wwPDB EM Validation Summary Report (i)

Mar 12, 2026 – 07:38 AM UTC

PDB ID	:	6QX9 / pdb_00006qx9
EMDB ID	:	EMD-4665
Title	:	Structure of a human fully-assembled precatalytic spliceosome (pre-B complex).
Authors	:	Charenton, C.; Wilkinson, M.E.; Nagai, K.
Deposited on	:	2019-03-07
Resolution	:	3.28 Å(reported)

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

We welcome your comments at *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 (i) 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 (i)) were used in the production of this report:

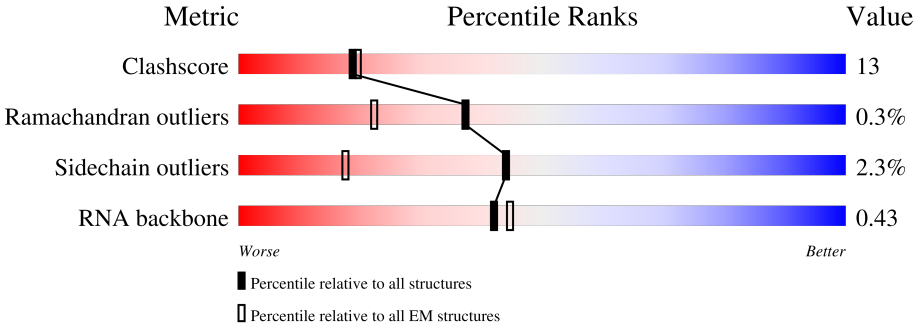
EMDB validation analysis	:	0.0.1.dev132
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4-5-2 with Phenix2.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

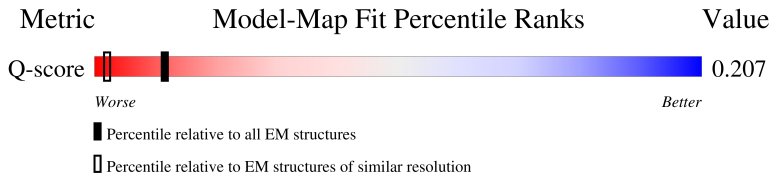
1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.28 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
RNA backbone	8273	3508	-
Q-score	-	25397	14492 (2.78 - 3.78)

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

Mol	Chain	Length	Quality of chain
1	1	164	100% 38% 39% 22% •
2	6	106	17% 26% 19% 5% 50%
3	5O	357	75% 66% 19% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	B4	424	18% 16% 82%
5	13	126	64% 42% 17% 36%
5	23	126	66% 48% 17% 34%
5	43	126	53% 44% 21% 34%
5	53	126	50% 17% 33%
6	4B	522	17% 53% 16% 31%
7	1e	92	84% 47% 29% 8% 16%
7	2e	92	88% 62% 26% 12%
7	4e	92	67% 48% 35% 17%
7	5e	92	30% 57% 27% 16%
8	I	62	40% 18% 18% 60%
9	1K	437	46% 27% 17% 54%
10	4C	499	6% 52% 8% 40%
11	11	119	68% 53% 13% 32%
11	21	119	67% 49% 18% 33%
11	41	119	23% 34% 28% 6% 32%
11	51	119	28% 35% 28% 32%
12	R	480	19% 78%
13	1f	86	86% 59% 22% 14%
13	2f	86	84% 66% 17% 16%
13	4f	86	56% 51% 31% 16%
13	5f	86	55% 59% 24% 15%
14	66	80	90% 69% 19% 10%
15	X	155	12% 15% 16% 68%
16	12	118	81% 60% 18% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
16	22	118	81% 60% 20% 19%
16	42	118	26% 42% 36% 22%
16	52	118	43% 51% 31% 17%
17	5	117	11% 38% 42% 9% 11%
18	67	103	75% 60% 13% 25%
19	62	95	100% 83% 15%
20	2B	225	41% 35% 6% 59%
21	A2	209	69% 49% 16% 31%
22	B2	895	23% 17% 6% 77%
23	5C	854	77% 22%
24	5X	820	48% 52% 18% 29%
25	1b	240	36% 25% 8% 64%
25	2b	240	34% 26% 8% 66%
25	4b	240	25% 25% 9% 66%
25	5b	240	6% 27% 70%
26	B5	86	80% 64% 16% 20%
27	1A	282	35% 28% 6% 65%
28	S	800	12% 85%
29	5J	850	14% 76% 18% 6%
30	4D	128	77% 20%
31	63	102	83% 61% 22% 17%
32	2	188	50% 27% 19% 50%
33	B3	1217	97% 76% 21%
34	1g	76	96% 54% 37%
34	2g	76	96% 72% 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	4g	76	86% 46% 42% 9%
34	5g	76	18% 47% 39% 11%
35	68	96	99% 72% 19% 8%
36	5A	2311	70% 25%
37	A3	501	76% 66% 9% 24%
38	U	555	65% 17% 18%
39	5D	142	73% 25%
40	64	139	53% 43% 8% 47%
41	BP	104	96% 84% 12%
42	1C	159	31% 16% 13% 69%
43	K	1007	32% 18% 11% 68%
44	4A	683	11% 26% 9% 65%
45	4	146	23% 42% 36% 8% 14%
46	2A	255	64% 46% 17% 36%
47	A1	647	23% 21% 74%
48	65	91	84% 60% 20% 16%
49	5B	2136	6% 74% 20% 6%
50	B1	1304	65% 54% 11% 35%

2 Entry composition (i)

There are 54 unique types of molecules in this entry. The entry contains 137494 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 U1 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	1	164	3485	1555	607	1159	164	0	0

- Molecule 2 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	6	53	1133	506	203	371	53	0	0

- Molecule 3 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	5O	306	2394	1501	422	457	14	0	0

- Molecule 4 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	B4	78	618	399	101	115	3	0	0

- Molecule 5 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	13	81	637	400	112	119	6	0	0
5	53	84	657	412	116	123	6	0	0
5	23	83	652	409	115	122	6	0	0
5	43	83	652	409	115	122	6	0	0

- Molecule 6 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	4B	359	2842	1793	509	521	19	0	0

- Molecule 7 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	5e	77	638	405	113	115	5	0	0
7	4e	76	631	400	112	114	5	0	0
7	1e	77	638	405	113	115	5	0	0
7	2e	81	669	424	119	121	5	0	0

- Molecule 8 is a RNA chain called AdML pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	I	25	530	237	92	177	24	0	0

- Molecule 9 is a protein called U1 small nuclear ribonucleoprotein 70 kDa.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	1K	201	1649	1036	317	291	5	0	0

- Molecule 10 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	4C	301	2375	1486	418	456	15	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	41	81	641	408	112	118	3	0	0
11	11	81	641	408	112	118	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	51	81	641	408	112	118	3	0	0
11	21	80	634	404	111	115	4	0	0

- Molecule 12 is a protein called RNA-binding protein 42.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	R	106	874	553	160	157	4	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	1f	74	576	373	95	103	5	0	0
13	2f	72	562	364	93	100	5	0	0
13	5f	73	567	367	94	101	5	0	0
13	4f	72	562	364	93	100	5	0	0

- Molecule 14 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	66	72	567	360	97	108	2	0	0

- Molecule 15 is a protein called U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	X	49	394	247	74	69	4	0	0

- Molecule 16 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	22	95	774	486	141	142	5	0	0
16	42	92	737	463	138	131	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	52	98	796	498	144	148	6	0	0
			Total	C	N	O	S		
16	12	95	777	486	141	144	6	0	0

- Molecule 17 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
17	5	104	2192	983	372	734	103	0	0

- Molecule 18 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	67	77	604	383	102	116	3	0	0

- Molecule 19 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	62	95	761	486	126	145	4	0	0

- Molecule 20 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	2B	92	745	480	130	130	5	0	0

- Molecule 21 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	A2	144	1221	782	219	214	6	0	0

- Molecule 22 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	B2	208	1699	1093	302	295	9	0	0

- Molecule 23 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	5C	852	6727	4300	1127	1266	34	0	0

- Molecule 24 is a protein called Probable ATP-dependent RNA helicase DDX23.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	5X	583	4780	3014	855	893	18	7	0

- Molecule 25 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	1b	86	692	435	126	124	7	0	0
25	2b	82	664	419	121	117	7	0	0
25	5b	73	594	376	108	103	7	0	0
25	4b	82	669	423	122	117	7	0	0

- Molecule 26 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	B5	69	567	360	99	103	5	0	0

- Molecule 27 is a protein called U1 small nuclear ribonucleoprotein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	1A	98	787	506	134	143	4	0	0

- Molecule 28 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	S	120	917	576	168	168	5	0	0

- Molecule 29 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	5J	803	6316	3963	1155	1170	28	0	0

- Molecule 30 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	4D	123	955	604	170	176	5	0	0

- Molecule 31 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	63	85	699	440	120	136	3	0	0

- Molecule 32 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
32	2	94	1984	887	331	672	94	0	0

- Molecule 33 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
33	B3	1186	9296	5898	1580	1773	45	0	0

- Molecule 34 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
34	1g	73	568	358	102	102	6	0	0
			Total	C	N	O	S		
34	2g	73	568	358	102	102	6	0	0
			Total	C	N	O	S		
34	5g	74	577	364	104	103	6	0	0
			Total	C	N	O	S		
34	4g	74	577	364	104	103	6	0	0

- Molecule 35 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
35	68	95	722	446	124	151	1	0	0

- Molecule 36 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
36	5A	2212	18366	11840	3193	3253	80	0	0

- Molecule 37 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
37	A3	383	3227	2029	566	618	14	0	0

- Molecule 38 is a protein called U4/U6.U5 tri-snRNP-associated protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
38	U	456	3750	2427	635	674	14	0	0

- Molecule 39 is a protein called Thioredoxin-like protein 4A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
39	5D	141	1169	751	194	214	10	0	0

- Molecule 40 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
40	64	73	596	376	105	109	6	0	0

- Molecule 41 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
41	BP	100	766	473	135	145	13	0	0

- Molecule 42 is a protein called U1 small nuclear ribonucleoprotein C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
42	1C	50	425	266	73	82	4	0	0

- Molecule 43 is a protein called Serine/threonine-protein kinase PRP4 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
43	K	322	2626	1682	462	467	15	0	0

- Molecule 44 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
44	4A	239	1946	1237	360	342	7	0	0

- Molecule 45 is a RNA chain called U4 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
45	4	126	2690	1202	474	888	126	0	0

- Molecule 46 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
46	2A	162	1282	820	219	240	3	0	0

- Molecule 47 is a protein called Splicing factor 3A subunit 1,Splicing factor 3A subunit 1,Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
47	A1	168	1339	855	237	245	2	0	0

- Molecule 48 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
48	65	76	587	373	96	114	4	0	0

- Molecule 49 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
49	5B	2001	16077	10235	2767	2991	84	0	0

- Molecule 50 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
50	B1	848	6749	4330	1160	1220	39	0	0

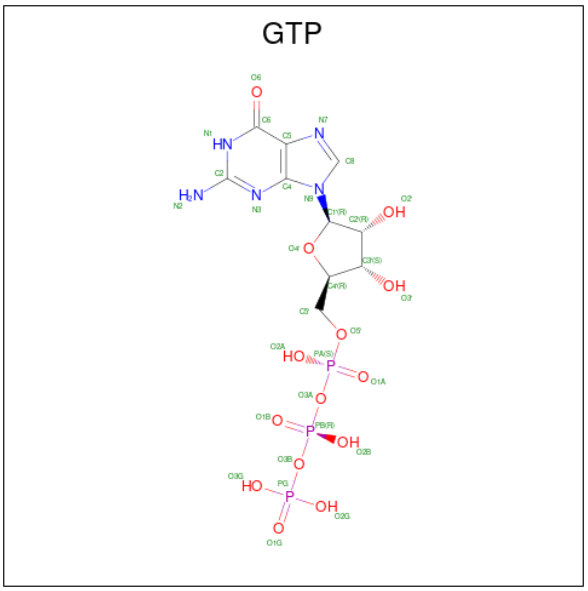
- Molecule 51 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
51	A2	1	1	1	0
51	U	1	1	1	0
51	BP	3	3	3	0
51	1C	1	1	1	0

- Molecule 52 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

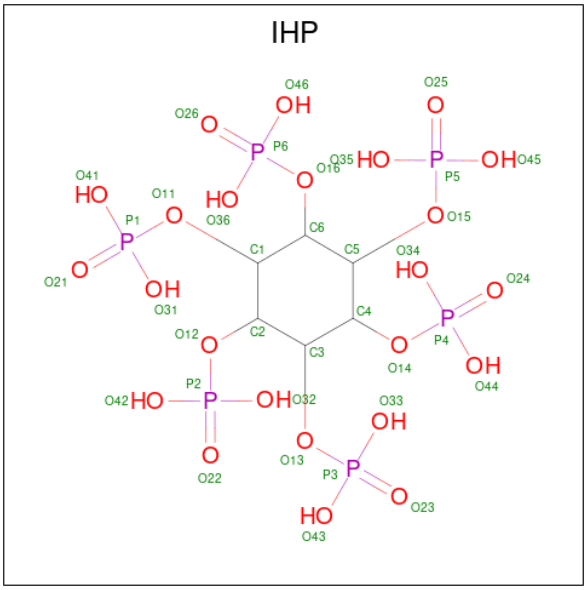
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
52	5C	1	1	1	0

- Molecule 53 is GUANOSINE-5'-TRIPHOSPHATE (CCD ID: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
53	5C	1	32	10	5	14	3	0

- Molecule 54 is INOSITOL HEXAKISPHOSPHATE (CCD ID: IHP) (formula: $C_6H_{18}O_{24}P_6$).



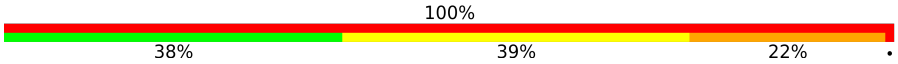
Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
54	5A	1	36	6	24	6	0

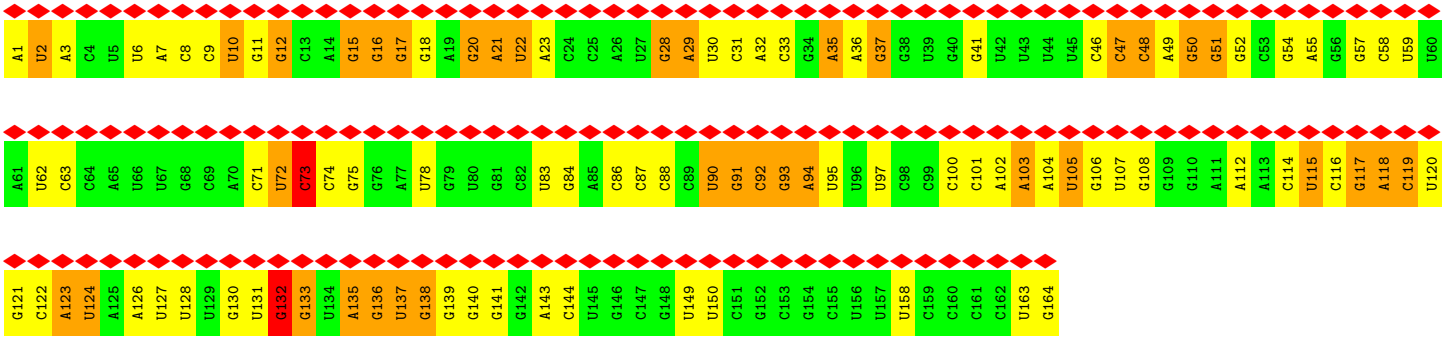
3 Residue-property plots (i)

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: U1 snRNA

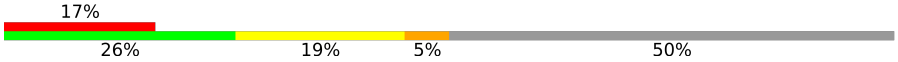
Chain 1:

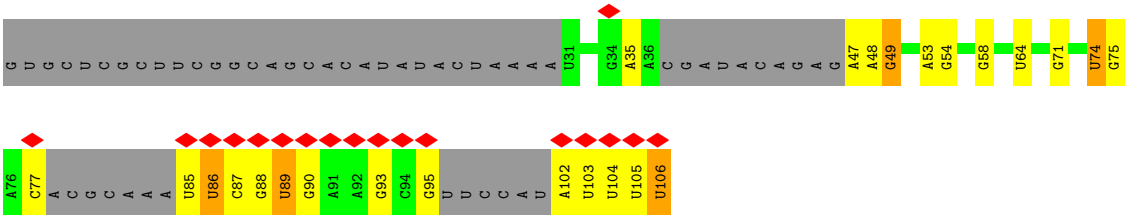




- Molecule 2: U6 snRNA

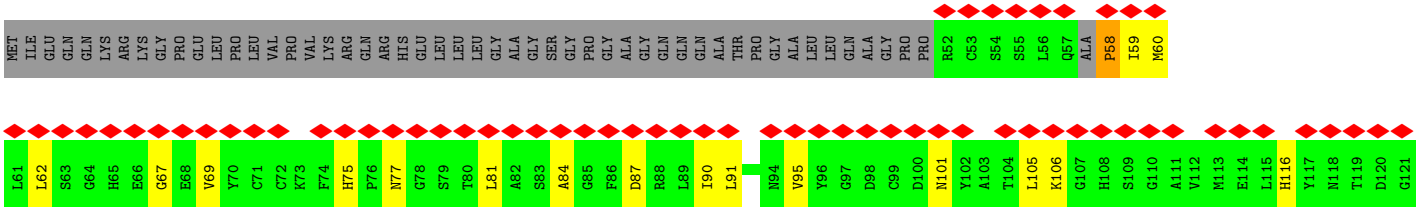
Chain 6:

