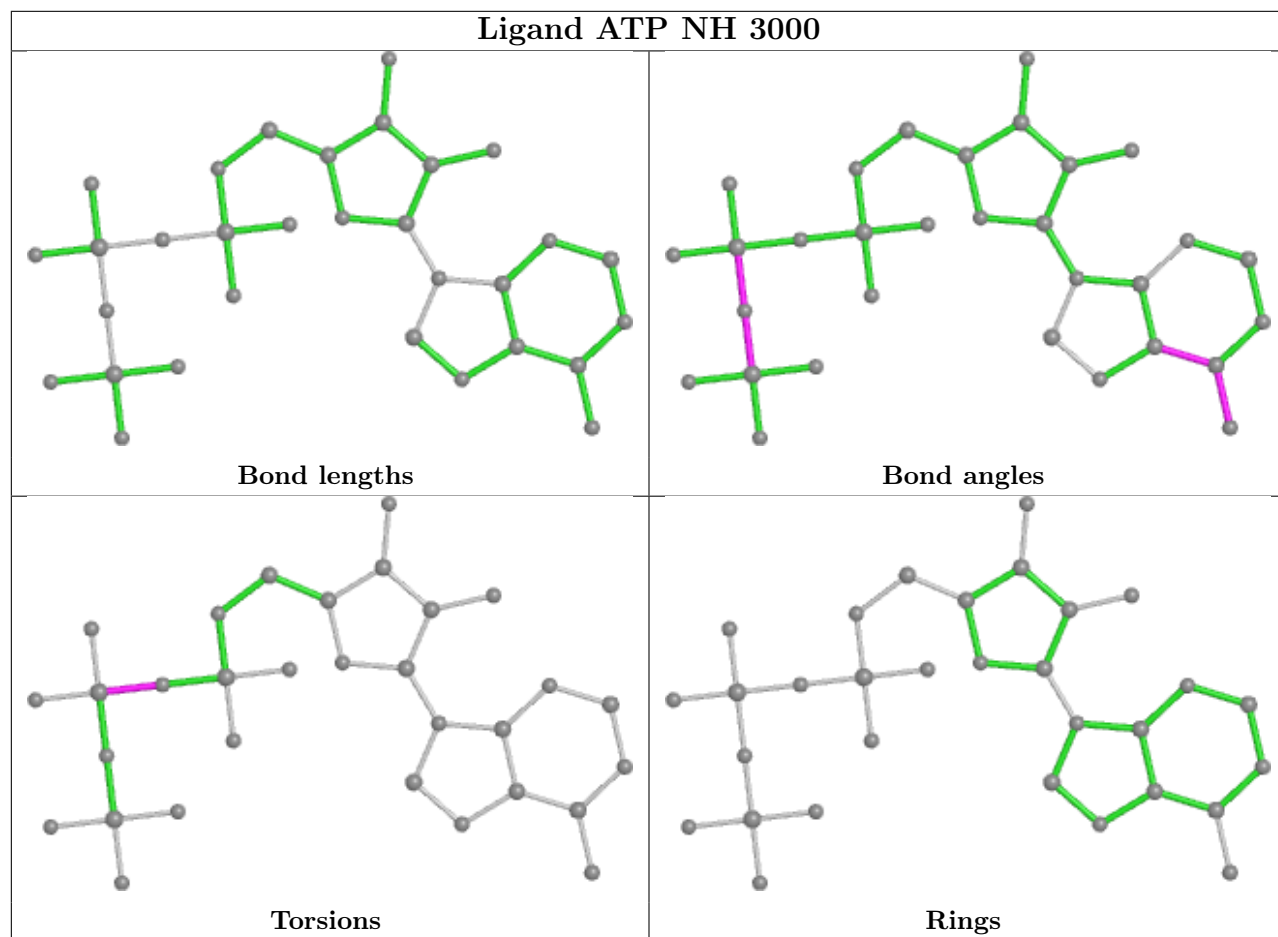
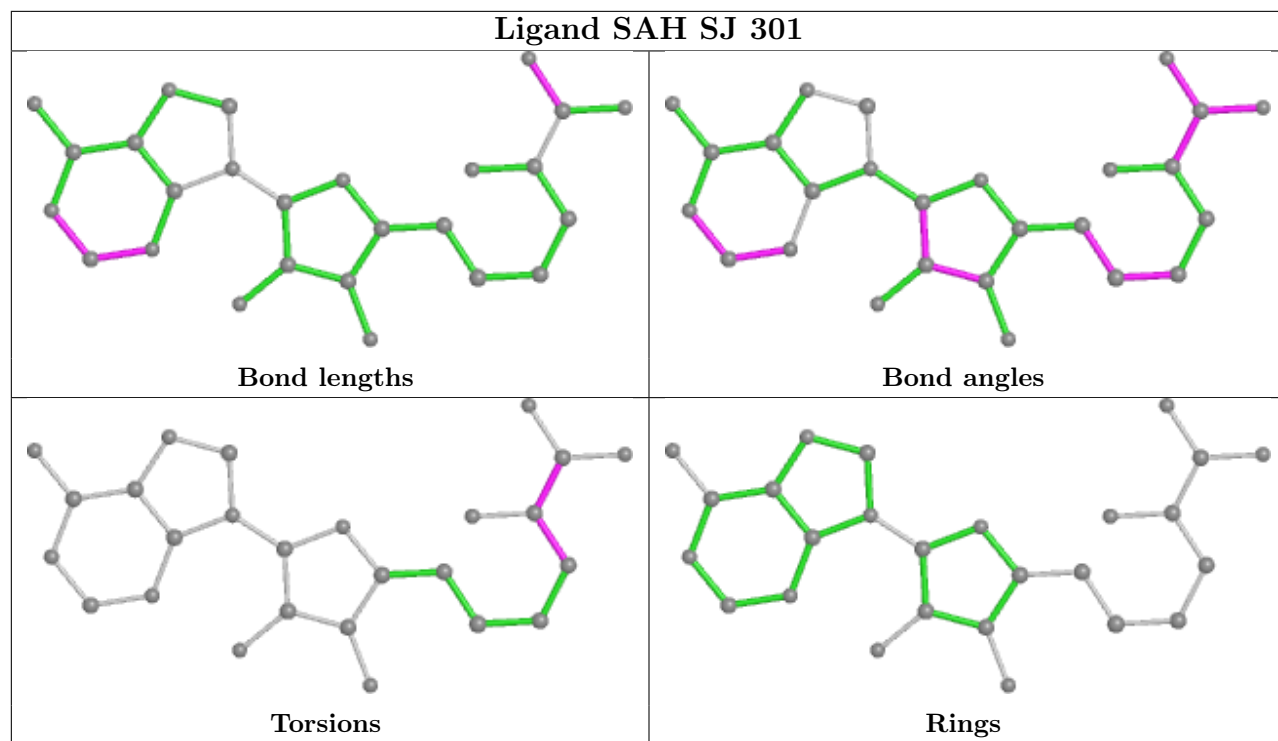
There are no ring outliers.