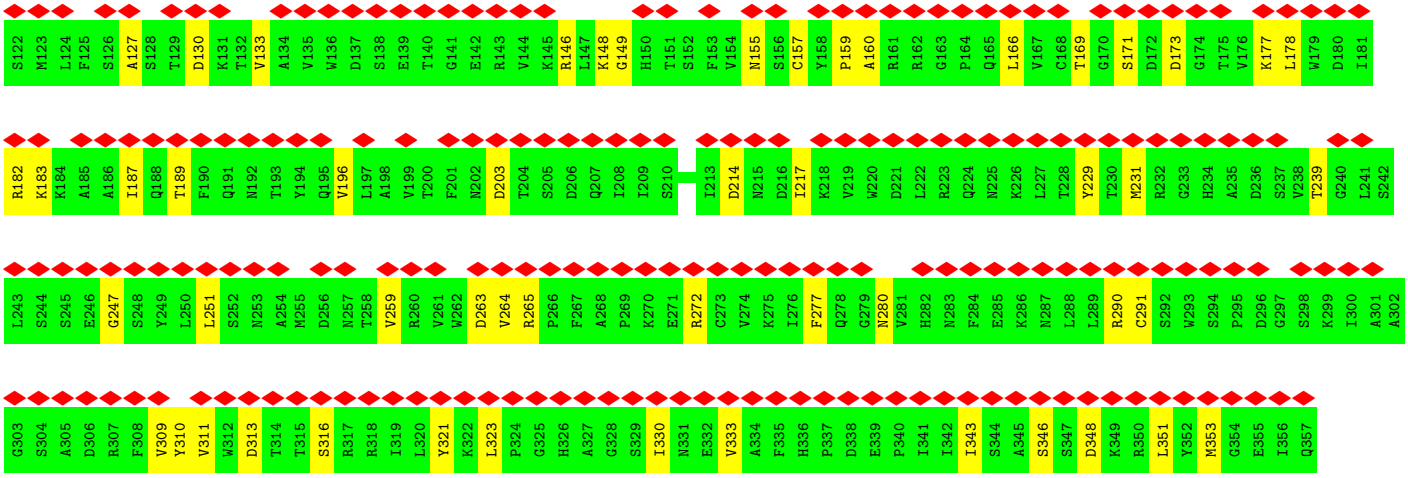




- Molecule 3: U5 small nuclear ribonucleoprotein 40 kDa protein

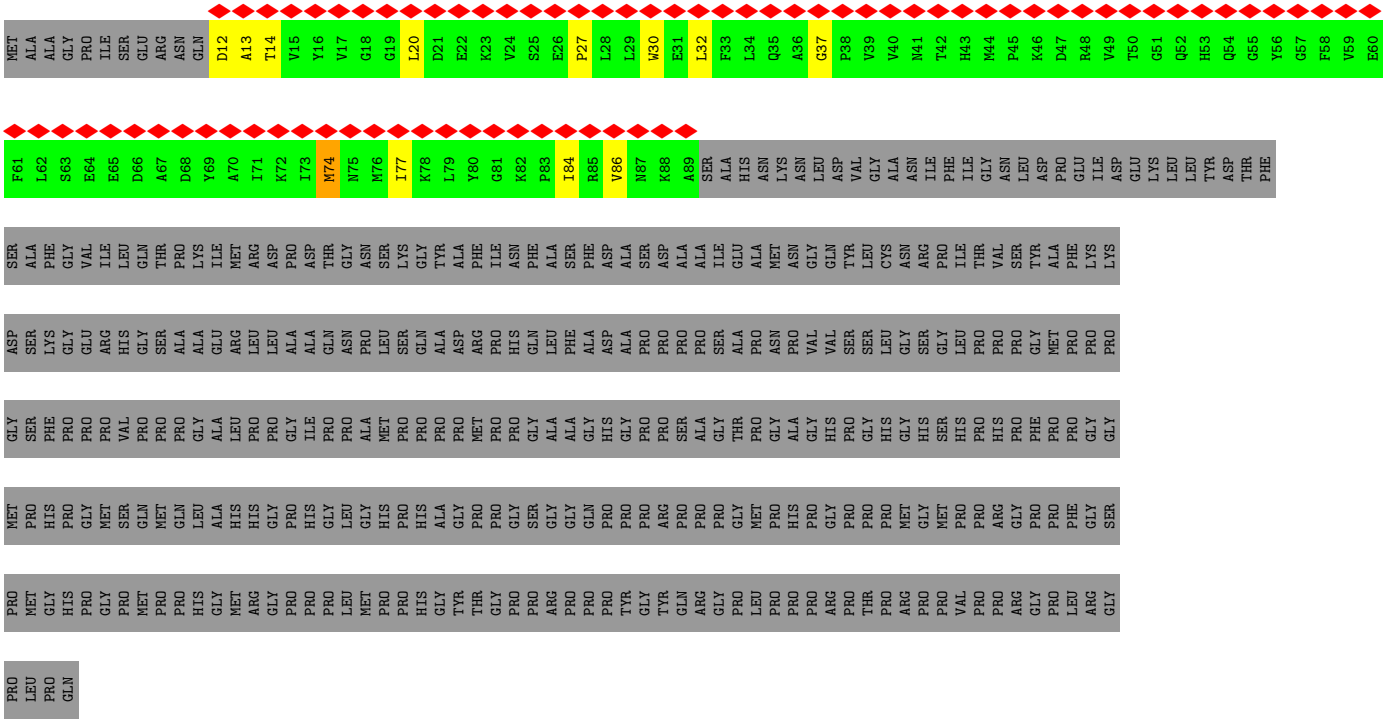
Chain 5O:





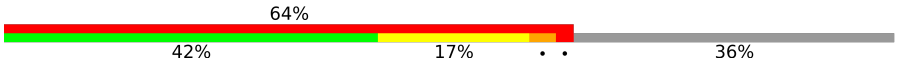
- Molecule 4: Splicing factor 3B subunit 4

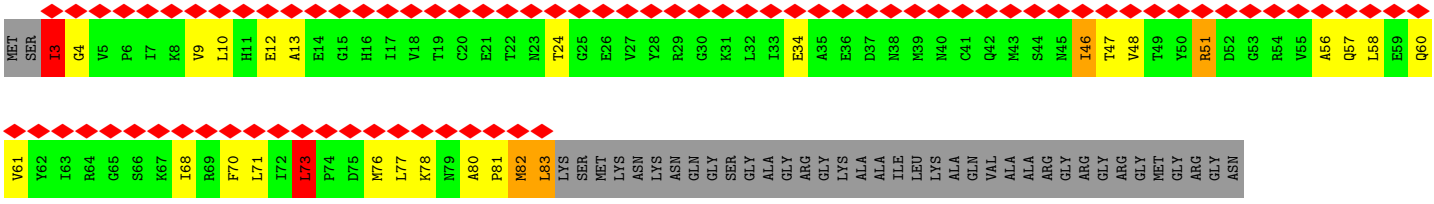
Chain B4:



- Molecule 5: Small nuclear ribonucleoprotein Sm D3

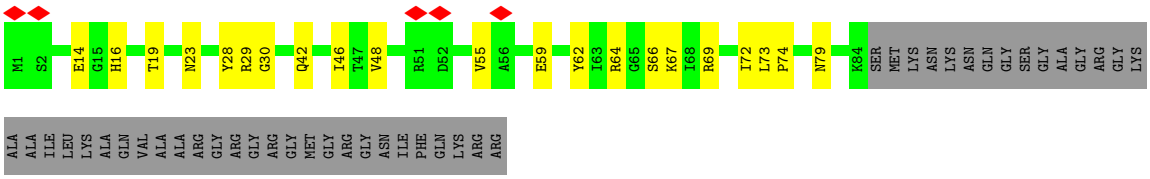
Chain 13:





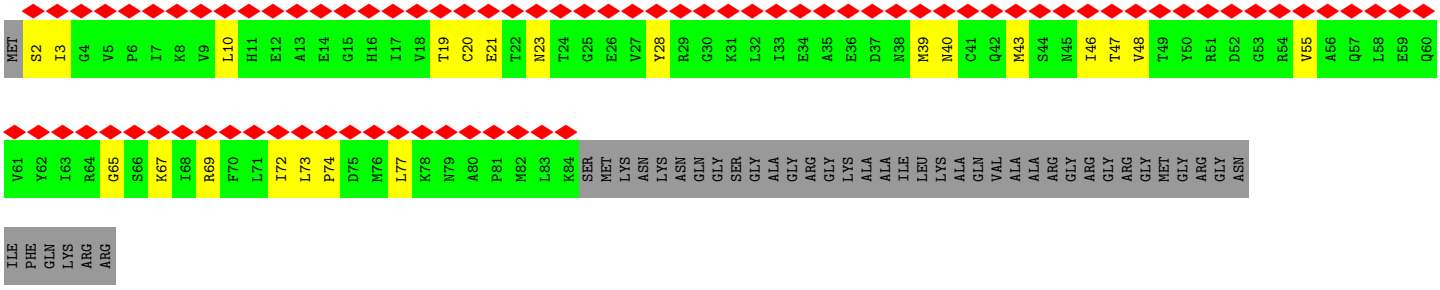
• Molecule 5: Small nuclear ribonucleoprotein Sm D3

Chain 53:



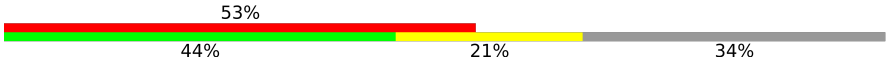
• Molecule 5: Small nuclear ribonucleoprotein Sm D3

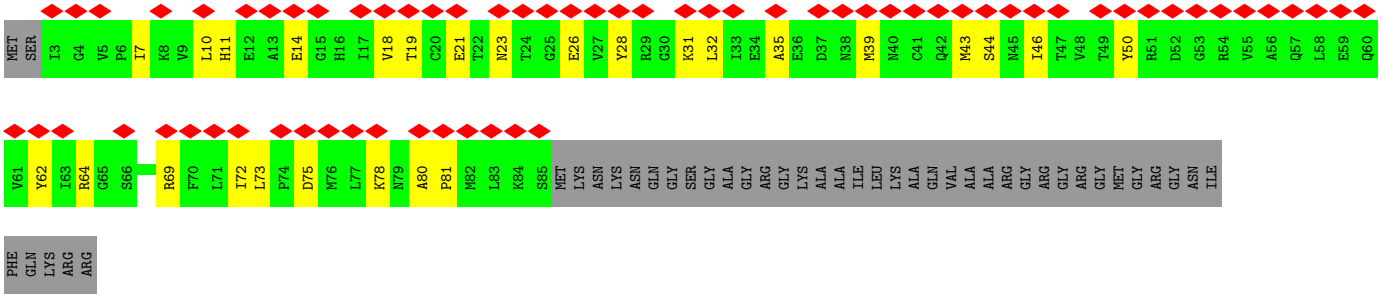
Chain 23:



• Molecule 5: Small nuclear ribonucleoprotein Sm D3

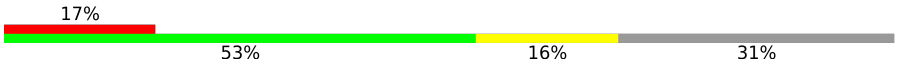
Chain 43:

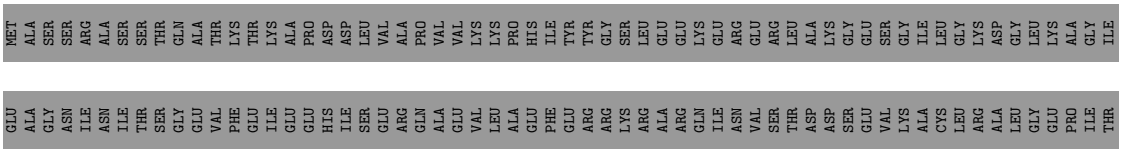


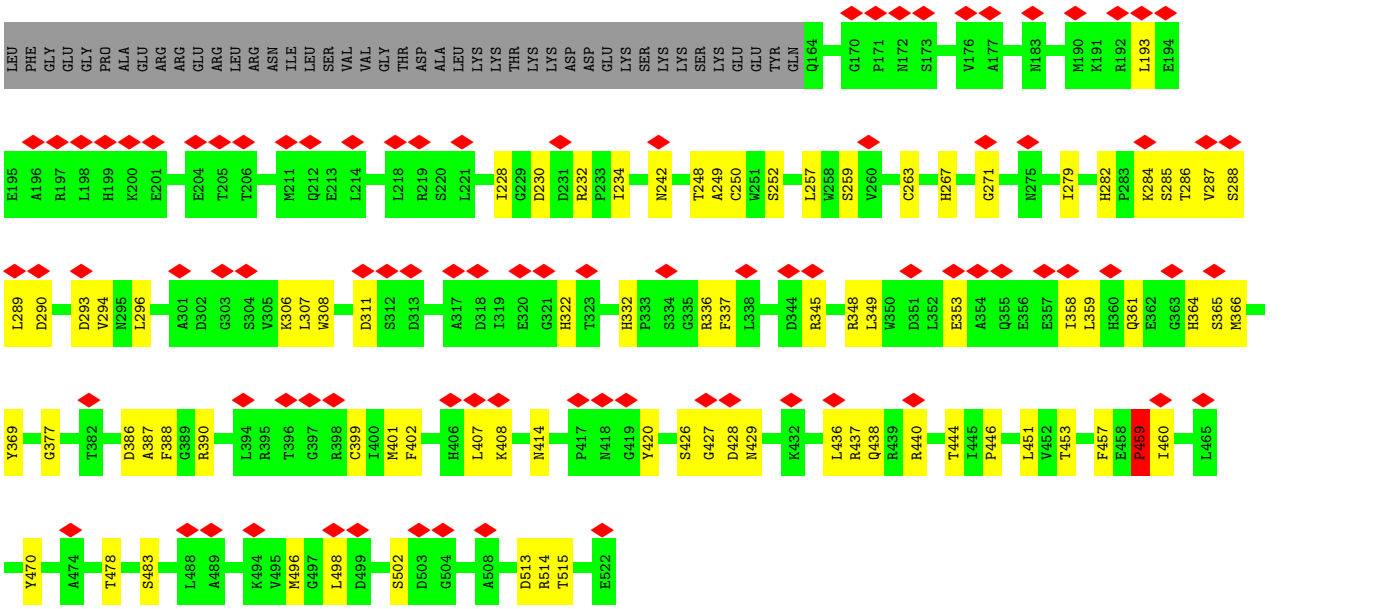


• Molecule 6: U4/U6 small nuclear ribonucleoprotein Prp4

Chain 4B:

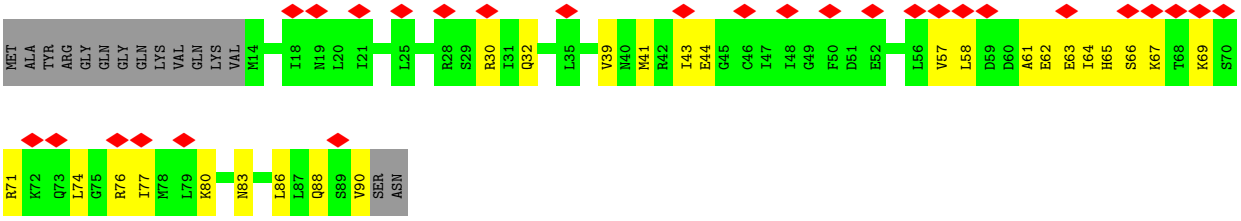






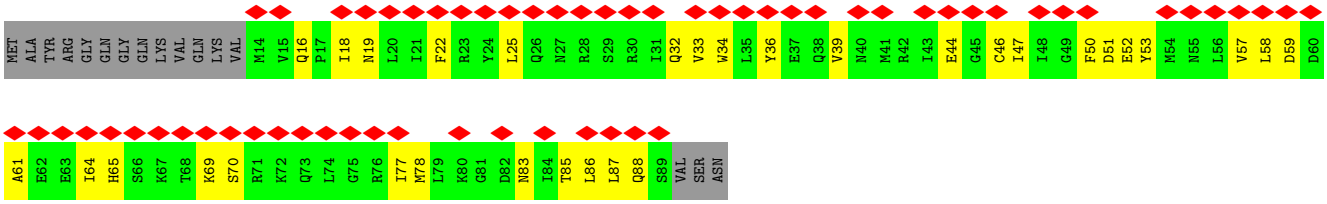
- Molecule 7: Small nuclear ribonucleoprotein E

Chain 5e:



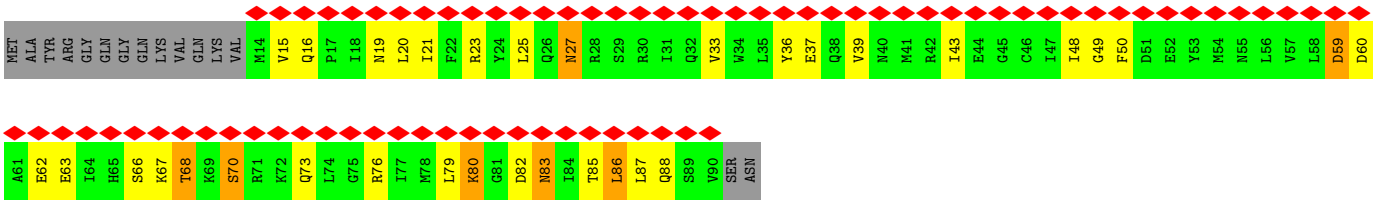
- Molecule 7: Small nuclear ribonucleoprotein E

Chain 4e:



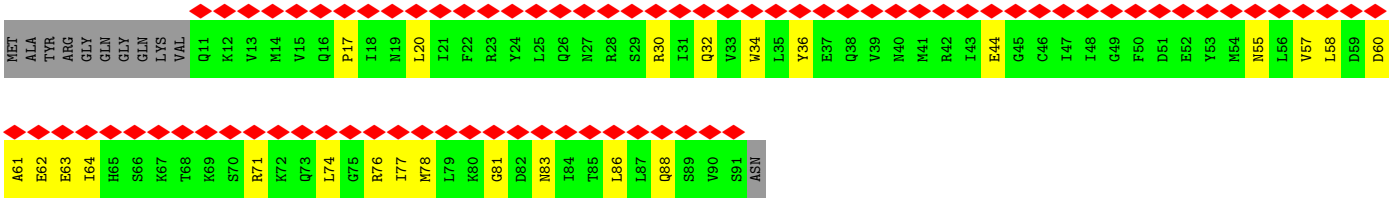
- Molecule 7: Small nuclear ribonucleoprotein E

Chain 1e:



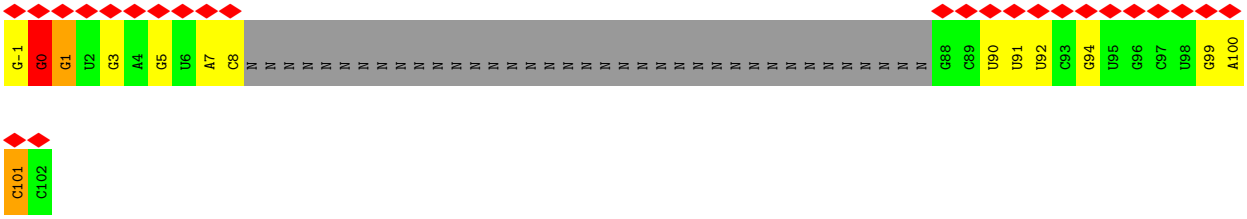
- Molecule 7: Small nuclear ribonucleoprotein E

Chain 2e:



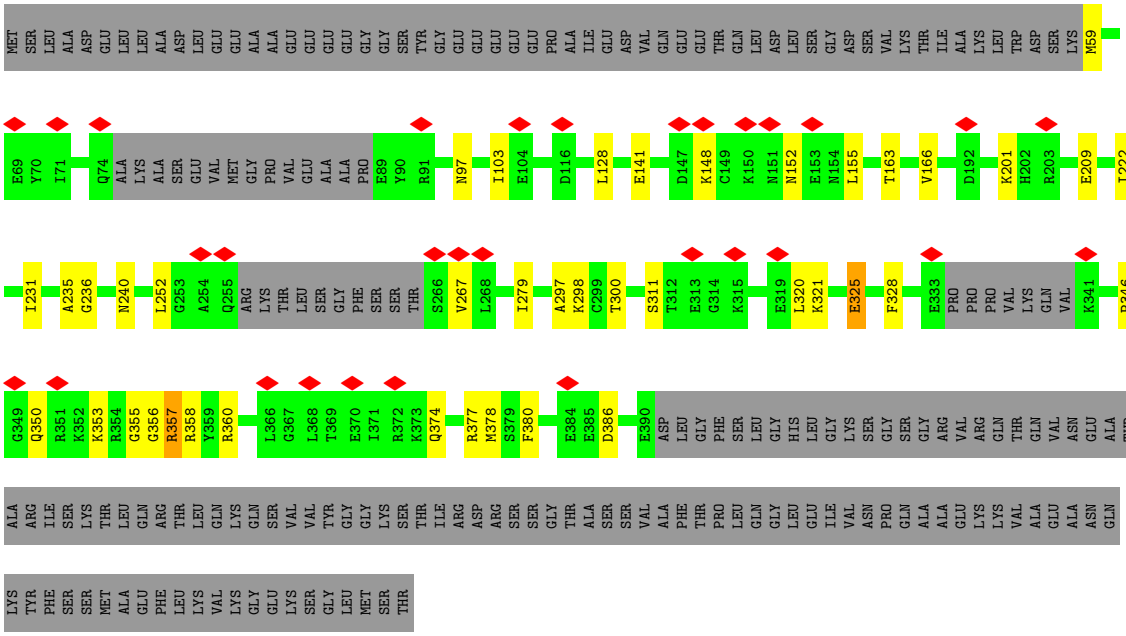
- Molecule 8: AdML pre-mRNA

Chain I:


G-1 G0 G1 U2 G3 A4 G5 U6 A7 C8 N N N

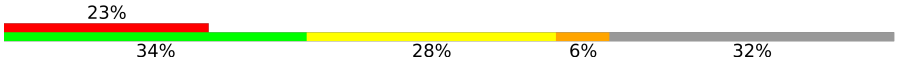
- Molecule 10: U4/U6 small nuclear ribonucleoprotein Prp31

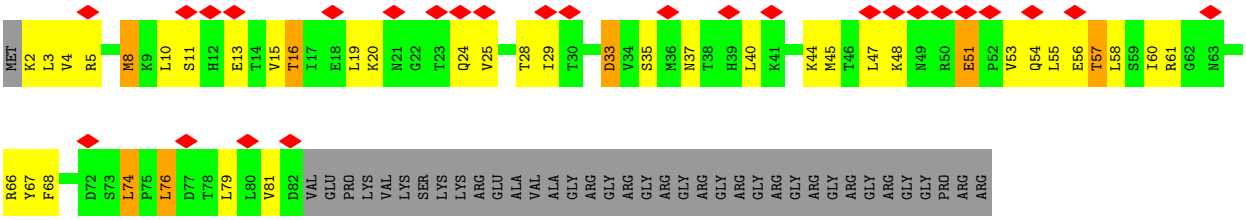
Chain 4C:



- Molecule 11: Small nuclear ribonucleoprotein Sm D1

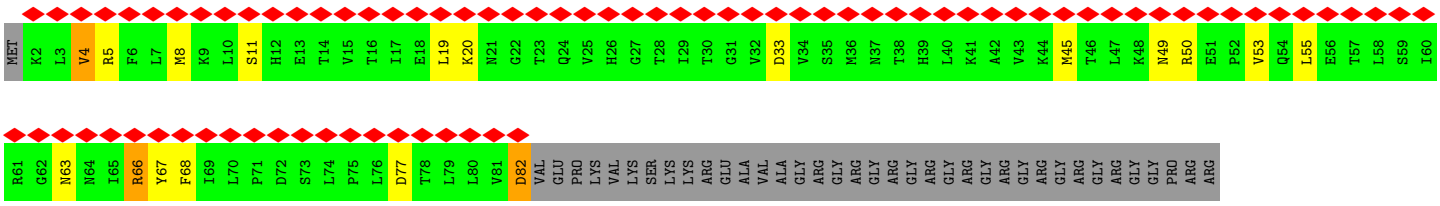
Chain 41:





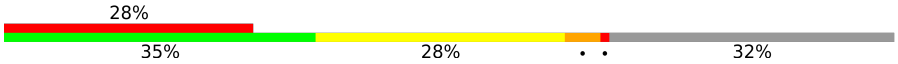
- Molecule 11: Small nuclear ribonucleoprotein Sm D1

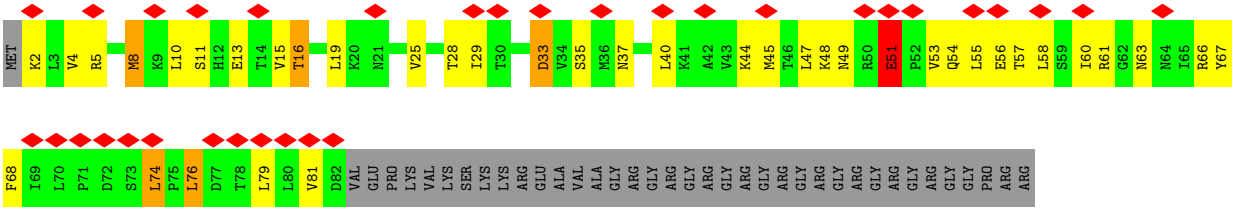
Chain 11:



- Molecule 11: Small nuclear ribonucleoprotein Sm D1

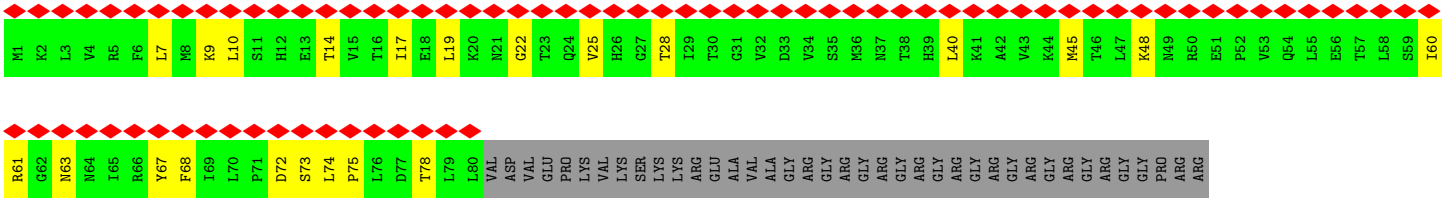
Chain 51:





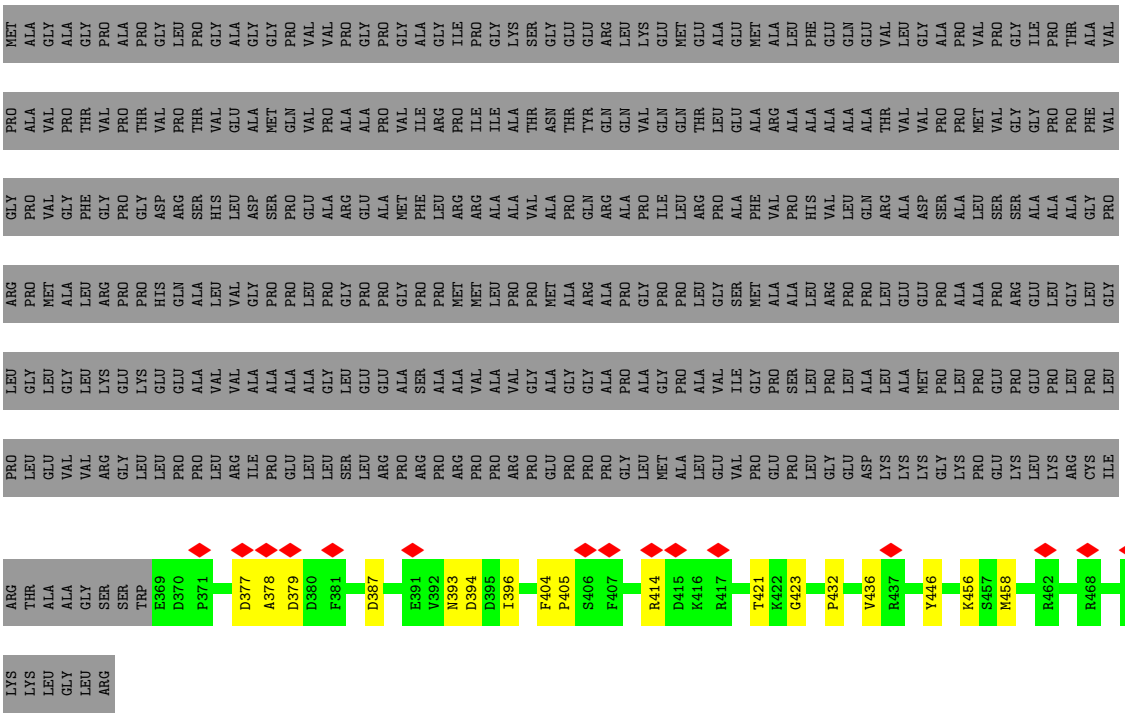
• Molecule 11: Small nuclear ribonucleoprotein Sm D1

Chain 21:



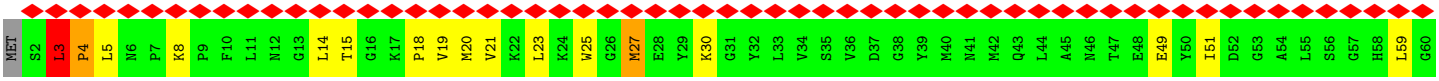
• Molecule 12: RNA-binding protein 42

Chain R:



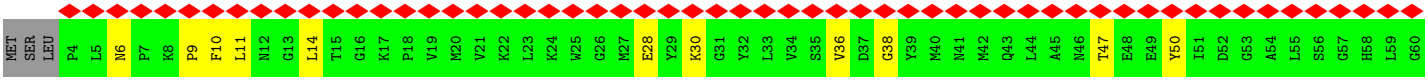
• Molecule 13: Small nuclear ribonucleoprotein F

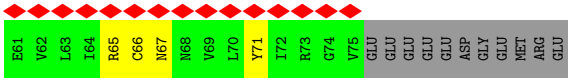
Chain 1f:



• Molecule 13: Small nuclear ribonucleoprotein F

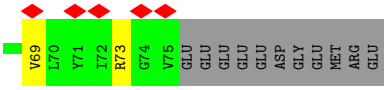
Chain 2f:





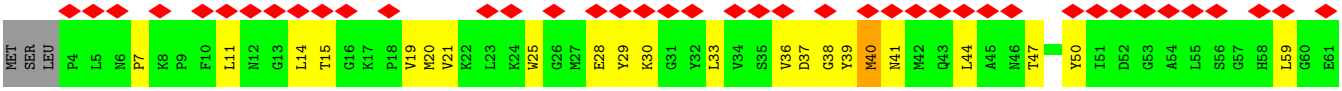
• Molecule 13: Small nuclear ribonucleoprotein F

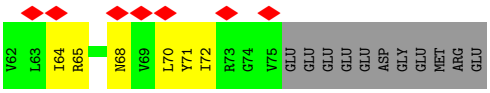
Chain 5f:



• Molecule 13: Small nuclear ribonucleoprotein F

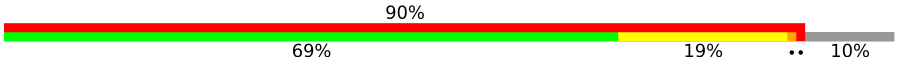
Chain 4f:

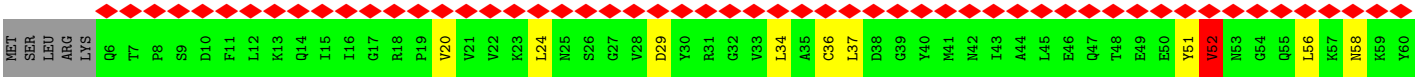


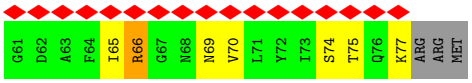


• Molecule 14: U6 snRNA-associated Sm-like protein LSm6

Chain 66:

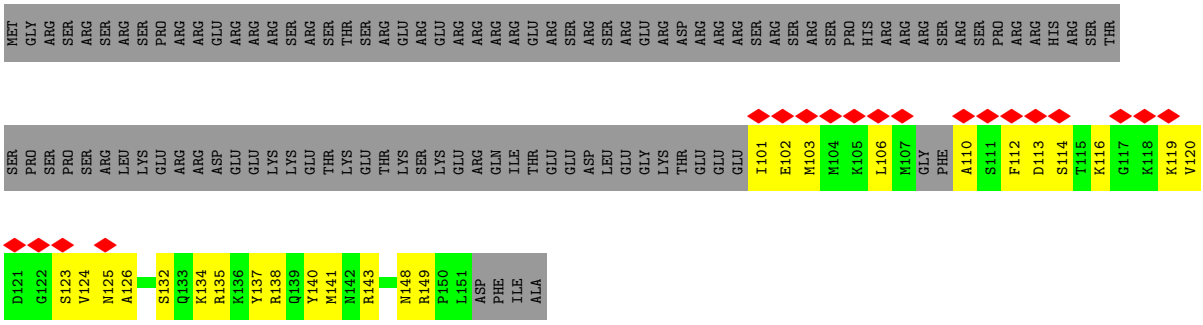






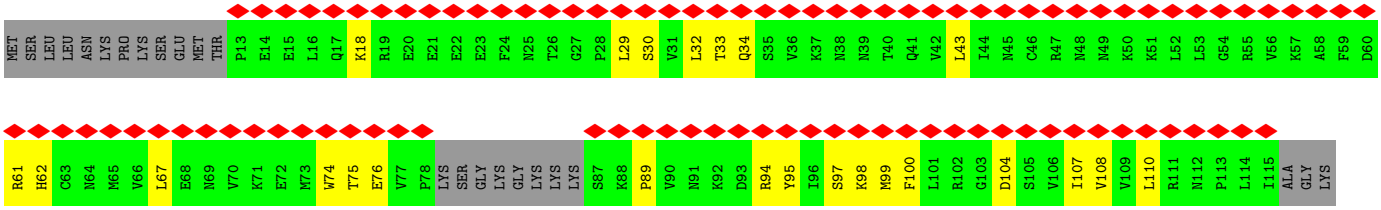
• Molecule 15: U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein

Chain X:



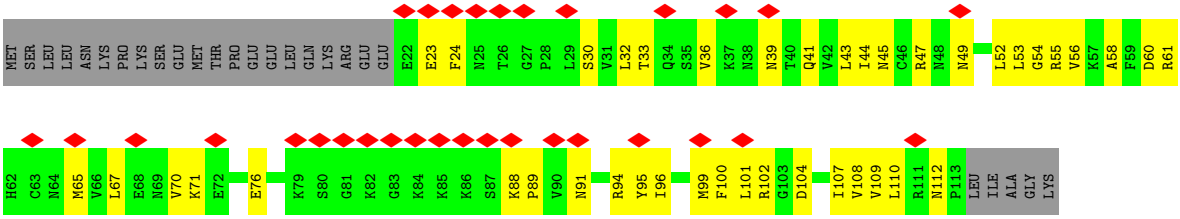
- Molecule 16: Small nuclear ribonucleoprotein Sm D2

Chain 22:



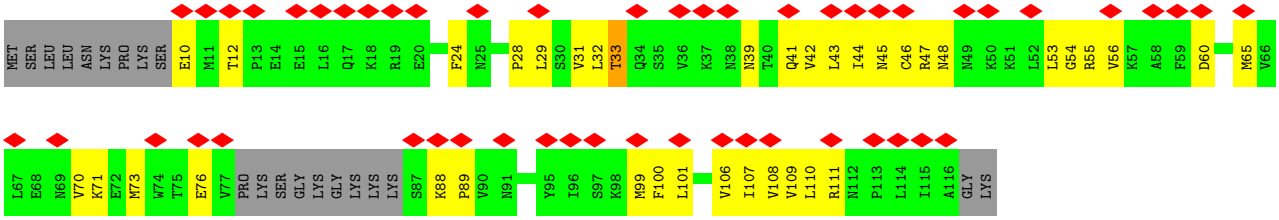
- Molecule 16: Small nuclear ribonucleoprotein Sm D2

Chain 42:



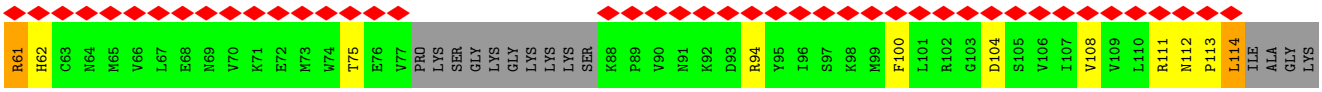
- Molecule 16: Small nuclear ribonucleoprotein Sm D2

Chain 52:



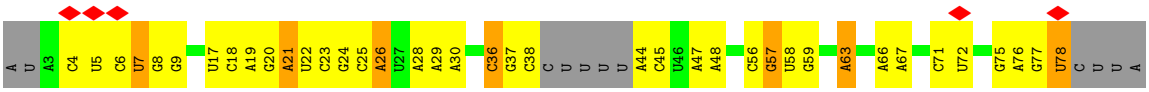
- Molecule 16: Small nuclear ribonucleoprotein Sm D2

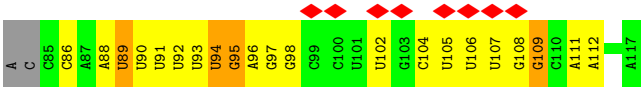
Chain 12:



• Molecule 17: U5 snRNA

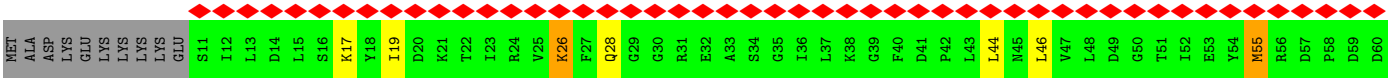
Chain 5:

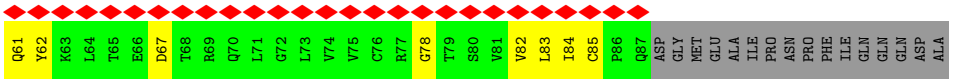




• Molecule 18: U6 snRNA-associated Sm-like protein LSm7

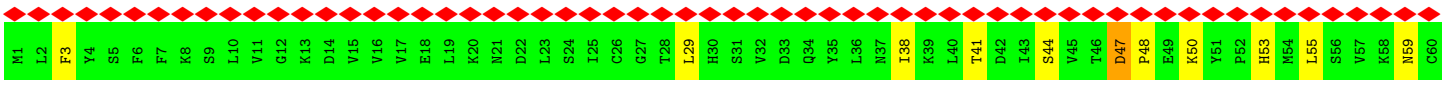
Chain 67:





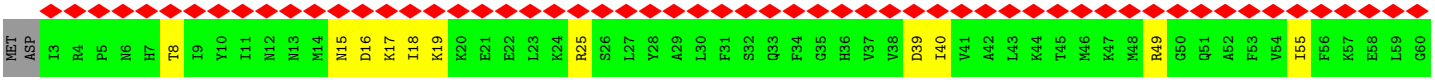
• Molecule 19: U6 snRNA-associated Sm-like protein LSm2

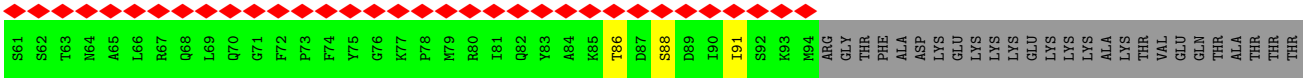
Chain 62:



• Molecule 20: U2 small nuclear ribonucleoprotein B"

Chain 2B:





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

- Molecule 21: Splicing factor 3A subunit 2

Chain A2: 69% 49% 16% • 31%

MET ASP PHE GLN HIS ARG PRO GLY GLY LYS THR GLY SER GLY GLY VAL ALA ALA SER SER SER GLU SER ASN ARG ASP ARG ARG GLU ARG LEU ARG GLN LEU ALA LEU GLU THR ILE ASP ILE ASN K42 D43 P44 Y45 F46 M47 K48 M49 H50 L51 G52 S53 Y54 E55 C56 K57 L58 C59 L60

T61 L62 H63 N64 N65 E66 G67 S68 Y69 L70 A71 H72 T73 Q74 G75 K76 K77 H78 Q79 T80 N81 L82 A83 R84 R85 ALA ALA LYS GLU ALA LYS GLU ALA PRO ALA GLN PRO ALA PRO GLU LYS VAL LYS V104 E105 V106 K107 K108 F109 V110 K111 I112 G113 R114 P115 G116 Y117 K118 V119 T120

K121 Q122 R123 D124 S125 E126 MET GLY GLN GLN SER LEU L133 F134 Q135 I136 D137 Y138 P139 E140 I141 A142 E143 G144 I145 M146 P147 R148 H149 R150 F151 M152 S153 A154 Y155 E156 Q157 R158 I159 E160 P161 P162 D163 R164 R165 W166 Q167 Y168 L169 L170 M171 A172 A173 E174 P175 Y176 E177 T178 I179 A180

F181 K182 V183 P184 S185 R186 E187 I188 D189 K190 A191 E192 G193 K194 F195 W196 T197 H198 W199 N200 R201 E202 T203 K204 Q205 F206 F207 L208 Q209

- Molecule 22: Splicing factor 3B subunit 2

Chain B2: 23% 17% 6% • 77%

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

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

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

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

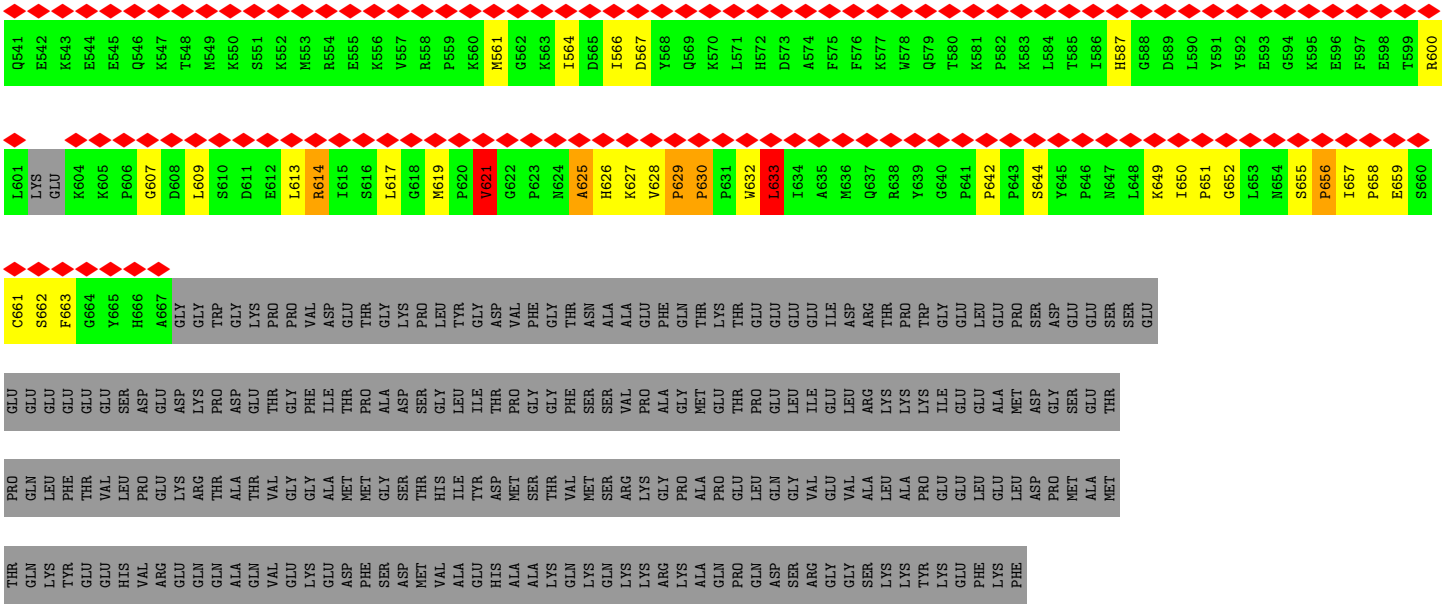
ALA PRO VAL PRO ARG PRO ARG GLY PRO PRO PRO PRO PRO GLY ASP GLU ASN ARG GLU MET ASP ASP PRO SER VAL GLY PRO LYS ILE PRO GLN ALA LEU GLU LYS ILE LEU GLN LEU LYS GLU SER ARG GLN GLU GLU MET ASN SER GLN GLN GLU GLU GLU GLU GLU MET GLU THR ASP ALA

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

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

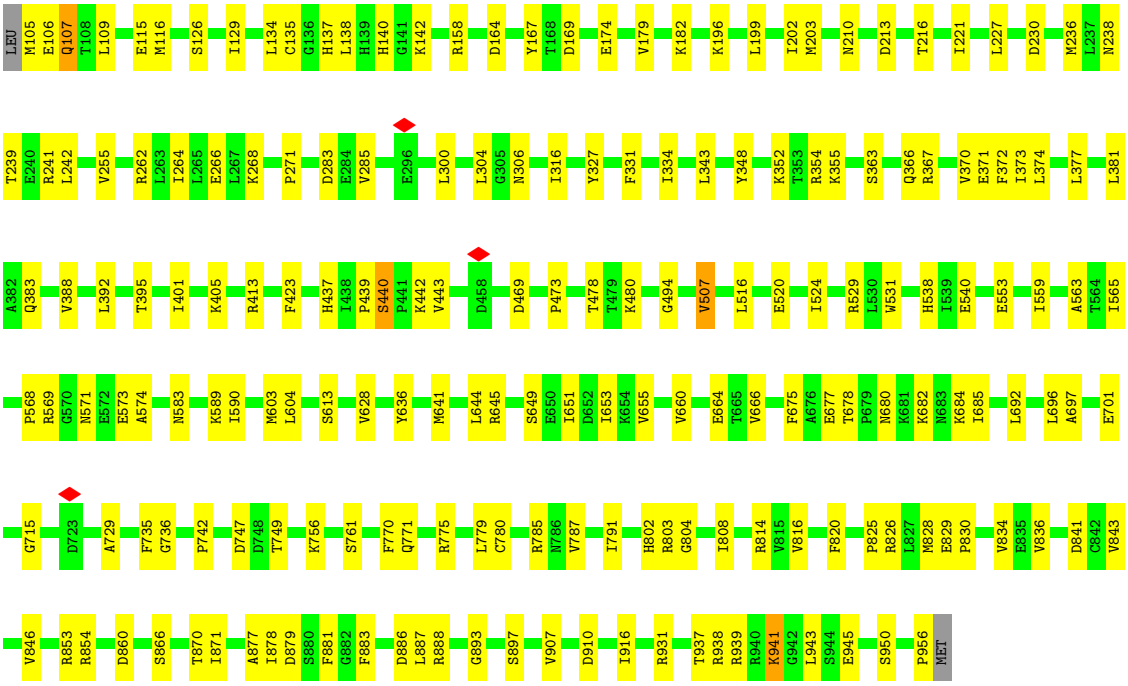
LYS LYS GLY PHE GLU GLU HIS LYS ASP SER ASP ASP SER SER ASP ASP GLU GLN GLU LYS LYS PRO GLU ALA PRO LYS LEU SER LYS LYS LYS LEU ARG ARG MET N458 R459 F460 T461 V462 A463 E464 L465 K466 Q467 L468 V469 A470 R471 P472 D473 V474 V475 E476 M477 H478 D479 V480

T481 A482 Q483 D484 P485 K486 L487 L488 V489 H490 L491 K492 A493 T494 R495 N496 S497 V498 P499 V500 P501 R502 H503 W504 C505 F506 K507 R508 K509 Y510 L511 Q512 G513 K514 R515 G516 I517 E518 K519 P520 P521 F522 E523 L524 P525 D526 F527 I528 K529 R530 T531 G532 I533 Q534 E535 M536 R537 E538 A539 L540



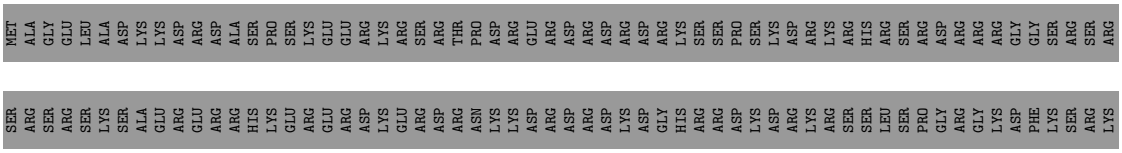
• Molecule 23: 116 kDa U5 small nuclear ribonucleoprotein component

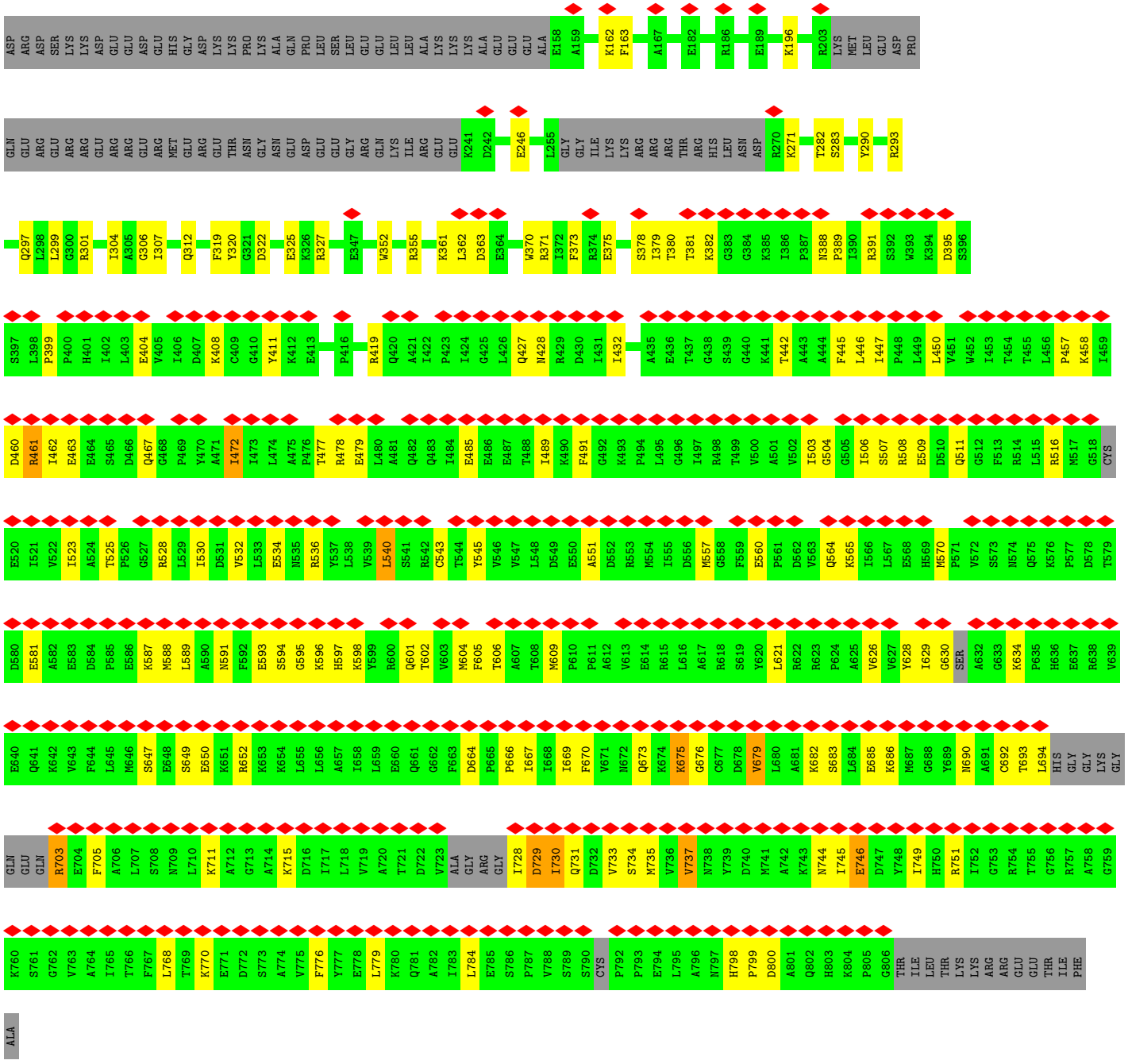
Chain 5C:



• Molecule 24: Probable ATP-dependent RNA helicase DDX23

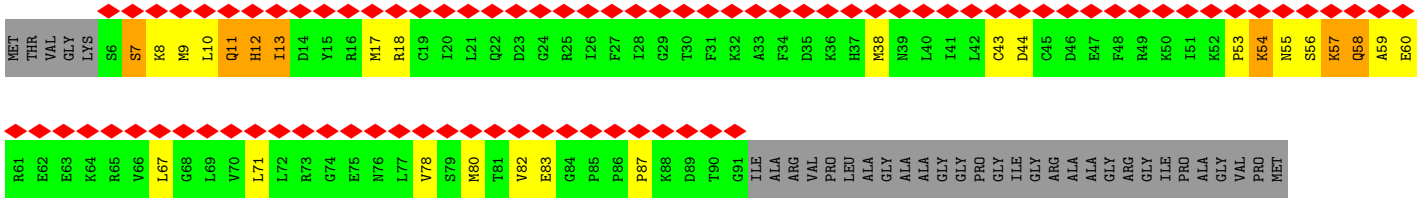
Chain 5X:

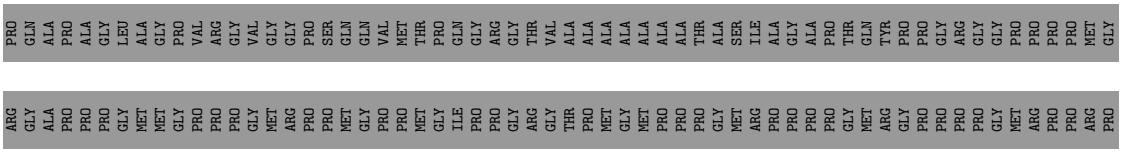




- Molecule 25: Small nuclear ribonucleoprotein-associated proteins B and B'

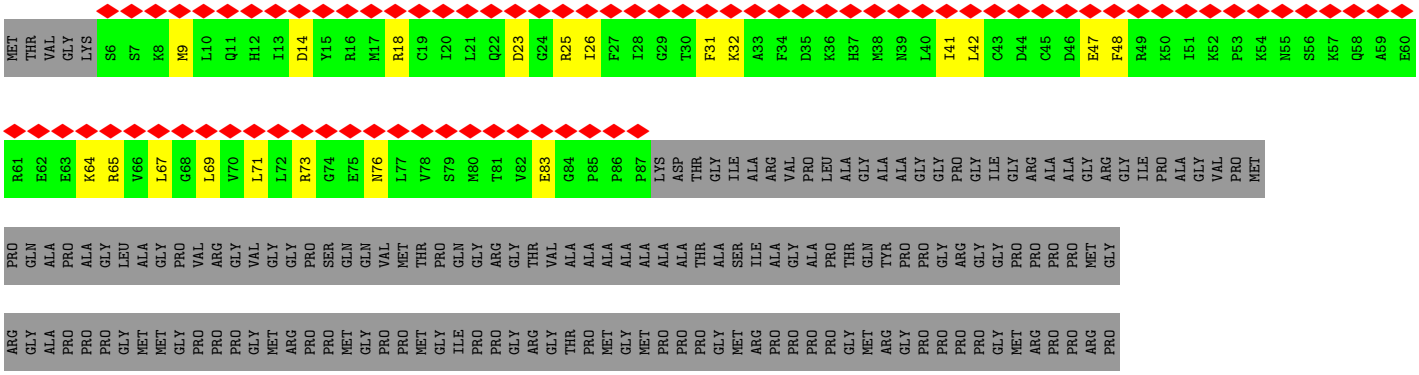
Chain 1b:





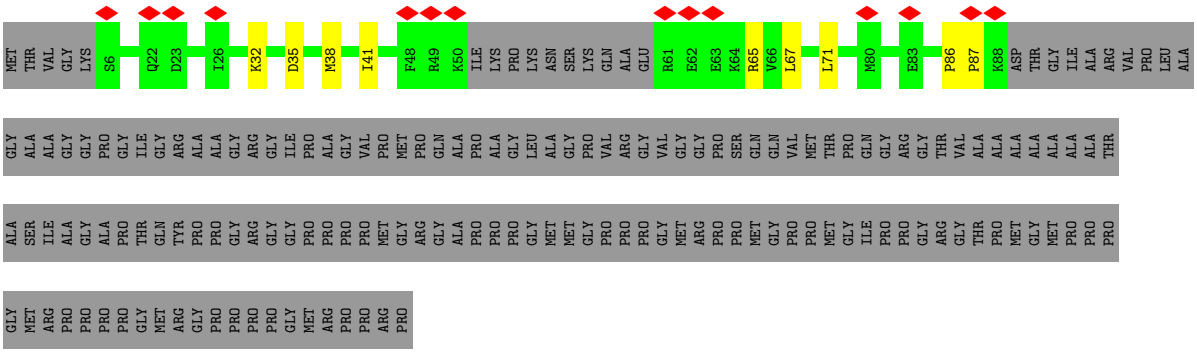
• Molecule 25: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain 2b:



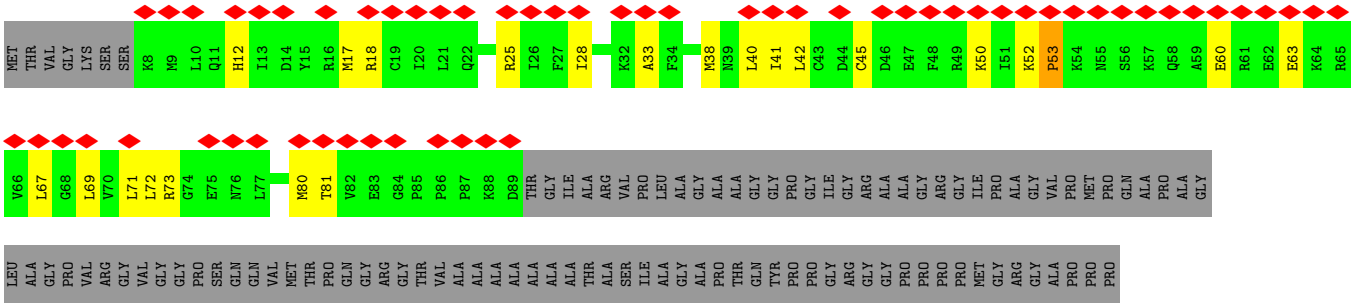
• Molecule 25: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain 5b:



• Molecule 25: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain 4b:



GLY MET MET GLY PRO PRO PRO GLY MET ARG PRO PRO PRO MET GLY PRO PRO MET GLY ILE PRO PRO GLY ARG GLY THR PRO MET GLY MET PRO PRO PRO GLY MET ARG PRO PRO PRO PRO GLY MET ARG GLY PRO PRO PRO PRO GLY MET ARG PRO PRO ARG PRO

- Molecule 26: Splicing factor 3B subunit 5

Chain B5:

MET THR ASP ARG TYR THR ILE HIS SER GLN LEU E12 H13 L14 Q15 S16 K17 Y18 I19 G20 T21 G22 H23 A24 D25 T26 T27 K28 W29 E30 W31 L32 V33 N34 Q35 H36 R37 D38 S39 Y40 C41 S42 Y43 M44 G45 H46 F47 D48 L49 L50 N51 Y52 F53 A54 I55 A56 E57 N58 E59 S60

K61 A62 R63 V64 R65 F66 N67 L68 M69 E70 K71 M72 L73 Q74 P75 C76 G77 P78 P79 A80 ASP LYS PRO GLU GLU ASN

- Molecule 27: U1 small nuclear ribonucleoprotein A

Chain 1A:

MET ALA VAL PRO GLU T6 R7 P8 N9 H10 T11 I12 Y13 I14 N15 N16 L17 N18 E19 K20 I21 K22 K23 D24 E25 L26 K27 K28 S29 L30 Y31 A32 I33 F34 S35 Q36 F37 G38 Q39 I40 L41 D42 I43 L44 V45 S46 R47 S48 L49 K50 M51 R52 G53 Q54 A55 F56 V57 I58 F59 K60