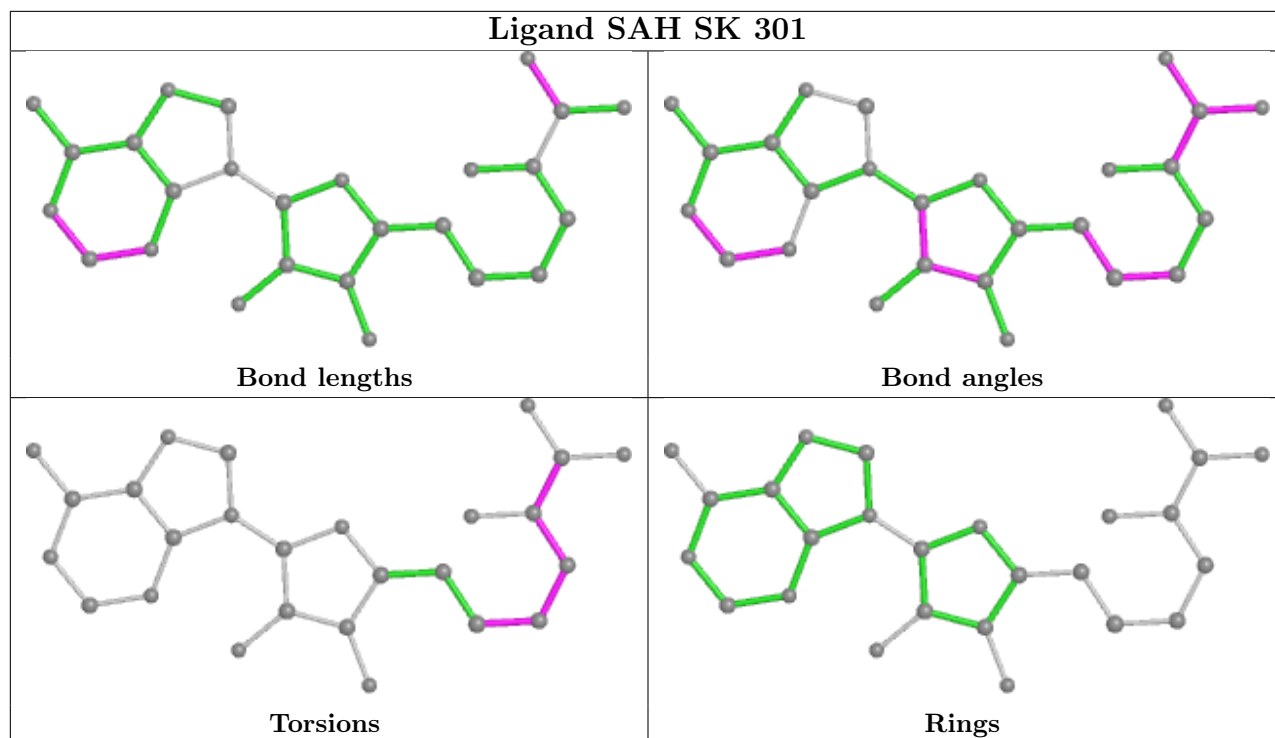
3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
73	SJ	301	SAH	5	0
70	NH	3000	ATP	1	0
73	SK	301	SAH	5	0

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 than 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.







## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
44	NR	41
66	SX	9
17	LI	5
64	SU	5
63	ST	3
68	SZ	1

The worst 5 of 64 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	SX	726:UNK	C	763:UNK	N	250.94
1	SX	779:UNK	C	1200:UNK	N	226.33
1	SX	1472:UNK	C	2269:UNK	N	103.10
1	SX	930:UNK	C	689:UNK	N	85.48
1	ST	298:LEU	C	414:UNK	N	52.29

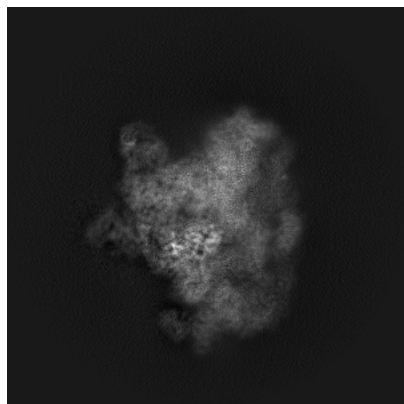
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-23938. These allow visual inspection of the internal detail of the map and identification of artifacts.

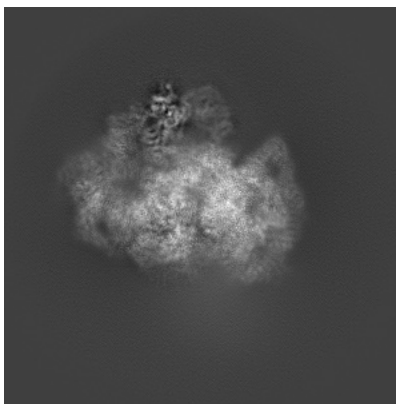
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