E61 V62 S63 S64 A65 T66 N67 A68 L69 R70 S71 M72 Q73 G74 F75 P76 F77 Y78 D79 K80 P81 M82 R83 I84 Q85 Y86 A87 K88 T89 D90 S91 D92 I93 I94 A95 K96 M97 K98 G99 T100 F101 V102 E103 ARG ASP ARG LYS ARG GLU LYS ARG LYS PRO LYS SER GLN GLU THR PRO ALA

THR LYS LYS ALA VAL GLN GLY GLY GLY ALA THR PRO VAL VAL GLY ALA VAL GLN GLY PRO VAL PRO GLY MET PRO PRO MET THR GLN ALA PRO ARG ILE MET HIS HIS MET PRO GLY GLN PRO PRO TYR MET PRO PRO PRO GLY MET ILE PRO PRO PRO GLY LEU ALA PRO GLY GLN ILE

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

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

- Molecule 28: U4/U6.U5 tri-snRNP-associated protein 1

Chain S:

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

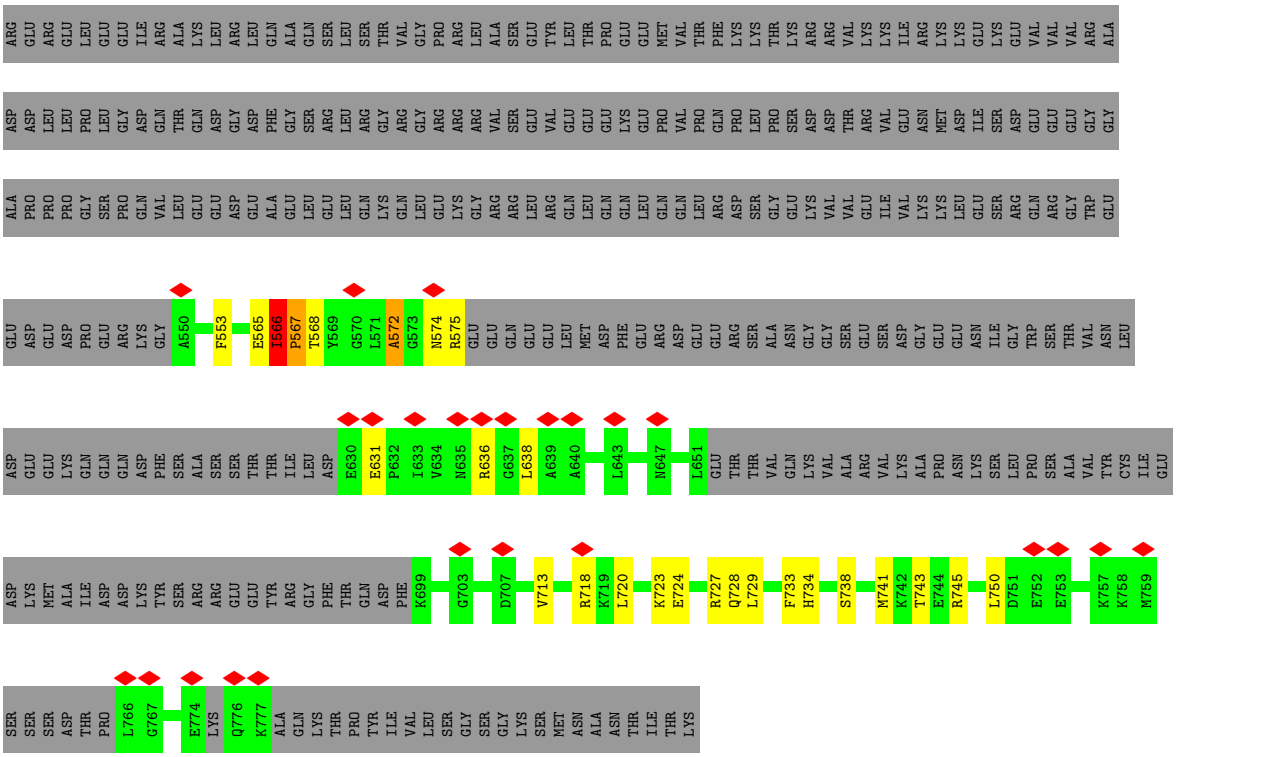
GLY GLU ARG GLY SER GLY ARG ARG GLY ALA GLU ALA GLU ALA ARG SER THR HIS GLY ARG GLU SER ARG SER GLN ALA GLU PRO SER GLU ARG ARG VAL LYS ARG GLU LYS ARG ASP ASP GLY TYR GLU ALA ALA ALA SER SER LYS THR SER SER GLY ASP ALA SER SER LEU SER ILE

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

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

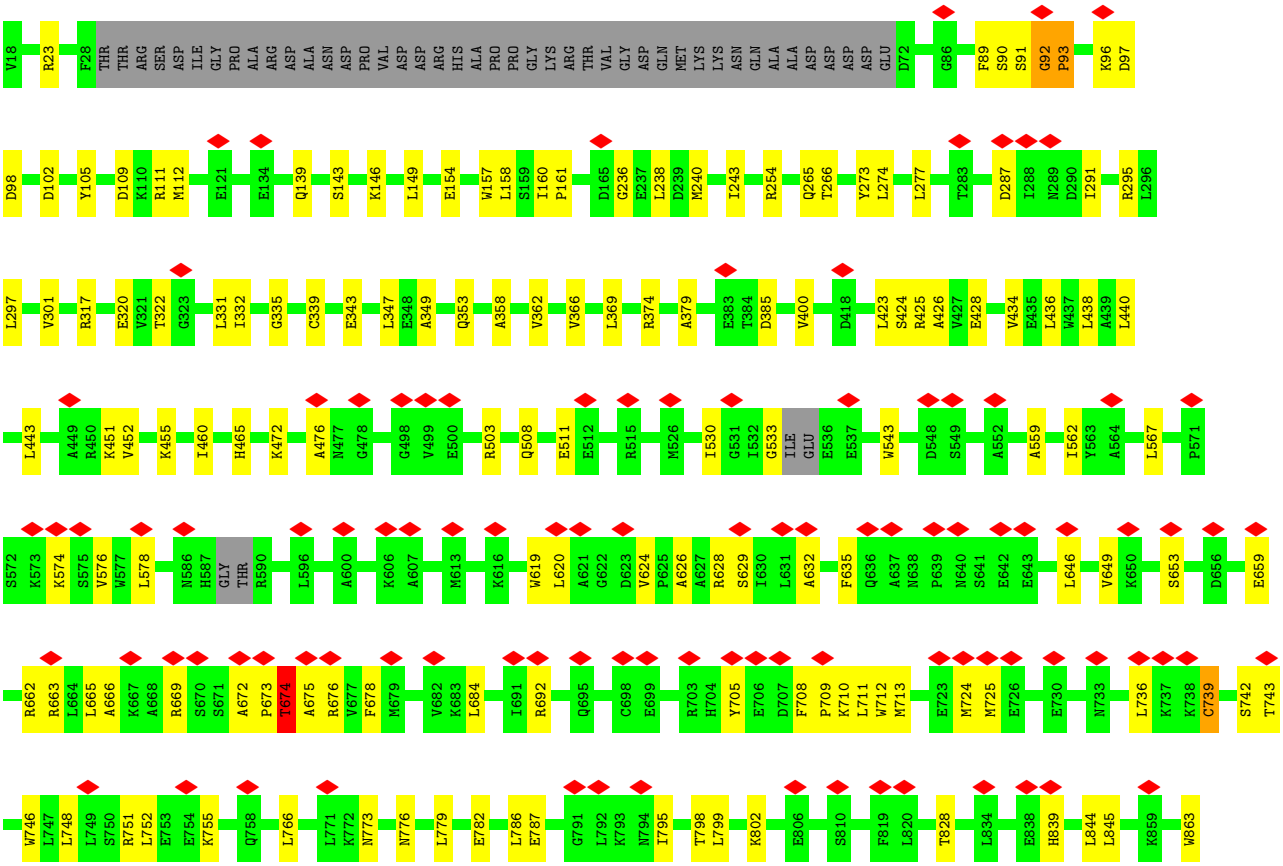
PHE GLY GLN ARG ARG GLN ASP LEU TYR SER ALA ARG ASP LEU GLN GLY LEU THR VAL GLU HIS ALA ILE ASP SER PHE ARG GLU GLY GLU THR MET ILE LEU THR LEU LYS ASP LYS GLY VAL LEU GLN GLU GLU GLU ASP VAL LEU VAL ASN VAL ASN LEU VAL ASP LYS GLU ARG ALA

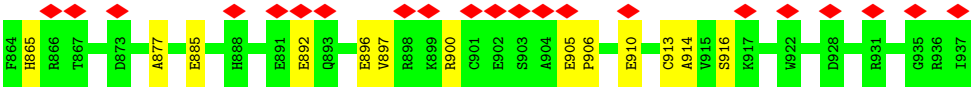
GLU LYS ASN VAL GLU LEU ARG LYS LYS PRO ASP TYR LEU PRO TYR ALA GLU ASP GLU SER VAL ASP ASP LEU ALA GLN GLN LYS PRO ARG SER ILE LEU SER LYS TYR ASP GLU GLU LEU GLU GLY GLU ARG PRO HIS SER PHE ARG LEU GLU GLN GLY GLY THR ALA ASP GLY LEU



- Molecule 29: Pre-mRNA-processing factor 6

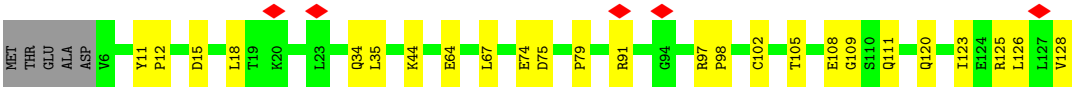
Chain 5J:





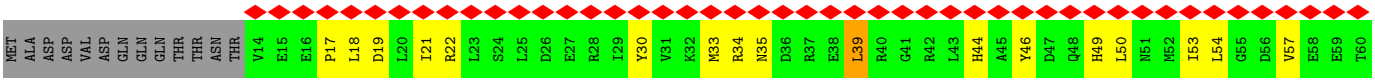
- Molecule 30: NHP2-like protein 1

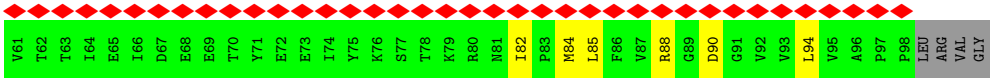
Chain 4D:



- Molecule 31: U6 snRNA-associated Sm-like protein LSm3

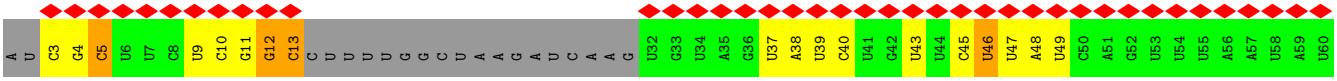
Chain 63:

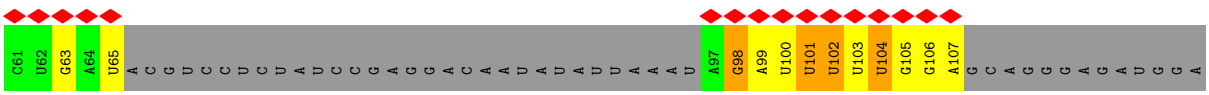


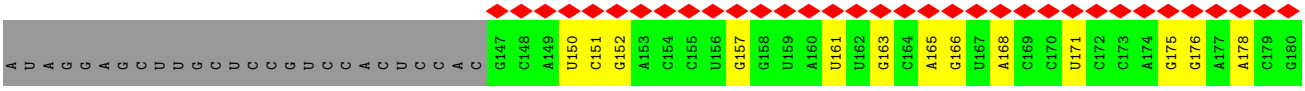


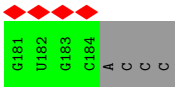
- Molecule 32: U2 snRNA

Chain 2:



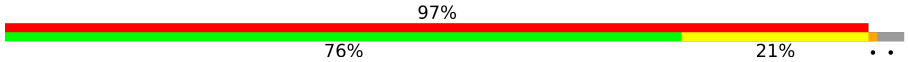


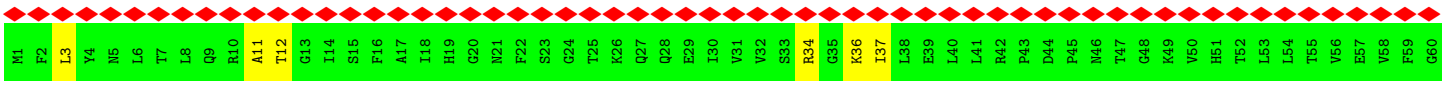


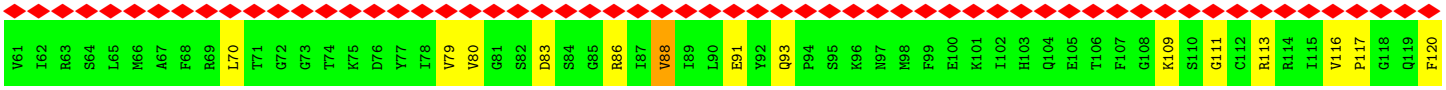


- Molecule 33: Splicing factor 3B subunit 3

Chain B3:







L121 A122 V123 D124 P125 K126 G127 R128 A129 V130 M131 I132 S133 A134 I135 E136 K137 Q138 K139 L140 V141 Y142 I143 L144 N145 R146 D147 A148 A149 A150 R151 L152 T153 I154 S155 S156 P157 L158 E159 A160 H161 K162 A163 N164 T165 L166 V167 Y168 H169 V170 V171 G172 V173 D174 V175 G176 F177 E178 N179 P180

M181 F182 A183 C184 L185 E186 M187 D188 Y189 E190 E191 A192 D193 N194 D195 P196 T197 G198 E199 A200 A201 A202 N203 T204 Q205 Q206 T207 L208 T209 F210 Y211 E212 L213 D214 L215 G216 L217 N218 H219 V220 V221 R222 K223 Y224 S225 E226 P227 L228 E229 E230 H231 G232 N233 F234 L235 I236 T237 V238 P239 G240

G241 S242 D243 G244 P245 S246 G247 V248 L249 I250 C251 S252 E253 N254 Y255 I256 T257 Y258 K259 N260 F261 G262 D263 Q264 P265 D266 I267 R268 C269 P270 I271 P272 R273 R274 R275 N276 D277 L278 D279 D280 P281 E282 R283 G284 M285 I286 F287 V288 C289 S290 A291 T292 H293 K294 T295 K296 S297 M298 F299 F300

F301 L302 A303 Q304 T305 E306 Q307 G308 D309 I310 F311 K312 I313 T314 L315 E316 T317 D318 E319 D320 M321 V322 T323 E324 I325 R326 L327 K328 Y329 F330 D331 T332 V333 P334 V335 A336 A337 A338 M339 C340 V341 L342 K343 T344 G345 F346 L347 F348 V349 A350 S351 E352 F353 G354 N355 H356 Y357 L358 Y359 Q360

I361 A362 H363 L364 G365 D366 D367 D368 E369 E370 P371 E372 F373 S374 S375 A376 M377 P378 L379 E380 E381 G382 D383 T384 F385 F386 F387 Q388 P389 R390 P391 L392 K393 N394 L395 V396 L397 V398 D399 E400 L401 D402 S403 L404 S405 P406 I407 L408 F409 C410 Q411 I412 A413 D414 L415 A416 N417 E418 D419 T420

P421 Q422 L423 Y424 V425 A426 C427 G428 R429 G430 P431 R432 S433 S434 L435 R436 V437 L438 R439 H440 G441 L442 E443 V444 S445 E446 M447 A448 V449 S450 E451 L452 P453 G454 N455 P456 N457 A458 V459 W460 T461 V462 R463 R464 H465 I466 E467 D468 E469 F470 D471 A472 Y473 I474 I475 V476 S477 F478 V479 N480

A481 T482 L483 V484 L485 S486 I487 G488 E489 T490 V491 E492 E493 V494 T495 D496 S497 G498 F499 L500 G501 T502 T503 P504 T505 L506 S507 C508 S509 L510 L511 G512 D513 D514 A515 L516 V517 Q518 V519 Y520 P521 D522 G523 I524 R525 H526 I527 R528 A529 D530 K531 R532 V533 N534 E535 W536 K537 T538 P539 G540

K541 K542 T543 I544 V545 K546 C547 A548 V549 N550 Q551 R552 Q553 V554 V555 I556 A557 L558 T559 G560 G561 E562 L563 V564 Y565 F566 E567 M568 D569 P570 S571 G572 Q573 L574 N575 E576 Y577 T578 E579 R580 K581 E582 M583 S584 A585 D586 V587 V588 C589 M590 S591 L592 A593 N594 V595 P596 P597 G598 E599 Q600

R601 S602 R603 F604 L605 A606 V607 G608 L609 V610 D611 N612 T613 V614 R615 I616 I617 S618 L619 D620 P621 S622 D623 C624 L625 Q626 P627 L628 S629 M630 Q631 A632 L633 P634 A635 Q636 P637 E638 S639 L640 C641 I642 V643 E644 M645 GLY GLY THR GLU LYS GLN ASP GLU LEU GLY GLU ARG GLY SER ILE

GLY F662 L663 Y664 L665 N666 I667 G668 L669 Q670 N671 G672 V673 L674 L675 R676 T677 V678 L679 D680 P681 V682 T683 G684 D685 L686 S687 D688 T689 R690 T691 ARG TYR LEU G695 S696 R697 P698 V699 K700 L701 F702 R703 V704 R705 M706 Q707 G708 Q709 E710 A711 V712 L713 A714 M715 S716 S717 R718 S719 W720

L721 S722 Y723 S724 Y725 Q726 S727 R728 F729 H730 L731 T732 P733 L734 S735 Y736 E737 T738 L739 E740 F741 A742 S743 G744 F745 A746 S747 E748 Q749 C750 P751 E752 G753 I754 V755 A756 I757 S758 T759 N760 T761 L762 R763 I764 L765 A766 L767 E768 K769 L770 G771 A772 V773 F774 N775 Q776 V777 A778 F779 P780

L781 Q782 Y783 T784 P785 R786 K787 F788 V789 T790 H791 F792 E793 S794 N795 N796 L797 T798 T799 I800 E801 T802 D803 H804 N805 A806 Y807 T808 E809 A810 T811 K812 A813 Q814 R815 K816 Q817 Q818 M819 A820 E821 E822 M823 V824 E825 A826 A827 G828 E829 D830 GLU ARG GLU L834 A835 A836 E837 M838 A839 A840

A841 F842 L843 N844 E845 N846 L847 P848 E849 S850 I851 F852 G853 A854 P855 K856 A857 G858 N859 G860 Q861 W862 A863 S864 V865 I866 R867 V868 M869 N870 P871 I872 Q873 G874 N875 T876 L877 D878 L879 V880 Q881 L882 E883 Q884 N885 E886 A887 A888 F889 S890 V891 A892 V893 C894 R895 F896 S897 N898 T899 G900

E901 D902 W903 Y904 V905 L906 V907 G908 V909 A910 K911 D912 L913 I914 L915 N916 P917 R918 S919 V920 A921 G922 G923 F924 V925 Y926 T927 Y928 K929 L930 V931 N932 N933 G934 E935 K936 L937 E938 F939 L940 H941 K942 T943 P944 V945 E946 E947 V948 P949 A950 A951 I952 A953 P954 F955 Q956 G957 R958 V959 L960

I961 G962 V963 G964 K965 L966 L967 R968 V969 Y970 D971 L972 G973 K974 K975 K976 L977 L978 R979 K980 C981 E982 N983 K984 H985 I986 A987 N988 Y989 I990 S991 G992 I993 Q994 T995 I996 G997 H998 R999 V1000 I1001 V1002 S1003 D1004 V1005 Q1006 E1007 S1008 F1009 I1010 W1011 V1012 R1013 Y1014 K1015 R1016 N1017 E1018 N1019 Q1020