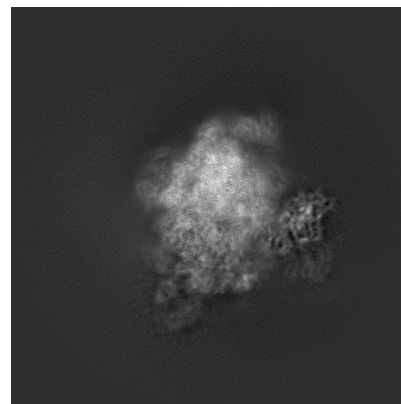
#### 6.1.1 Primary map



X

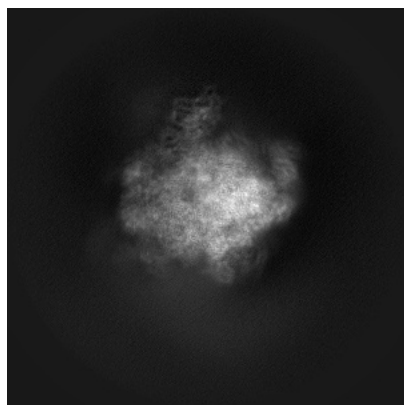


Y

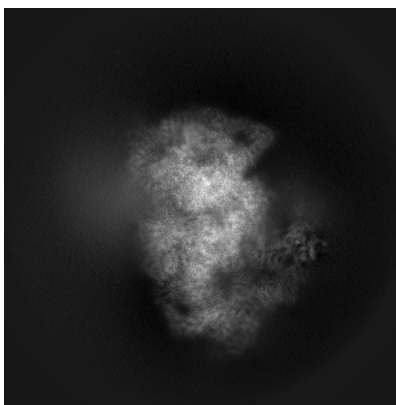


Z

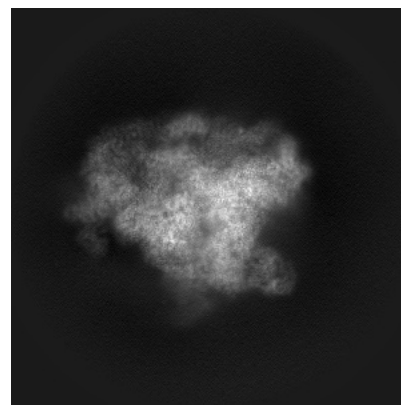
#### 6.1.2 Raw map



X



Y



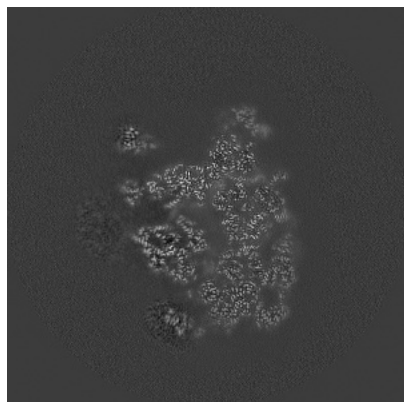
Z

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

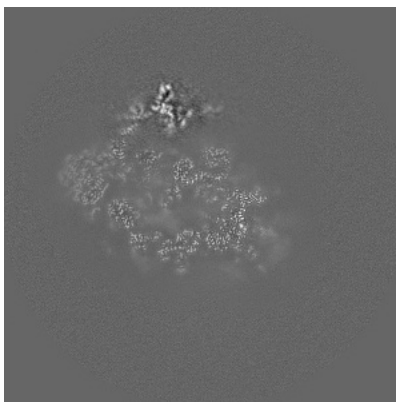


## 6.2 Central slices [i](#)

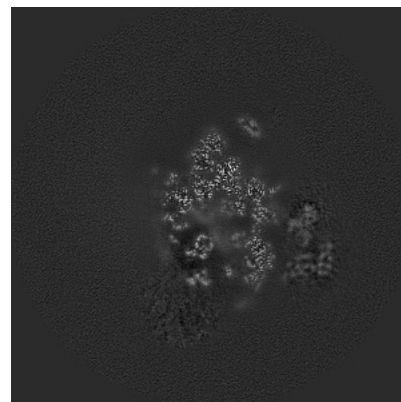
### 6.2.1 Primary map



X Index: 280

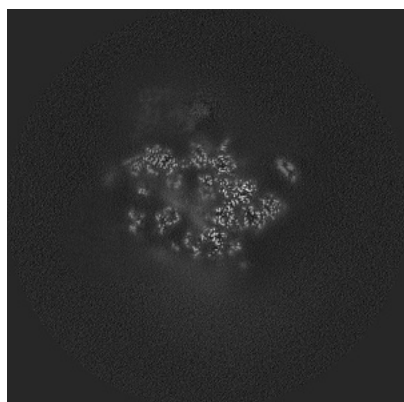


Y Index: 280

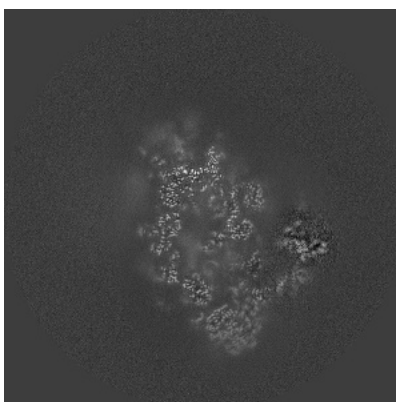


Z Index: 280

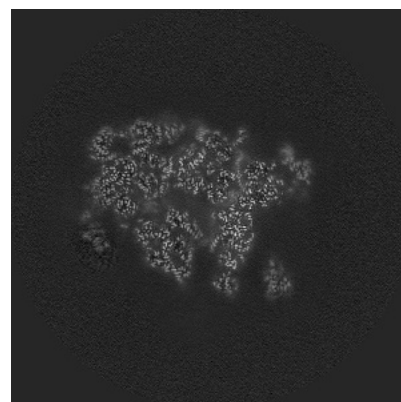
### 6.2.2 Raw map



X Index: 280



Y Index: 280

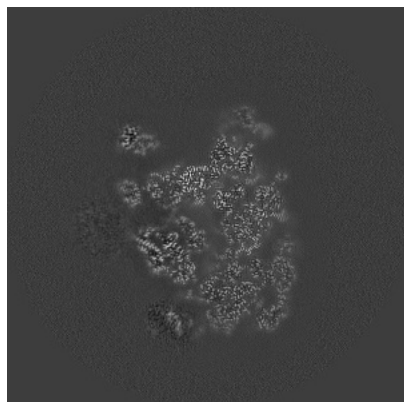


Z Index: 280

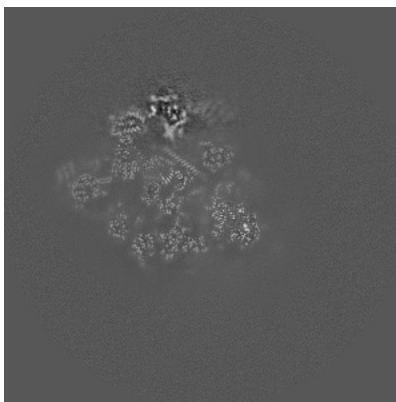
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