L1021 I1022 I1023 F1024 A1025 D1026 D1027 T1028 Y1029 P1030 R1031 W1032 V1033 T1034 T1035 A1036 S1037 L1038 L1039 D1040 Y1041 D1042 T1043 V1044 A1045 G1046 A1047 D1048 K1049 F1050 G1051 N1052 I1053 C1054 V1055 V1056 R1057 L1058 P1059 P1060 N1061 T1062 N1063 D1064 E1065 V1066 D1067 E1068 ASP PRO THR GLY ASN LYS ALA LEU TRP D1078 R1079 G1080

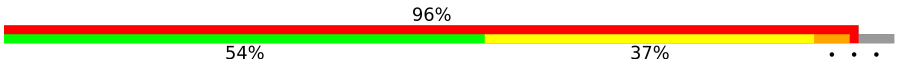
L1081 L1082 N1083 G1084 A1085 S1086 Q1087 K1088 A1089 E1090 V1091 I1092 M1093 N1094 Y1095 H1096 V1097 G1098 E1099 T1100 V1101 L1102 S1103 L1104 Q1105 K1106 T1107 T1108 L1109 I1110 P1111 G1112 G1113 S1114 E1115 S1116 L1117 V1118 Y1119 T1120 T1121 L1122 S1123 G1124 G1125 I1126 G1127 I1128 L1129 V1130 P1131 F1132 T1133 S1134 H1135 E1136 D1137 H1138 D1139 F1140

F1141 Q1142 H1143 V1144 E1145 M1146 H1147 L1148 R1149 S1150 E1151 H1152 P1153 P1154 L1155 C1156 G1157 R1158 D1159 H1160 L1161 S1162 F1163 R1164 S1165 Y1166 Y1167 F1168 P1169 V1170 K1171 N1172 V1173 I1174 D1175 G1176 D1177 L1178 C1179 E1180 Q1181 F1182 N1183 S1184 M1185 E1186 P1187 N1188 K1189 Q1190 K1191 N1192 V1193 S1194 E1195 E1196 L1197 D1198 R1199 T1200

P1201 P1202 E1203 V1204 S1205 K1206 K1207 L1208 E1209 D1210 I1211 R1212 T1213 R1214 Y1215 A1216 F1217

- Molecule 34: Small nuclear ribonucleoprotein G

Chain 1g:

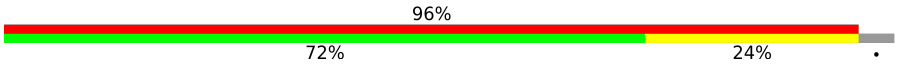


MET SER LYS A4 H5 P6 P7 E8 L9 K10 K11 F12 M13 D14 K15 K16 L17 S18 L19 K20 L21 N22 G23 G24 R25 H26 V27 Q28 G29 I30 L31 R32 G33 F34 D35 P36 F37 M38 N39 L40 V41 I42 D43 E44 C45 V46 E47 M48 A49 T50 S51 G52 Q53 Q54 N55 N56 I57 G58 M59 V60

V61 I62 R63 G64 N65 S66 I67 I68 M69 L70 E71 A72 L73 E74 R75 V76

- Molecule 34: Small nuclear ribonucleoprotein G

Chain 2g:



MET SER LYS A4 H5 P6 P7 E8 L9 K10 K11 F12 M13 D14 K15 K16 L17 S18 L19 K20 L21 N22 G23 G24 R25 H26 V27 Q28 G29 I30 L31 R32 G33 F34 D35 P36 F37 M38 N39 L40 V41 I42 D43 E44 C45 V46 E47 M48 A49 T50 S51 G52 Q53 Q54 N55 N56 I57 G58 M59 V60

V61 I62 R63 G64 N65 S66 I67 I68 M69 L70 E71 A72 L73 E74 R75 V76

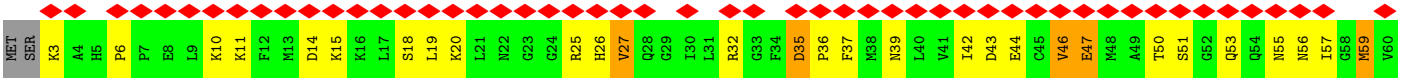
- Molecule 34: Small nuclear ribonucleoprotein G

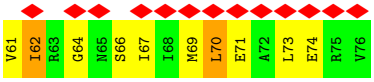
Chain 5g:

MET SER K3 A4 H5 P6 K10 K11 F12 M13 D14 K15 K16 L17 S18 L19 K20 R25 H26 V27 D35 P36 N39 I42 D43 E44 C45 V46 E47 T50 S51 G52 Q53 Q54 N55 N56 I57 G58 M59 V60 V61 I62 R63 G64 N65 S66 I67 I68 M69 L70 E71 A72 L73 E74

• Molecule 34: Small nuclear ribonucleoprotein G

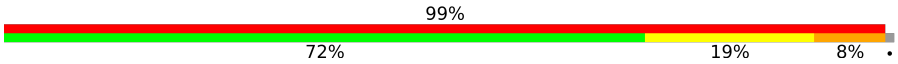
Chain 4g:

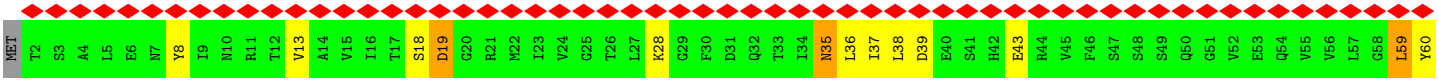


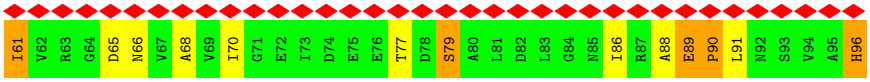


• Molecule 35: U6 snRNA-associated Sm-like protein LSm8

Chain 6S:

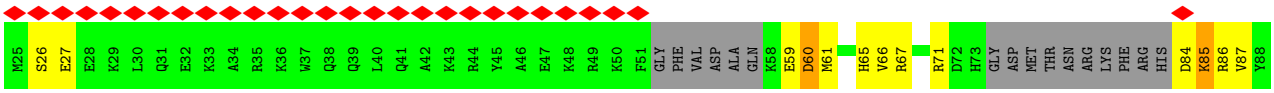


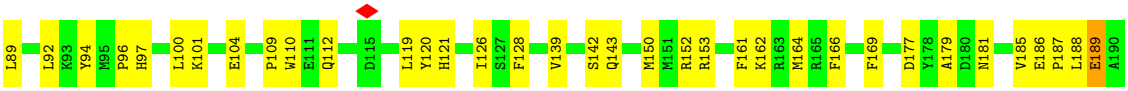


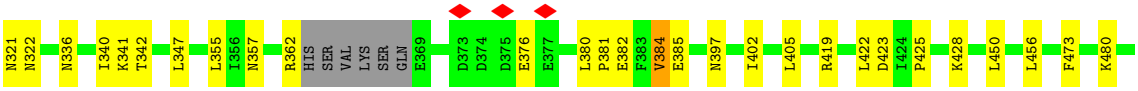


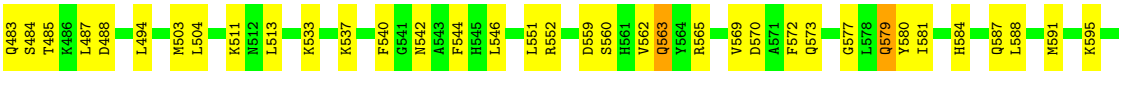
• Molecule 36: Pre-mRNA-processing-splicing factor 8

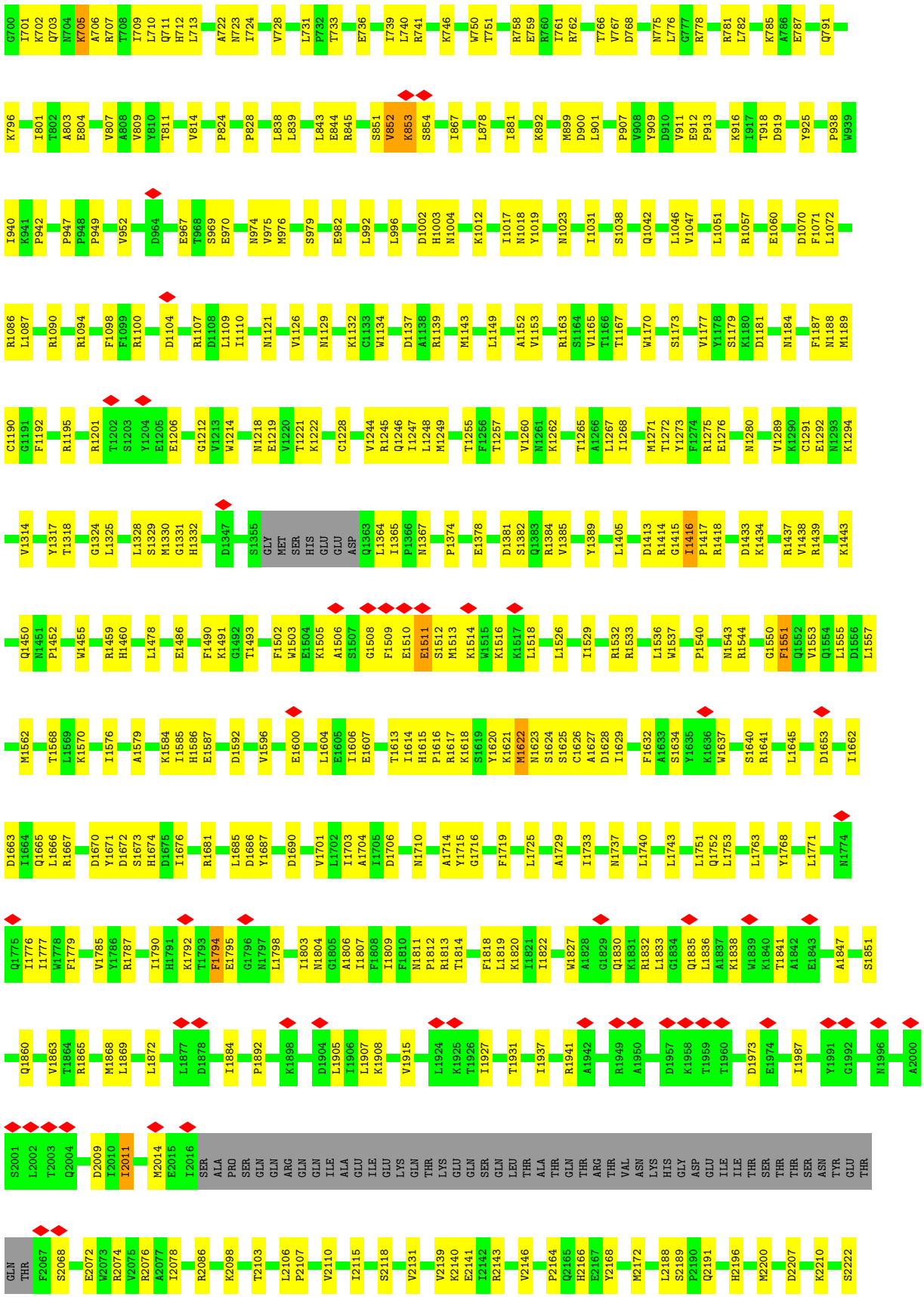
Chain 5A:










G700 I701 K702 Q703 N704 K705 A706 R707 T708 I709 L710 Q711 H712 L713 A722 N723 I724 V728 L731 P732 T733 E736 I739 L740 R741 K746 W750 T751 R758 E759 R760 I761 R762 T766 V767 D768 N775 L776 G777 R778 R781 L782 K785 A786 E787 Q791
K796 I801 T802 A803 E804 V807 A808 V809 Y810 T811 V814 P824 P828 L838 L839 L843 E844 R845 S851 V852 K853 S854 I867 L878 I881 K892 M899 D900 L901 P907 V908 Y909 D910 V911 E912 P913 K916 I917 T918 D919 Y925 P938 W939
I940 K941 P942 P947 P948 P949 V952 D964 E967 T968 S969 E970 N974 V975 M976 S979 E982 L992 L996 D1002 H1003 N1004 K1012 I1017 N1018 Y1019 N1023 I1031 S1038 Q1042 L1046 V1047 L1051 R1057 E1060 D1070 F1071 L1072
R1086 L1087 R1090 R1094 F1098 F1099 R1100 D1104 R1107 D1108 L1109 I1110 N1121 V1126 N1129 K1132 C1133 W1134 D1137 A1138 R1139 M1143 L1149 A1152 V1153 R1163 S1164 V1165 T1166 T1167 W1170 S1173 V1177 Y1178 S1179 K1180 D1181 N1184 F1187 N1188 M1189
C1190 G1191 F1192 R1195 R1201 T1202 S1203 Y1204 E1205 E1206 G1212 V1213 W1214 N1218 E1219 V1220 T1221 K1222 C1228 V1244 R1245 Q1246 I1247 L1248 M1249 T1255 F1256 T1257 V1260 N1261 K1262 T1265 A1266 L1267 I1268 M1271 T1272 Y1273 F1274 R1275 E1276 N1280 V1289 K1290 C1291 E1292 N1293 K1294
V1314 Y1317 T1318 G1324 L1325 L1328 S1329 M1330 G1331 H1332 D1347 S1355 GLY MET SER HIS GLU GLU ASP Q1363 L1364 I1365 P1366 N1367 P1374 E1378 D1381 S1382 Q1383 R1384 V1385 Y1389 L1405 D1413 R1414 G1415 I1416 P1417 R1418 D1433 K1434 R1437 V1438 R1439 K1443
Q1450 N1451 P1452 W1455 R1459 H1460 L1478 E1486 F1490 K1491 G1492 T1493 F1502 W1503 E1504 K1505 A1506 S1507 G1508 F1509 E1510 E1511 S1512 M1513 K1514 W1515 K1516 K1517 L1518 L1526 I1529 R1532 R1533 L1536 W1537 P1540 N1543 R1544 G1550 F1551 Q1552 V1553 Q1554 L1555 D1556 L1557
M1562 T1568 L1569 K1570 I1576 A1579 K1584 I1585 H1586 E1587 D1592 V1596 E1600 L1604 E1605 I1606 E1607 T1613 I1614 H1615 P1616 R1617 K1618 S1619 Y1620 K1621 M1622 N1623 S1624 S1625 C1626 A1627 D1628 I1629 F1632 A1633 S1634 Y1635 K1636 W1637 S1640 R1641 L1645 D1653 I1662
D1663 I1664 Q1665 L1666 R1667 D1670 Y1671 D1672 S1673 H1674 D1675 I1676 R1681 L1685 D1686 Y1687 D1690 V1701 L1702 I1703 A1704 I1705 D1706 N1710 A1714 Y1715 G1716 F1719 L1725 A1729 I1733 N1737 L1740 L1743 L1751 Q1752 L1753 L1763 Y1768 L1771 N1774
Q1775 I1776 I1777 W1778 F1779 V1785 Y1786 R1787 I1790 H1791 K1792 T1793 F1794 E1795 G1796 N1797 L1798 I1803 N1804 G1805 A1806 I1807 F1808 I1809 F1810 N1811 P1812 R1813 T1814 F1818 L1819 K1820 I1821 I1822 W1827 A1828 G1829 Q1830 K1831 R1832 L1833 G1834 Q1835 L1836 A1837 K1838 W1839 K1840 T1841 A1842 E1843 A1847 S1851
Q1860 V1863 T1864 R1865 M1868 L1869 L1872 L1877 D1878 I1884 P1892 K1898 D1904 L1905 I1906 L1907 K1908 V1915 L1924 K1925 T1926 I1927 T1931 I1937 R1941 A1942 R1949 A1950 D1957 K1958 T1959 T1960 D1973 E1974 I1987 Y1991 G1992 N1996 A2000
S2001 L2002 T2003 Q2004 D2009 I2010 I2011 M2014 E2015 I2016 SER ALA PRO SER GLN GLN ARG GLN GLN ILE ALA GLU ILE GLU LYS GLN THR LYS GLU GLN SER GLN LEU THR ALA THR GLN THR ARG THR VAL ASN LYS HIS GLY ASP GLU ILE ILE THR SER THR THR SER ASN TYR GLU THR
GLN THR F2067 S2068 E2072 W2073 R2074 V2075 R2076 A2077 I2078 R2086 K2098 T2103 L2106 P2107 V2110 I2115 S2118 V2131 V2139 K2140 E2141 I2142 R2143 V2146 P2164 Q2165 H2166 E2167 Y2168 M2172 L2188 S2189 P2190 Q2191 H2196 M2200 D2207 K2210 S2222

L2225 L2230 Y2251 L2252 H2255 M2261 M2272 V2273 N2280 M2284 R2287 L2296 H2306 E2307 V2308 H2309 R2310 P2311 L2315 L2319 E2324 S2327 A2328 D2329 R2330 A2335

- Molecule 37: Splicing factor 3A subunit 3

Chain A3: 76% 66% 9% 24%

M1 E2 T3 I4 L5 E6 Q7 Q8 R9 R10 Y11 H12 E13 E14 K15 E16 R17 L18 M19 D20 V21 M22 A23 K24 E25 M26 L27 T28 K29 K30 S31 T32 L33 R34 D35 Q36 I37 N38 S39 D40 H41 R42 T43 R44 A45 M46 Q47 D48 R49 Y50 M51 E52 V53 S54 G55 N56 L57 R58 D59 L60

Y61 D62 D63 K64 D65 G66 L67 R68 K69 E70 E71 L72 N73 A74 I75 S76 G77 P78 N79 E80 F81 A82 E83 F84 Y85 N86 R87 L88 K89 Q90 T91 K92 E93 F94 H95 R96 K97 H98 P99 N100 E101 I102 C103 V104 P105 M106 S107 V108 E109 F110 E111 E112 L113 L114 K115 A116 R117 E118 N119 P120

S121 E122 E123 A124 Q125 N126 L127 V128 E129 F130 T131 D132 E133 E134 G135 Y136 G137 R138 Y139 L140 D141 L142 H143 D144 C145 Y146 L147 K148 Y149 I150 N151 L152 K153 A154 S155 E156 K157 L158 D159 Y160 I161 T162 Y163 L164 S165 I166 F167 D168 Q169 L170 F171 D172 I173 P174 K175 E176 R177 K178 N179 A180

E181 Y182 K183 R184 Y185 L186 E187 M188 L189 L190 E191 Y192 L193 Q194 D195 Y196 T197 D198 R199 V200 K201 P202 L203 Q204 D205 Q206 N207 E208 L209 F210 G211 K212 I213 Q214 A215 E216 F217 E218 K219 K220 W221 E222 N223 G224 T225 F226 P227 G228 W229 PRO LYS GLU THR SER SER ALA LEU THR HIS ALA

GLY ALA HIS LEU ASP LEU SER ALA PHE SER SER TRP GLU GLU LEU ALA SER LEU GLY LEU ASP ARG LEU LYS SER ALA LEU LEU ALA LEU GLY LEU LYS CYS GLY GLY THR LEU E279 E280 R281 A282 Q283 R284 L285 F286 S287 T288 K289 G290 K291 S292 L293 E294 S295 L296 D297 T298 S299 L300

F301 A302 K303 N304 P305 K306 S307 K308 G309 T310 K311 R312 D313 T314 E315 R316 N317 K318 D319 I320 A321 F322 L323 E324 A325 Q326 I327 Y328 E329 Y330 V331 E332 I333 L334 G335 E336 Q337 R338 H339 L340 T341 H342 E343 N344 V345 Q346 R347 K348 Q349 A350 R351 T352 G353 E354 E355 R356 E357 E358 E359 E360