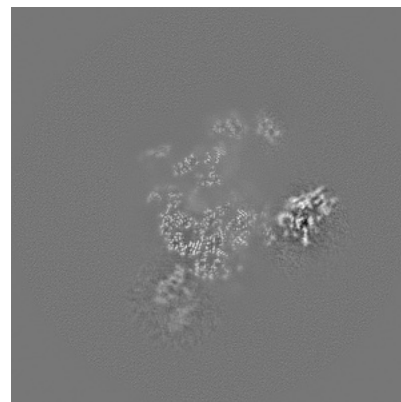
### 6.3.1 Primary map



X Index: 277

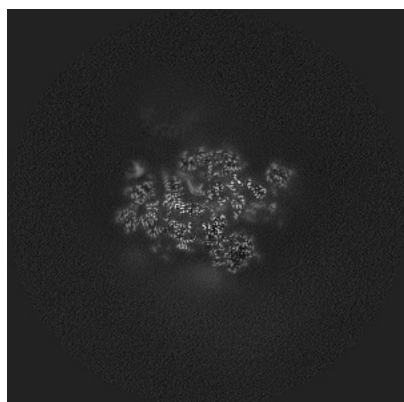


Y Index: 268

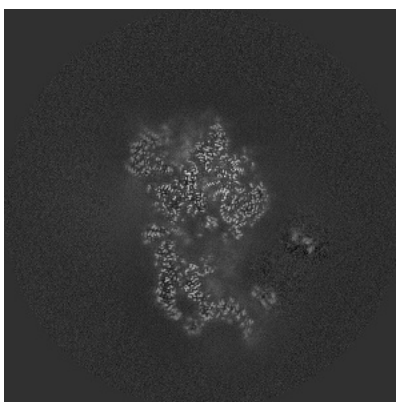


Z Index: 234

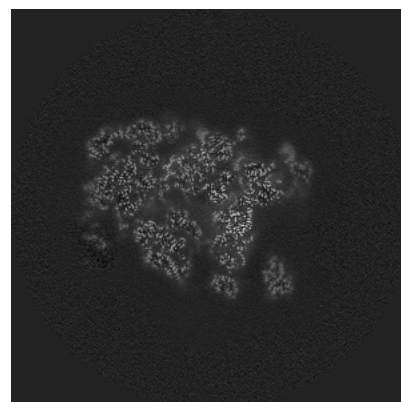
### 6.3.2 Raw map



X Index: 313



Y Index: 299



Z Index: 277

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

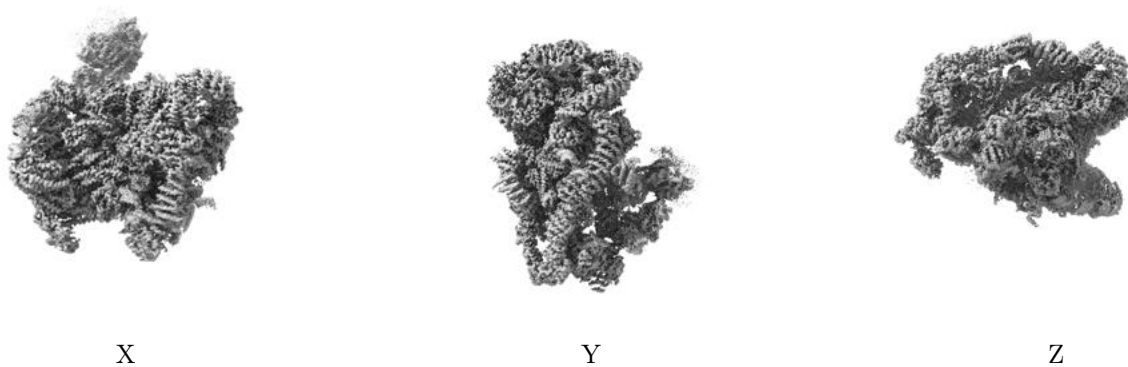
## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 3.28. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.4.2 Raw map



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.

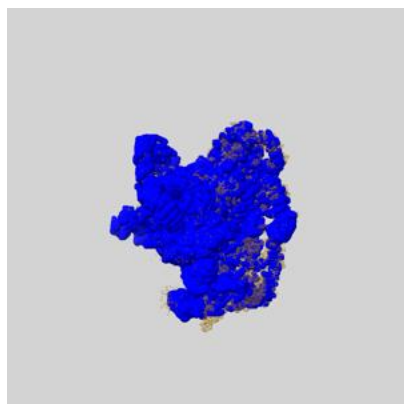
## 6.5 Mask visualisation [i](#)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

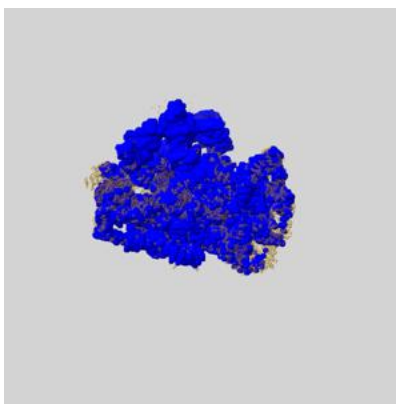
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

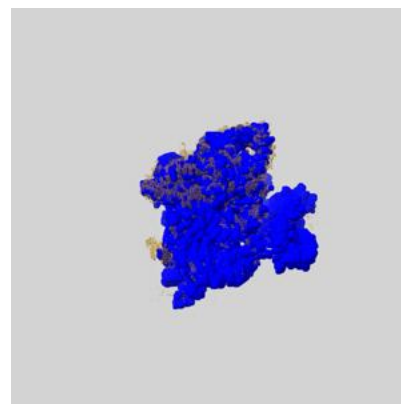
### 6.5.1 emd\_23938\_msk\_1.map [i](#)



X



Y

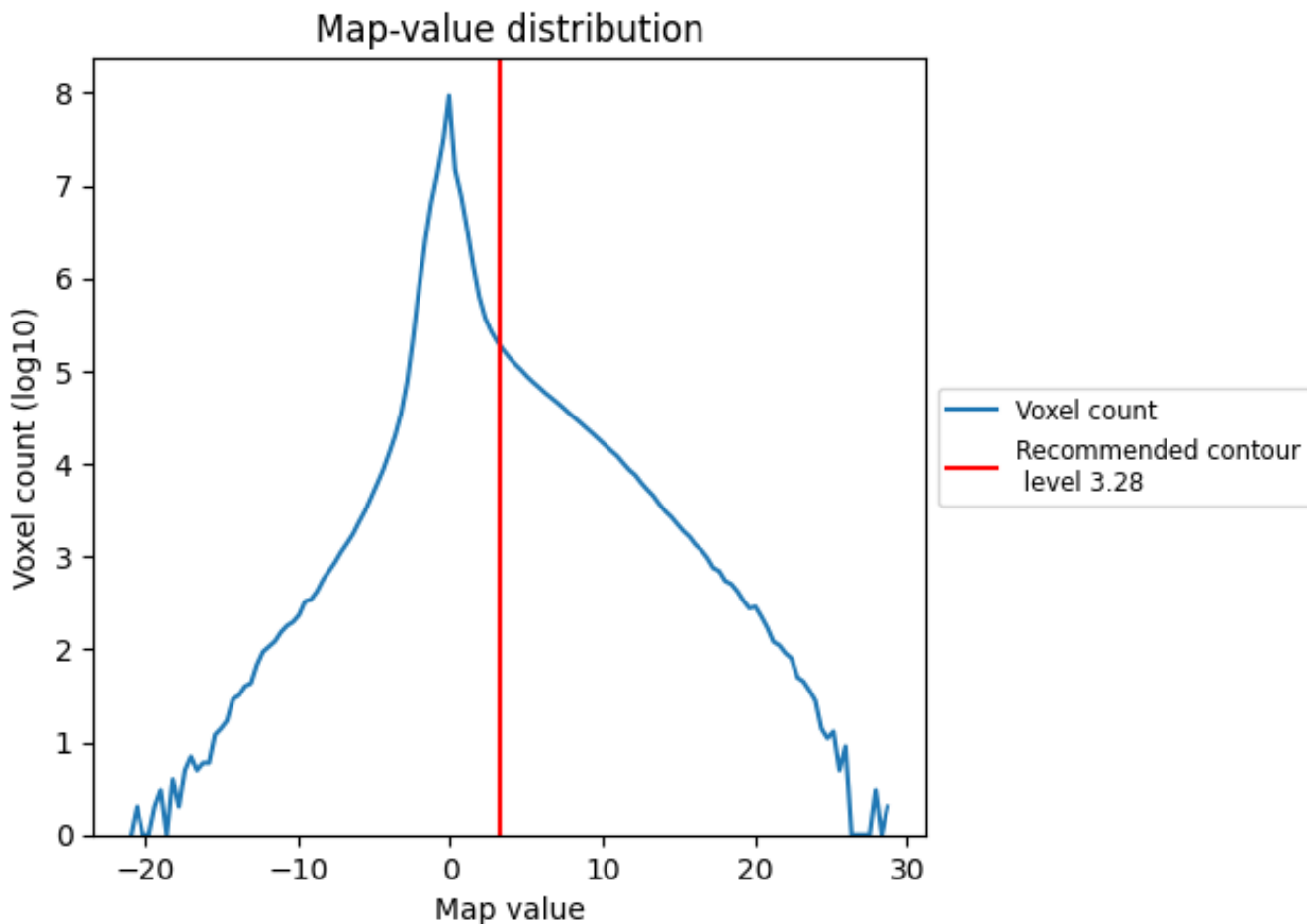


Z

## 7 Map analysis [i](#)

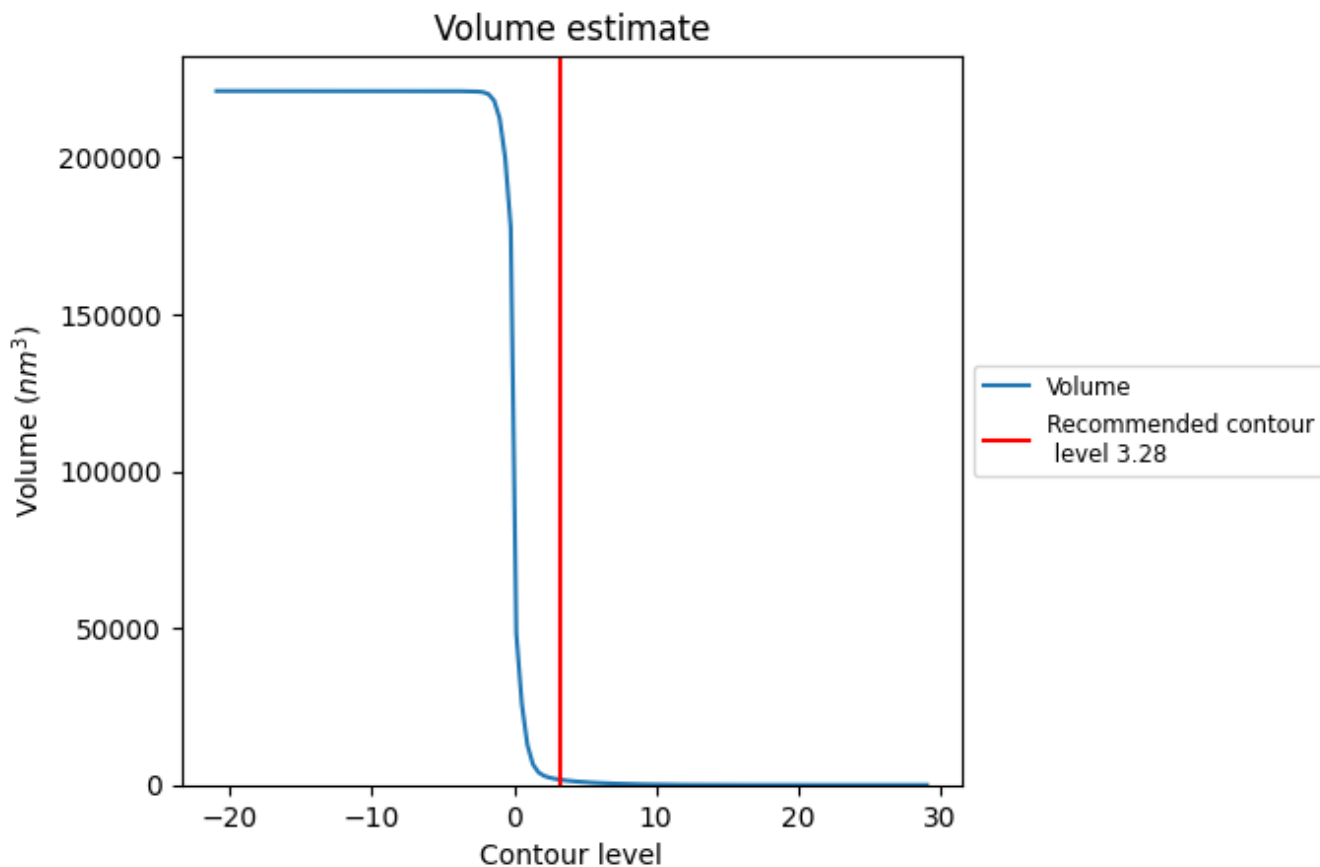
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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.