E361 E362 GLN ILE SER GLU SER GLU SER GLU ASP GLU GLU ASN GLU ILE ILE TYR ASN PRO LYS ASN LEU PRO LEU GLY TRP ASP GLY LYS PRO ILE P393 Y394 W395 L396 Y397 K398 L399 H400 G401 L402 N403 I404 N405 Y406 N407 C408 E409 I410 C411 G412 N413 Y414 T415 Y416 R417 G418 P419 K420

A421 F422 Q423 R424 H425 F426 A427 E428 W429 R430 H431 A432 H433 G434 M435 R436 C437 L438 G439 I440 P441 N442 T443 A444 H445 F446 A447 N448 V449 T450 Q451 I452 E453 D454 A455 V456 S457 L458 W459 A460 K461 L462 LYS LEU GLN LYS ALA SER GLU ARG TRP GLN PRO ASP THR GLU GLU GLU TYR GLU

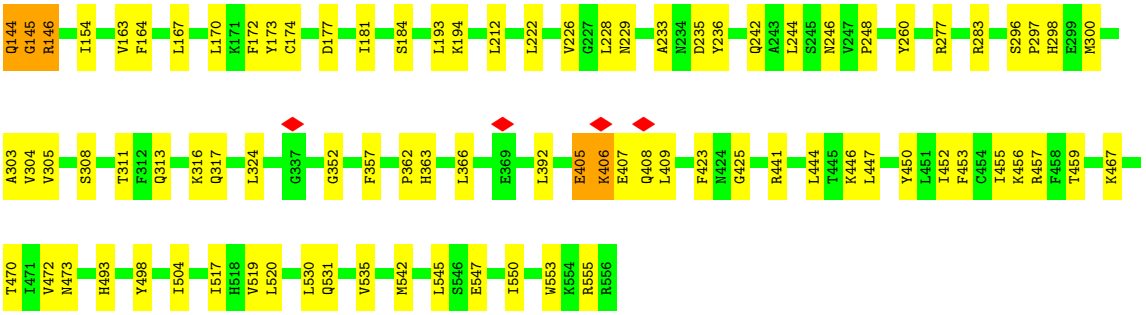
ASP SER SER GLY ASN VAL VAL ASN LYS LYS THR TYR GLU ASP LEU LYS ARG GLN GLY LEU LEU

- Molecule 38: U4/U6.U5 tri-snRNP-associated protein 2

Chain U: 65% 17% 18%

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

GLU ALA PRO ALA SER VAL VAL PRO PHE VAL ARG VAL LYS ARG GLU ARG GLU VAL ASP GLU ASP SER GLU PRO GLU ARG GLU VAL ARG ALA LYS ASN GLY ARG VAL ASP SER GLU ASP R100 C105 P106 Y107 L108 I111 K122 L128 N132 A133 Y134 A135 C136 L137 F143

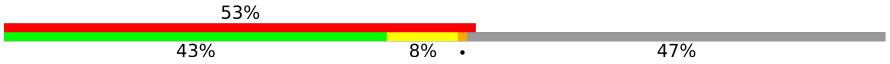


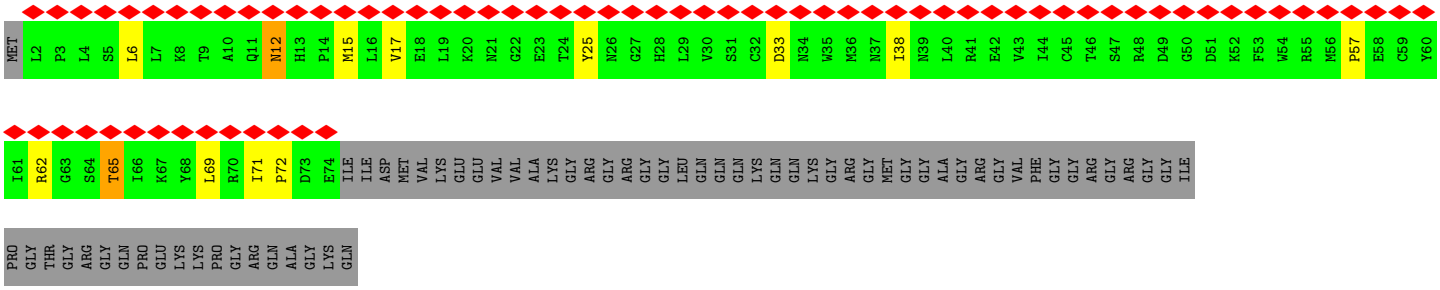
- Molecule 39: Thioredoxin-like protein 4A

Chain 5D:

- Molecule 40: U6 snRNA-associated Sm-like protein LSm4

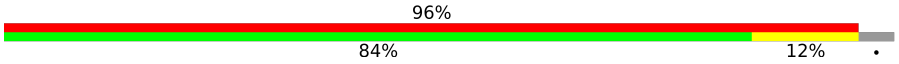
Chain 64:

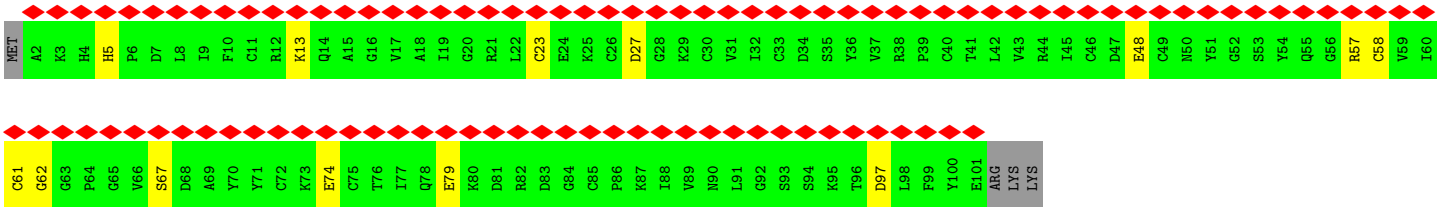




- Molecule 41: PHD finger-like domain-containing protein 5A

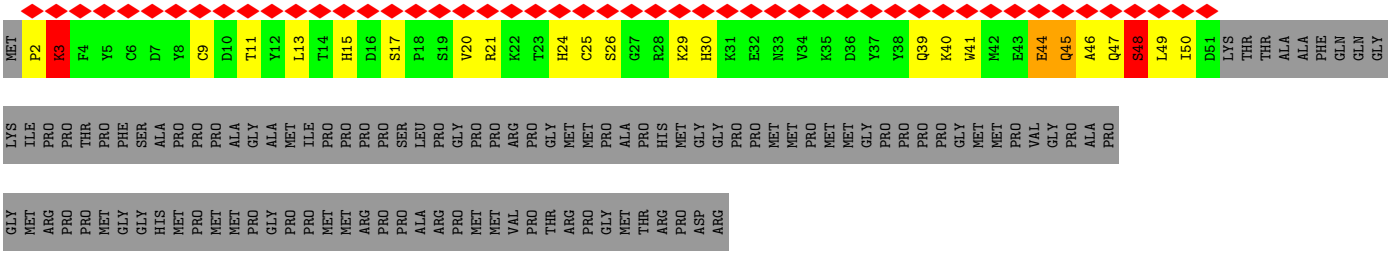
Chain BP:





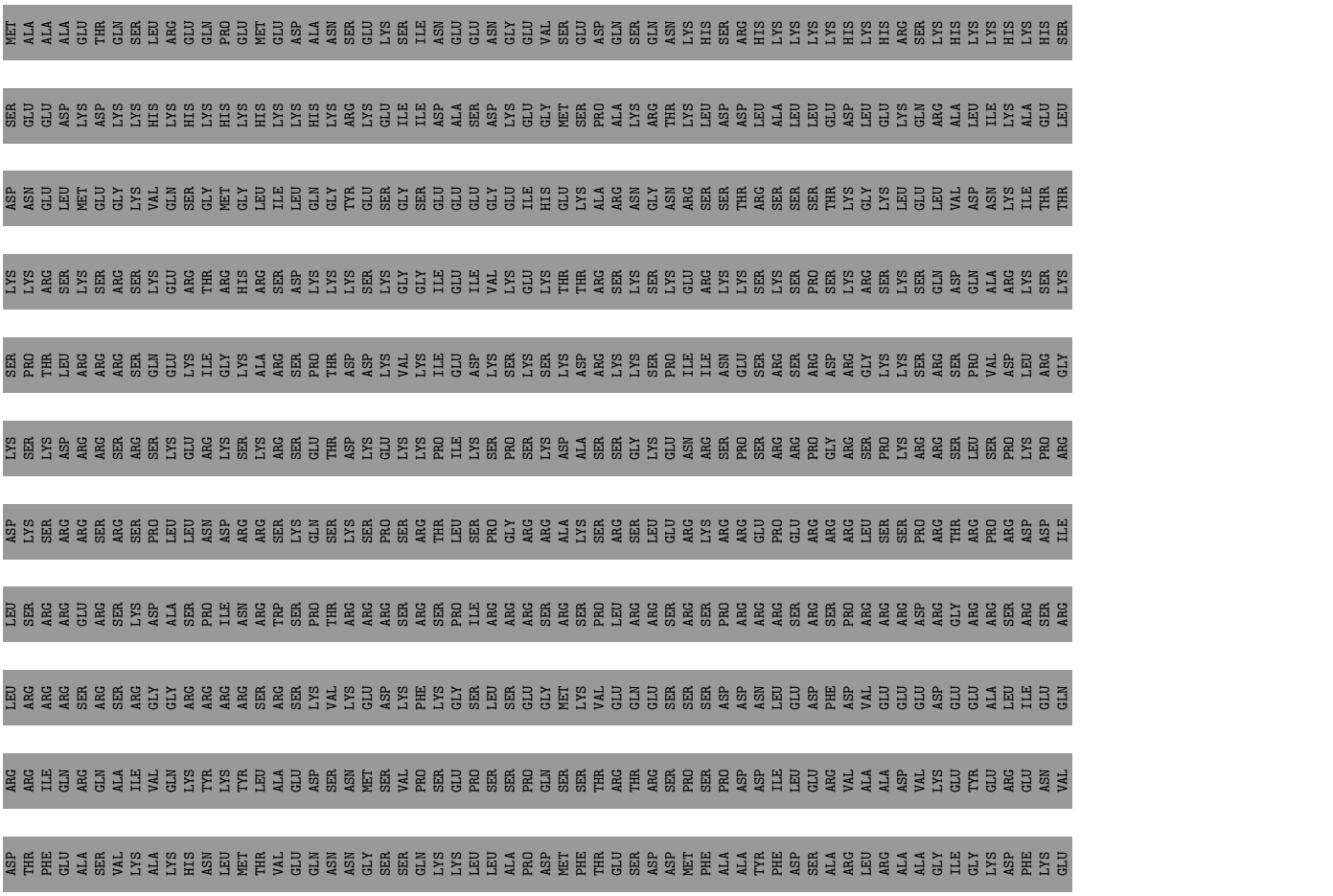
- Molecule 42: U1 small nuclear ribonucleoprotein C

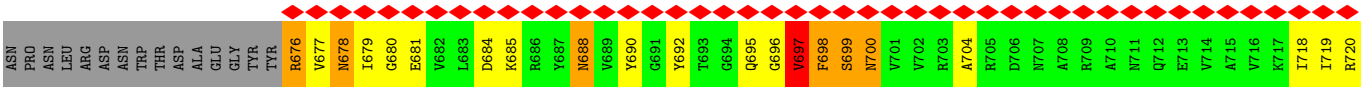
Chain 1C:

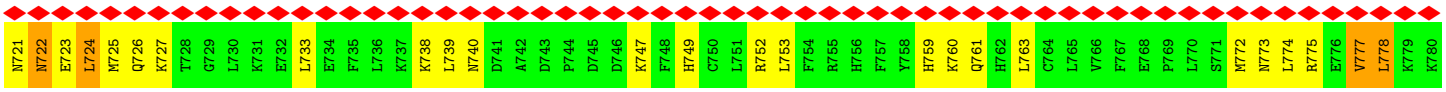


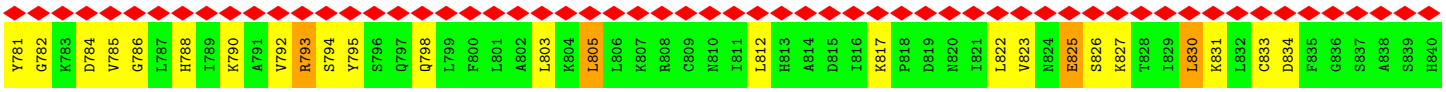
• Molecule 43: Serine/threonine-protein kinase PRP4 homolog

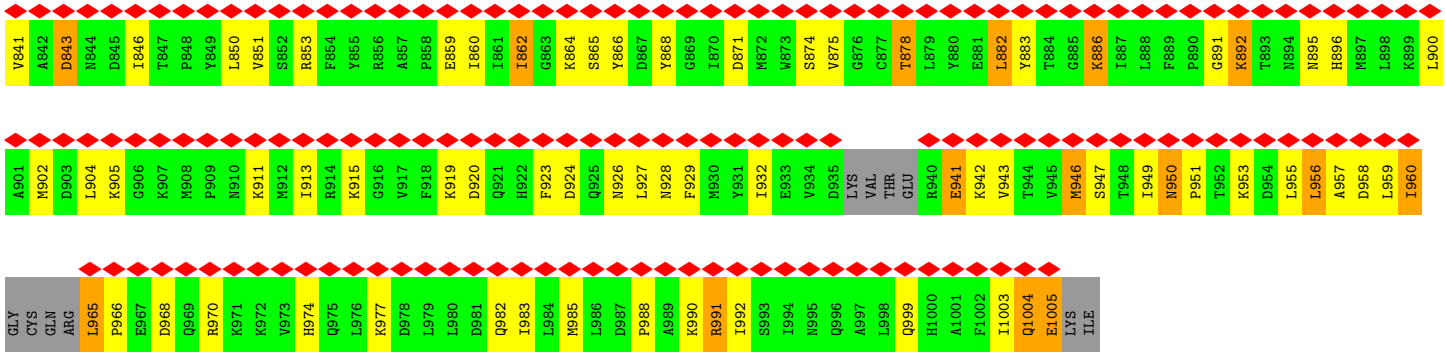
Chain K:





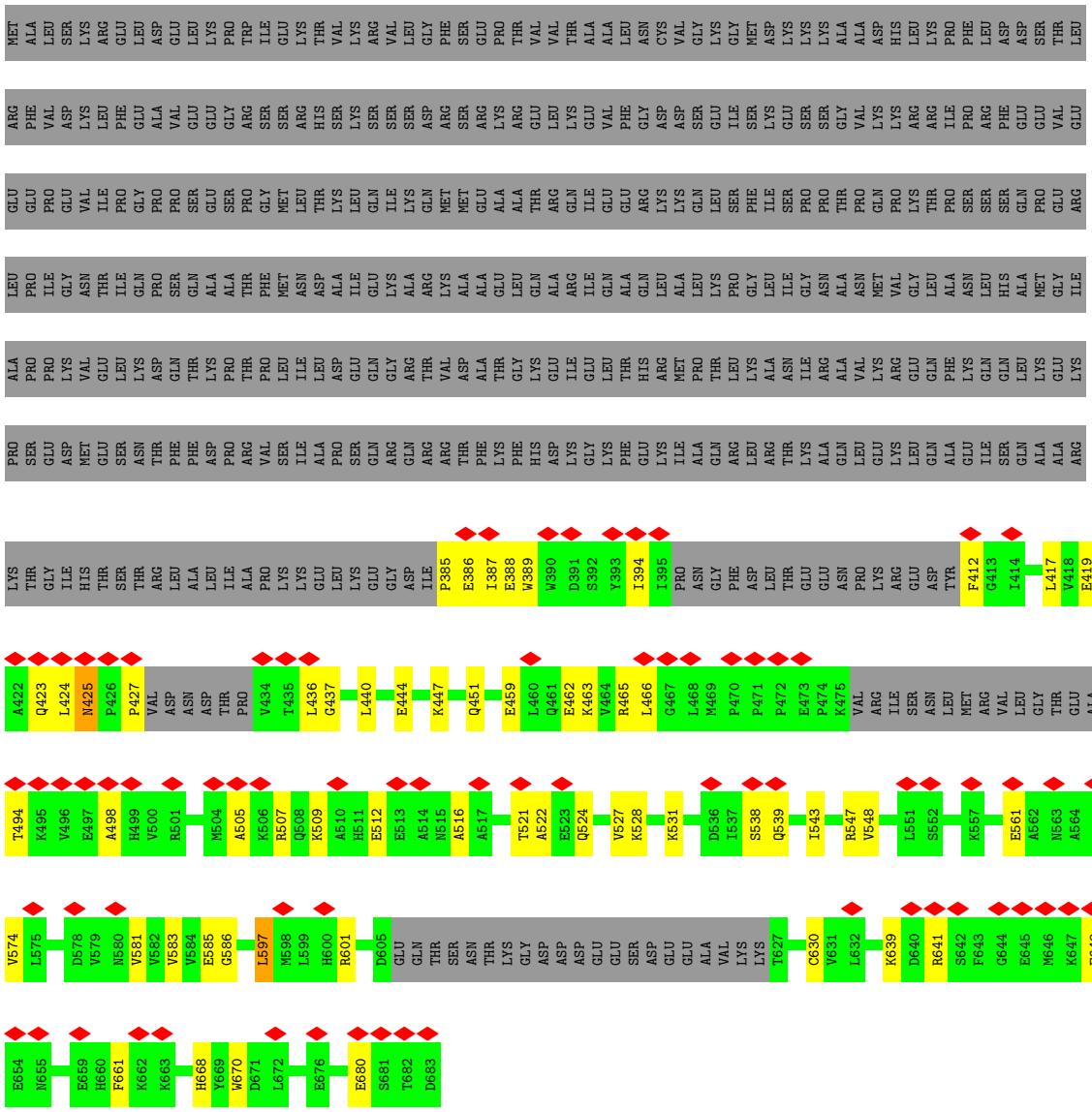






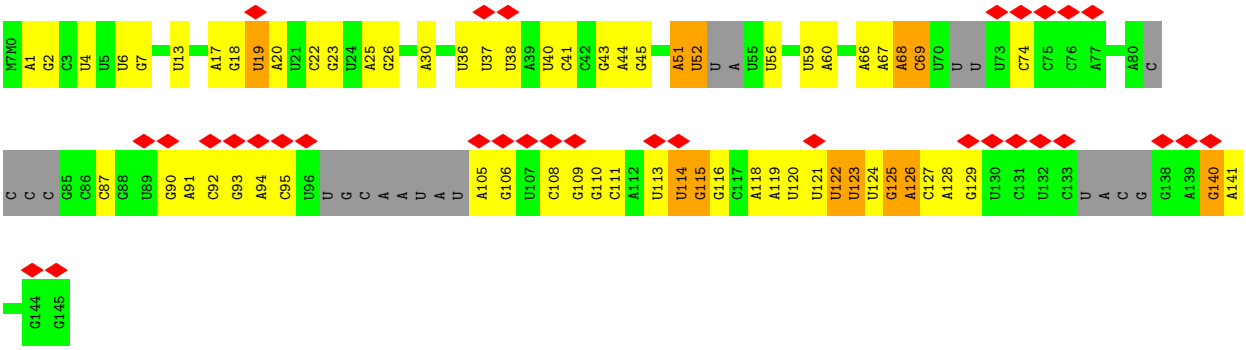
- Molecule 44: U4/U6 small nuclear ribonucleoprotein Prp3