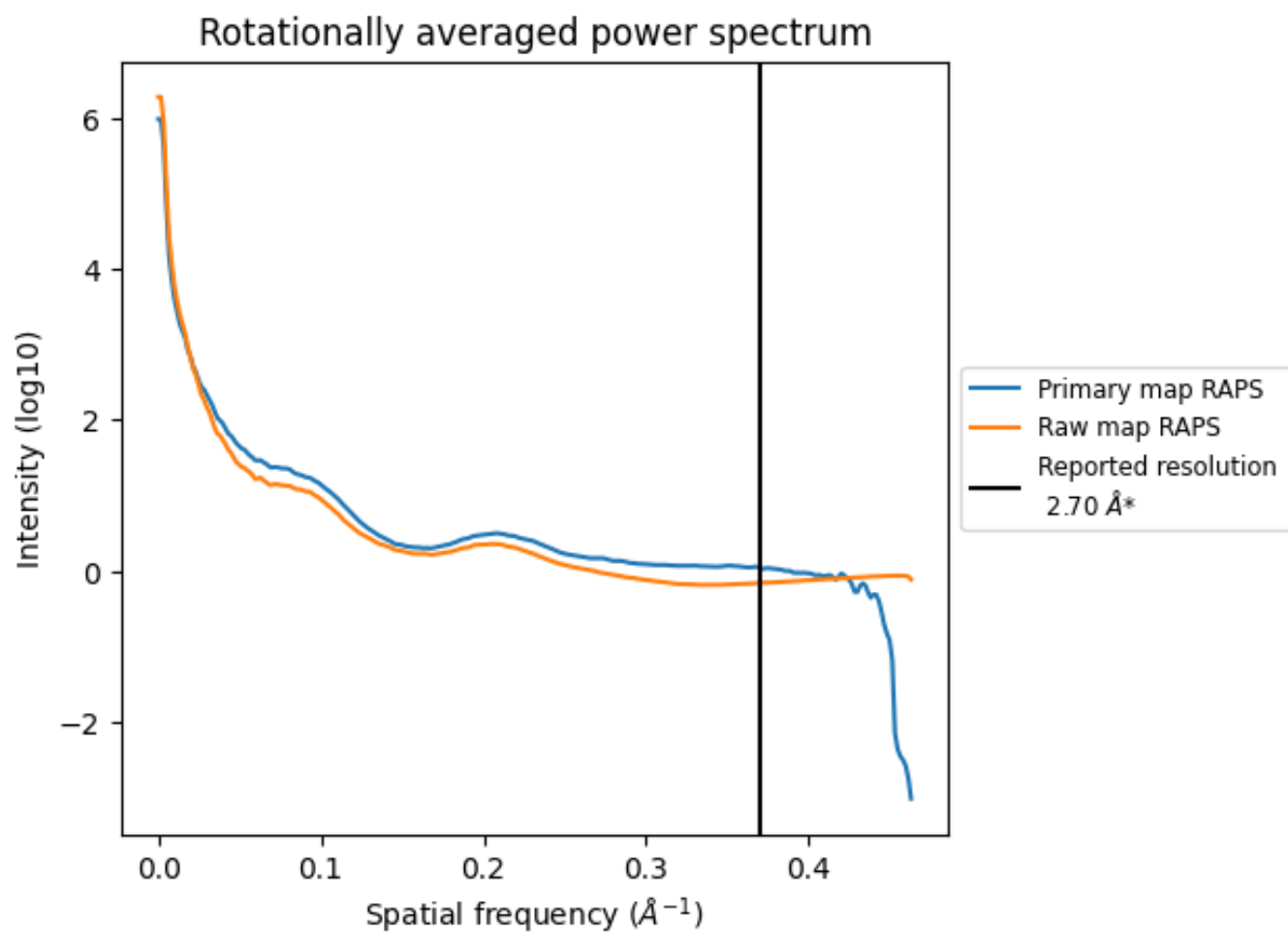
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 1662  $\text{nm}^3$ ; this corresponds to an approximate mass of 1501 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 i

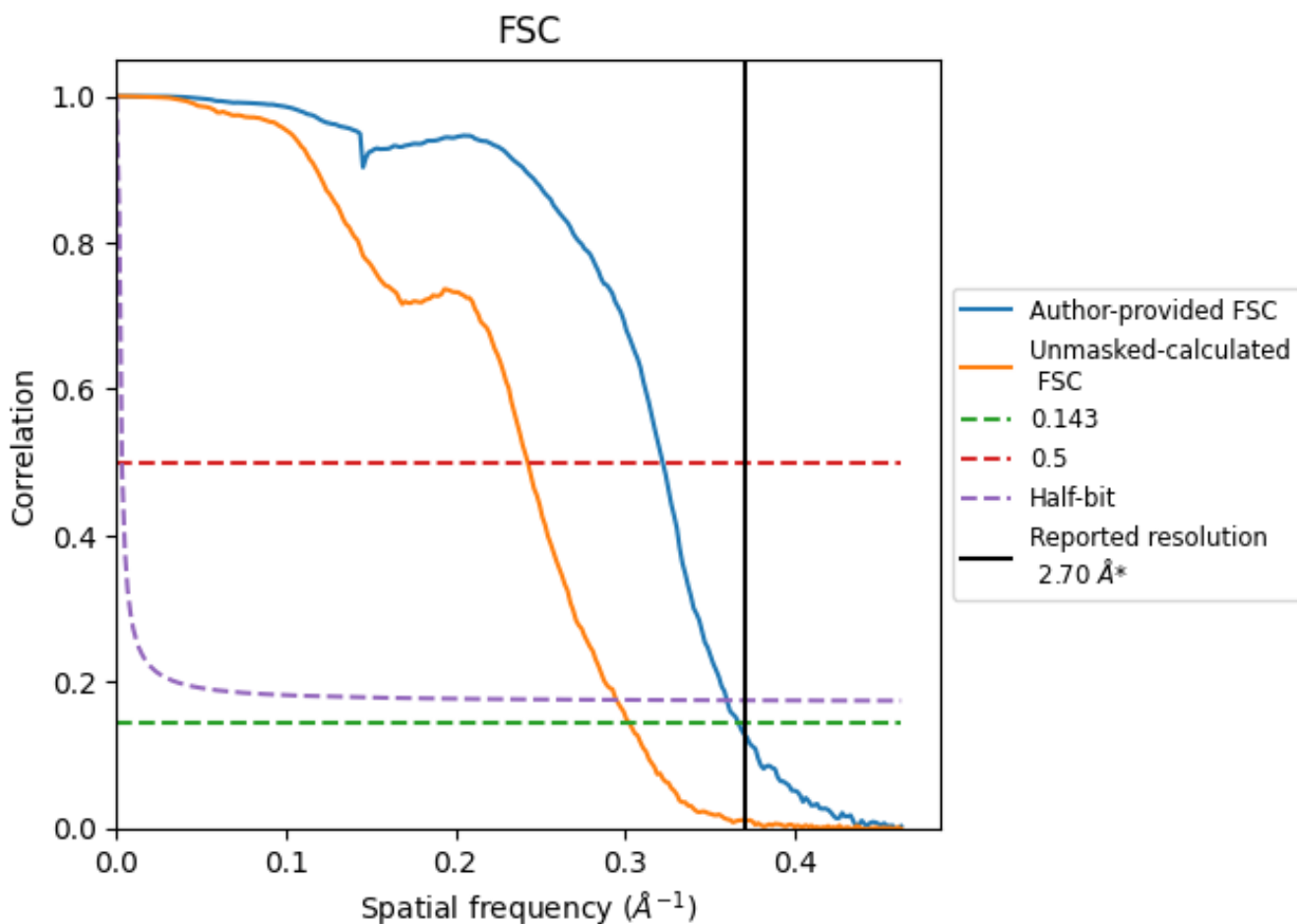


\*Reported resolution corresponds to spatial frequency of 0.370 Å<sup>-1</sup>

## 8 Fourier-Shell correlation [i](#)

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 [i](#)



\*Reported resolution corresponds to spatial frequency of 0.370 Å<sup>-1</sup>



## 8.2 Resolution estimates [i](#)

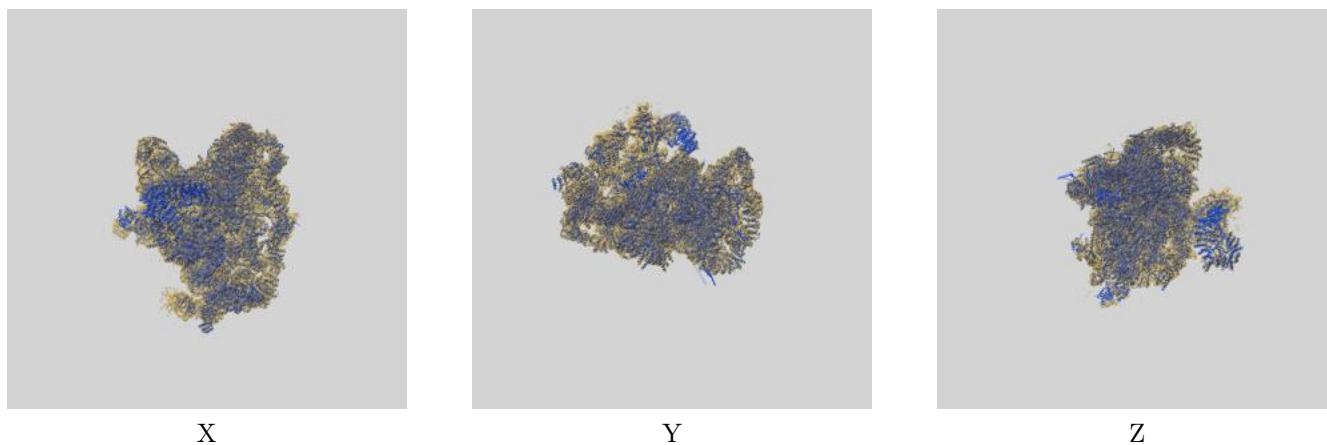
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.70	-	-
Author-provided FSC curve	2.72	3.10	2.78
Unmasked-calculated*	3.30	4.12	3.39

\*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 3.30 differs from the reported value 2.7 by more than 10 %

## 9 Map-model fit [i](#)

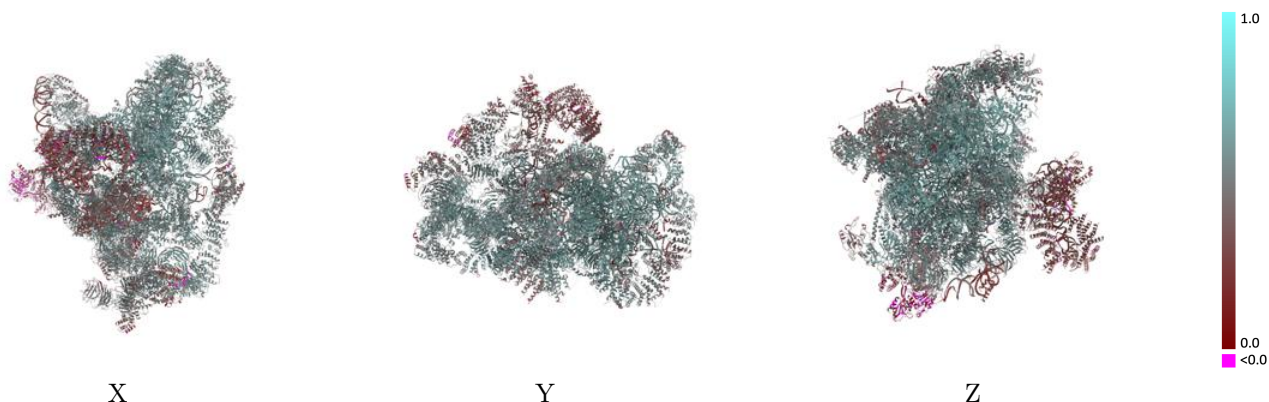
This section contains information regarding the fit between EMDB map EMD-23938 and PDB model 7MQA. Per-residue inclusion information can be found in section [3](#) on page [19](#).

### 9.1 Map-model overlay [i](#)



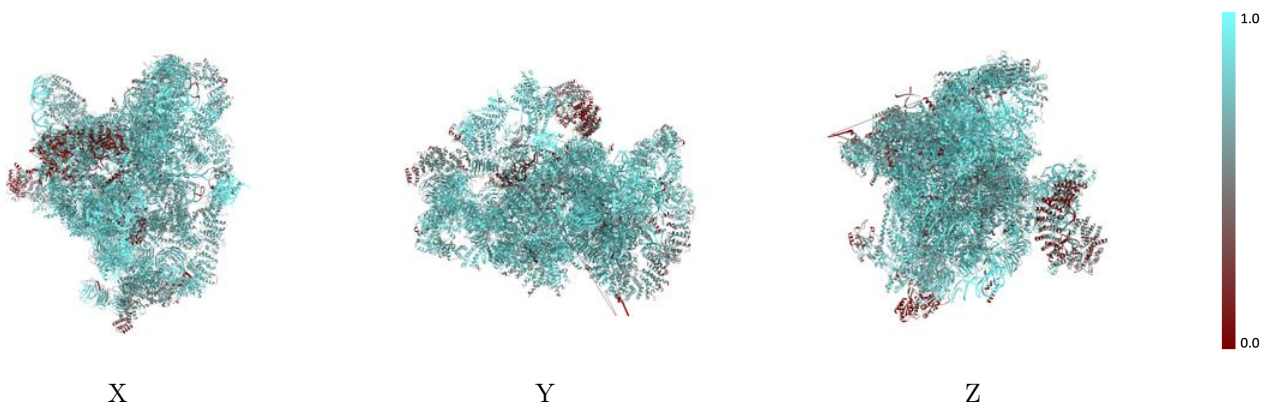
The images above show the 3D surface view of the map at the recommended contour level 3.28 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [i](#)



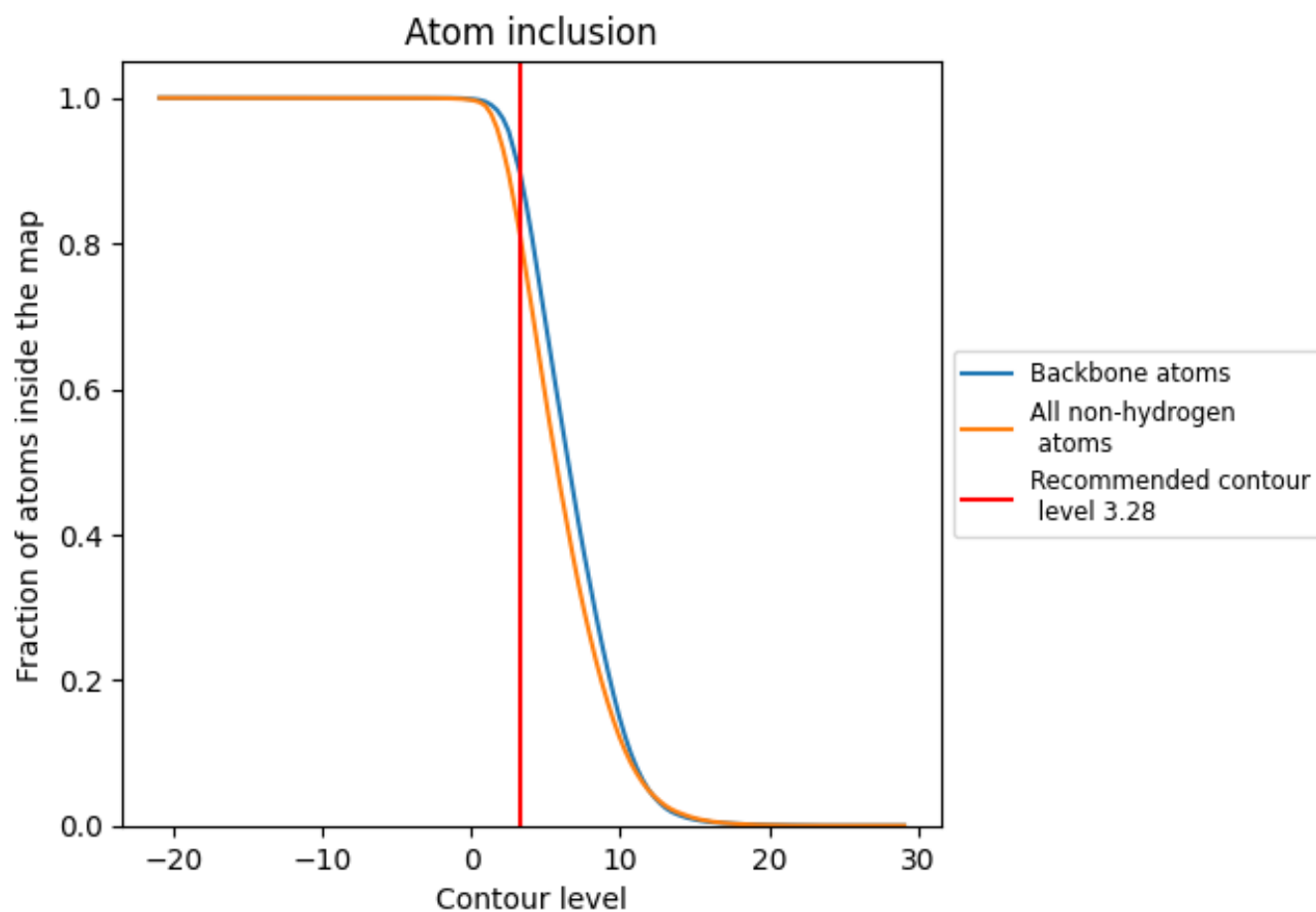
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (3.28).







































































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 90% 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 (3.28) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8121	 0.5260
L0	 0.9481	 0.6010
L1	 0.8953	 0.5410
L2	 0.8820	 0.5390
L3	 0.6209	 0.5020
L4	 0.9230	 0.6290
L5	 0.8658	 0.5630
L6	 0.8578	 0.5760
L7	 0.8075	 0.5570
L8	 0.9027	 0.6100
L9	 0.9486	 0.6370
LA	 0.8277	 0.2670
LC	 0.8861	 0.5840
LD	 0.8878	 0.6140
LF	 0.9400	 0.6390
LG	 0.8255	 0.5360
LH	 0.7806	 0.5370
LI	 0.6160	 0.4110
LJ	 0.7987	 0.5220
LK	 0.5665	 0.4120
LL	 0.7787	 0.5350
LM	 0.7769	 0.5150
LN	 0.8772	 0.5840
LO	 0.9006	 0.5550
LP	 0.7973	 0.5260
LQ	 0.8115	 0.5220
LR	 0.8458	 0.5230
LS	 0.8619	 0.5810
LT	 0.8888	 0.5640
LU	 0.9174	 0.6230
LW	 0.8973	 0.5810
LY	 0.4333	 0.4920
LZ	 0.8675	 0.5480
NA	 0.7602	 0.5050
NB	 0.9098	 0.6170



*Continued on next page...*

Continued from previous page...

Chain	Atom inclusion	Q-score
ND	0.6772	0.5130
NF	0.9057	0.6190
NG	0.8890	0.5660
NH	0.7383	0.5010
NI	0.5747	0.4670
NL	0.8221	0.5640
NM	0.8528	0.5930
NO	0.9683	0.6610
NP	0.2720	0.5080
NQ	0.9300	0.6270
NR	0.3638	0.2700
NS	0.5400	0.2530
NT	0.8337	0.3110
NU	0.0565	0.2180
NV	0.8077	0.5760
SA	0.8778	0.5790
SB	0.8282	0.5690
SC	0.8711	0.5850
SD	0.9482	0.6340
SE	0.8783	0.5880
SF	0.9170	0.6120
SG	0.8742	0.5830
SH	0.8667	0.5770
SI	0.8630	0.5820
SJ	0.8281	0.3410
SK	0.7857	0.4580
SL	0.9068	0.6120
SM	0.8600	0.5680
SP	0.7623	0.5380
SQ	0.7520	0.5640
SR	0.9225	0.6160
SS	0.7680	0.5200
ST	0.7722	0.4210
SU	0.8632	0.4700
SW	0.8638	0.5390
SX	0.7561	0.3760
SY	0.8053	0.5620
SZ	0.8363	0.3420