Chain 4A:



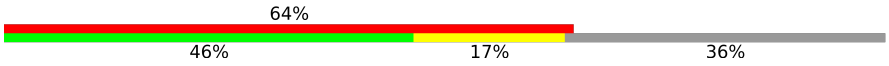
- Molecule 45: U4 snRNA

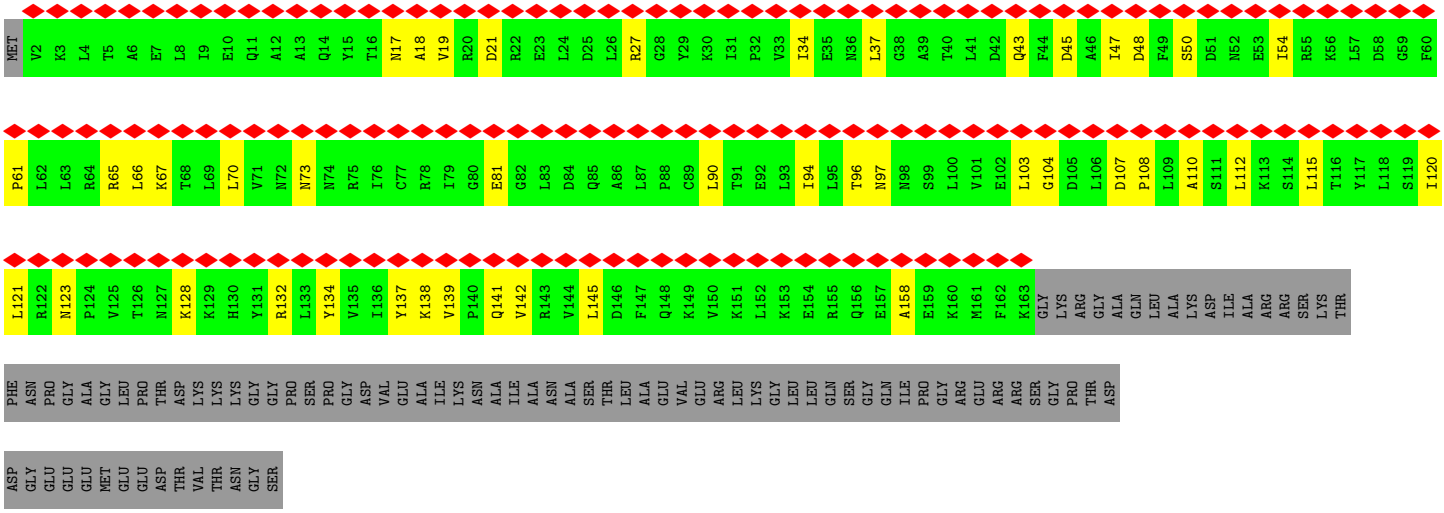
Chain 4:



• Molecule 46: U2 small nuclear ribonucleoprotein A'

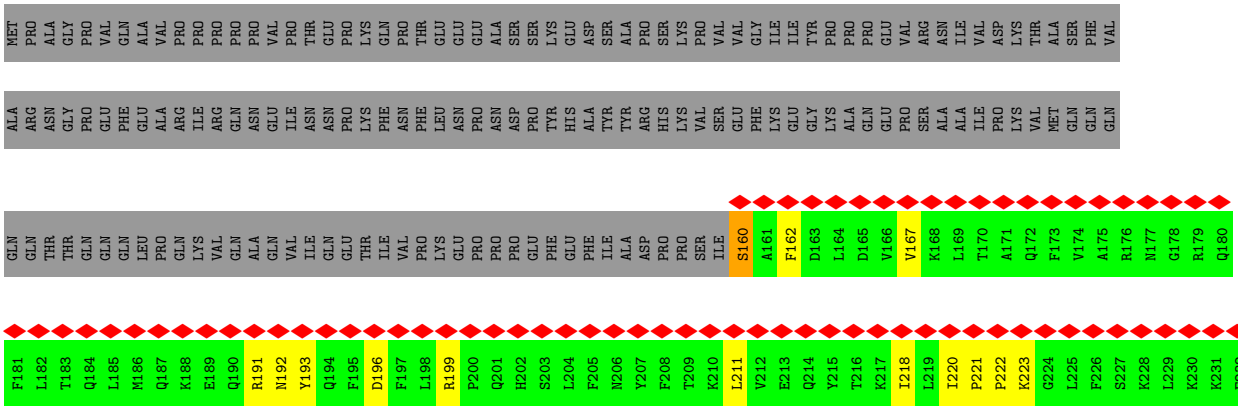
Chain 2A:

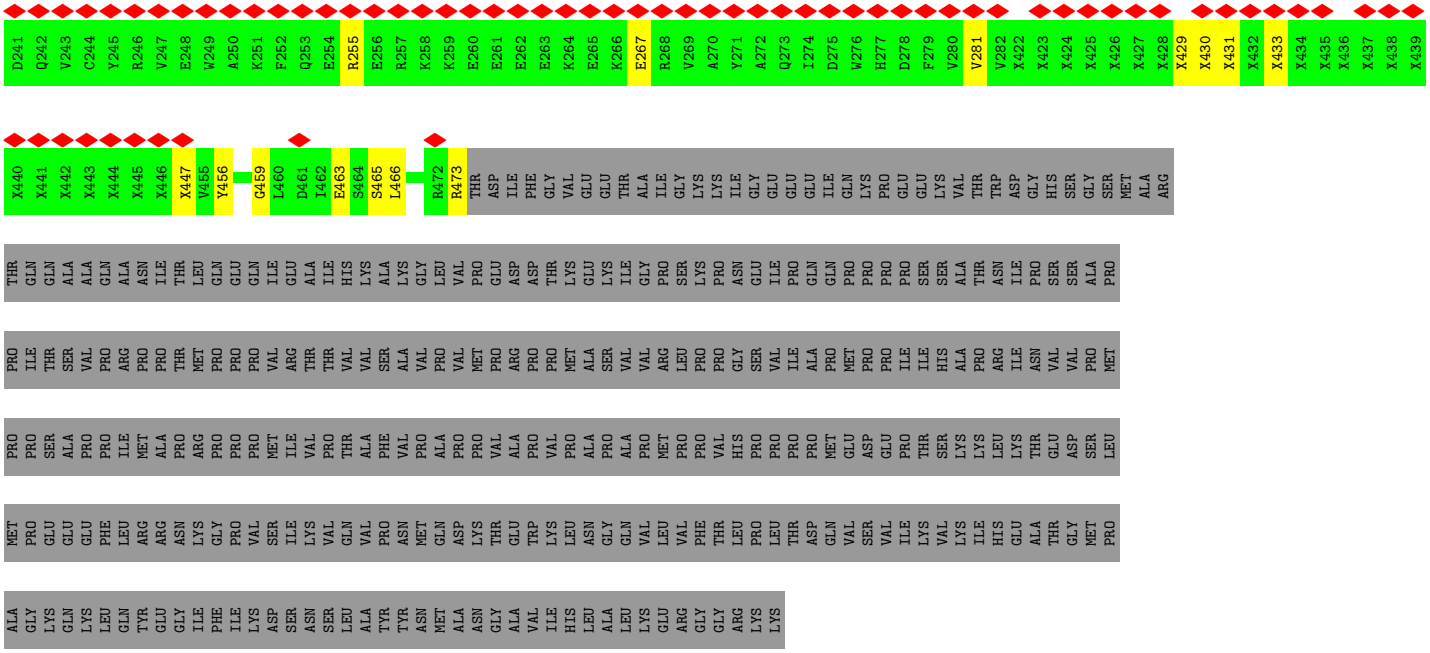




• Molecule 47: Splicing factor 3A subunit 1,Splicing factor 3A subunit 1,Splicing factor 3A subunit 1

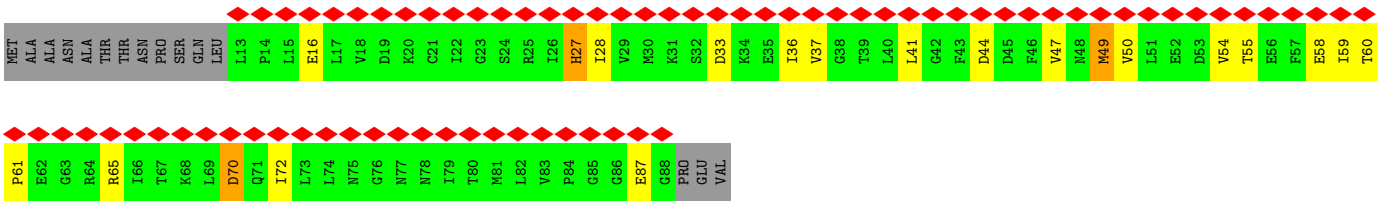
Chain A1:





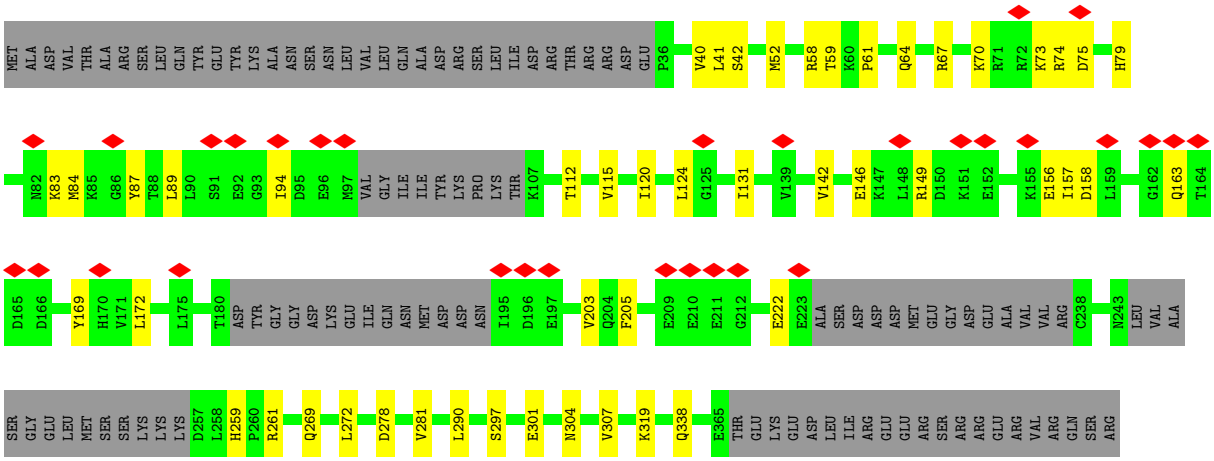
• Molecule 48: U6 snRNA-associated Sm-like protein LSm5

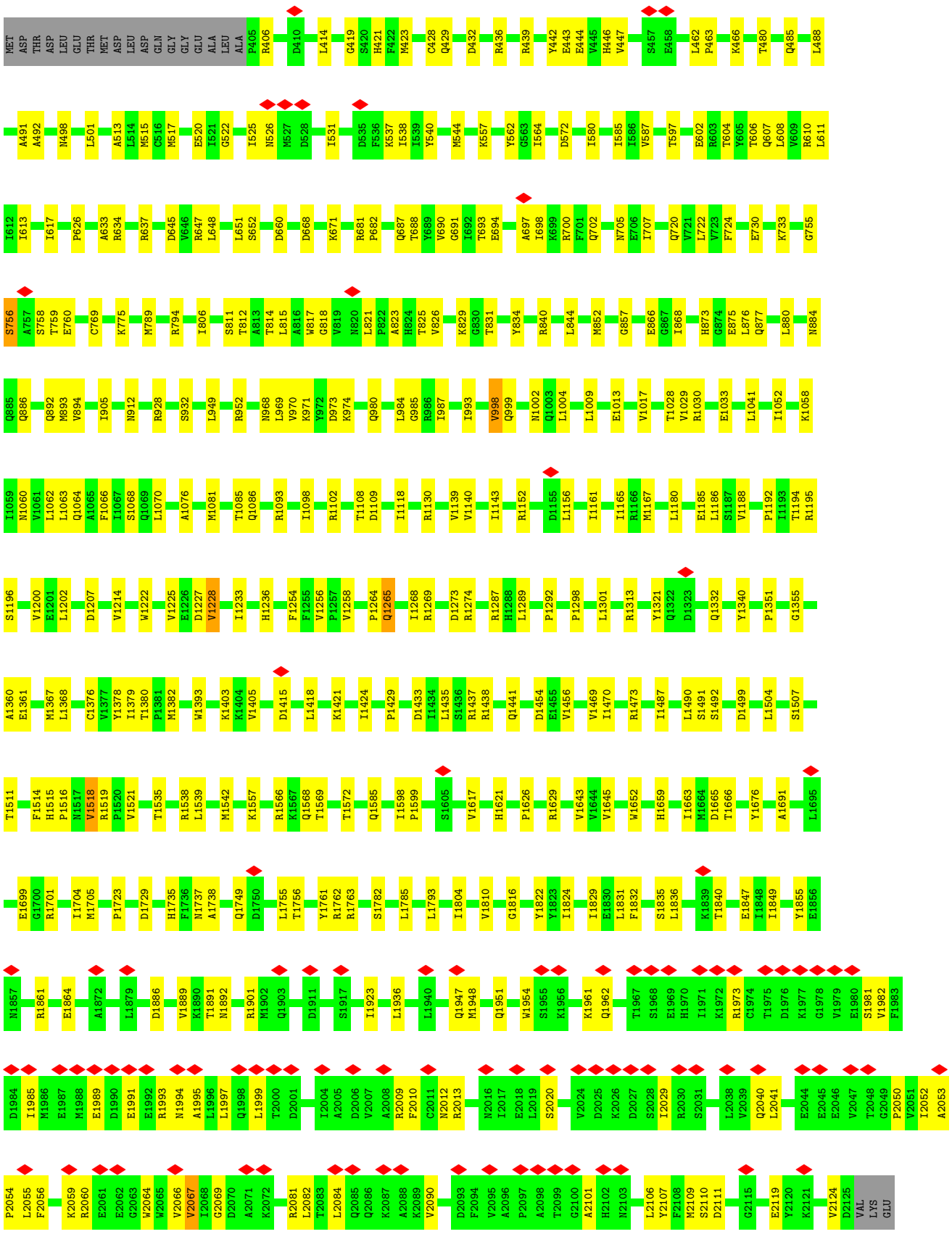
Chain 65:



• Molecule 49: U5 small nuclear ribonucleoprotein 200 kDa helicase

Chain 5B:




MET ASP THR ASP LEU GLU THR MET ASP LEU ASP GLN GLY GLY GLU ALA LEU ALA P405 R406 D410 L414 G419 S420 H421 F422 M423 C428 Q429 D432 R436 R439 Y442 E443 E444 V445 H446 V447 S457 E458 L462 P463 K466 T480 Q485 L488
A491 A492 N498 L501 A513 L514 M515 C516 M517 E520 I521 G522 I525 N526 M527 D528 I531 D535 F536 K537 I538 I539 Y540 M544 K557 V562 G563 I564 D572 I580 I585 I586 V587 T597 E602 R603 T604 Y605 T606 Q607 L608 V609 R610 L611
I612 I613 I617 P626 A633 R634 R637 D645 V646 R647 L648 L651 S652 D660 D668 K671 R681 P682 Q687 T688 Y689 V690 G691 I692 T693 E694 A697 I698 K699 R700 F701 Q702 N705 E706 I707 Q720 V721 L722 V723 F724 E730 K733 G755
S756 A757 S758 T759 E760 C769 K775 M789 R794 I806 S811 T812 A813 T814 L815 A816 W817 G818 V819 N820 L821 P822 A823 H824 T825 V826 K829 G830 T831 Y834 R840 L844 M852 G857 E866 G867 I868 H873 G874 E875 L876 Q877 L880 N884
Q885 Q886 Q892 M893 V894 I905 N912 R928 S932 L949 R952 N968 L969 V970 K971 Y972 D973 K974 Q980 L984 G985 R986 I987 I993 V998 Q999 N1002 Q1003 L1004 L1009 E1013 V1017 T1028 V1029 R1030 E1033 L1041 I1052 K1058
I1059 N1060 V1061 L1062 L1063 Q1064 A1065 F1066 I1067 S1068 Q1069 L1070 A1076 M1081 T1085 Q1086 R1093 I1098 R1102 T1108 D1109 I1118 R1130 V1139 V1140 I1143 R1152 D1155 L1156 I1161 I1165 R1166 M1167 L1180 E1185 L1186 S1187 V1188 P1192 I1193 T1194 R1195
S1196 V1200 E1201 L1202 D1207 V1214 W1222 V1225 E1226 D1227 V1228 I1233 H1236 F1254 F1255 V1256 P1257 V1258 P1264 Q1265 I1268 R1269 D1273 R1274 R1287 H1288 L1289 P1292 P1298 L1301 R1313 Y1321 Q1322 D1323 Q1332 Y1340 P1351 G1355
A1360 E1361 M1367 L1368 C1376 V1377 Y1378 I1379 T1380 P1381 M1382 W1393 K1403 K1404 V1405 D1415 L1418 K1421 I1424 P1429 D1433 I1434 L1435 S1436 R1437 R1438 Q1441 D1454 E1455 V1456 V1469 I1470 R1473 I1487 L1490 S1491 S1492 D1499 L1504 S1507
T1511 F1514 H1515 P1516 N1517 V1518 R1519 P1520 V1521 T1535 R1538 L1539 M1542 K1557 R1566 K1567 Q1568 T1569 T1572 Q1585 I1598 P1599 S1605 V1617 H1621 P1626 R1629 V1643 V1644 V1645 W1652 H1659 I1663 M1664 D1665 T1666 Y1676 A1691 L1695
E1699 G1700 R1701 I1704 M1705 P1723 D1729 H1735 F1736 N1737 A1738 Q1749 D1750 L1755 T1756 Y1761 R1762 R1763 S1782 L1785 L1793 I1804 V1810 G1816 Y1822 Y1823 I1824 I1829 E1830 L1831 F1832 S1835 L1836 K1839 T1840 E1847 I1848 I1849 Y1855 E1856
N1857 R1861 E1864 A1872 L1879 D1886 V1889 K1890 T1891 N1892 R1901 M1902 Q1903 D1911 S1917 I1923 L1936 L1940 Q1947 M1948 Q1951 W1954 S1955 K1956 K1961 Q1962 T1967 S1968 E1969 H1970 I1971 K1972 R1973 C1974 T1975 D1976 K1977 G1978 V1979 E1980 S1981 V1982 F1983
D1984 I1985 M1986 E1987 M1988 E1989 D1990 E1991 E1992 R1993 N1994 A1995 L1996 L1997 Q1998 L1999 T2000 D2001 I2004 A2005 D2006 V2007 A2008 R2009 F2010 C2011 N2012 R2013 N2016 I2017 E2018 L2019 S2020 V2024 D2025 K2026 D2027 S2028 I2029 R2030 S2031 L2038 V2039 Q2040 L2041 E2044 E2045 E2046 V2047 T2048 G2049 P2050 V2051 I2052 A2053
P2054 L2055 F2056 K2059 R2060 E2061 E2062 G2063 W2064 W2065 V2066 V2067 I2068 G2069 D2070 A2071 K2072 R2081 L2082 T2083 L2084 Q2085 Q2086 K2087 A2088 K2089 V2090 D2093 F2094 V2095 A2096 P2097 A2098 T2099 G2100 A2101 H2102 N2103 L2106 Y2107 F2108 M2109 S2110 D2111 G2115 E2119 Y2120 K2121 V2124 D2125 VAL LYS GLU

ALA GLU THR ASP SER ASP SER ASP

- Molecule 50: Splicing factor 3B subunit 1

Chain B1: 65% 54% 11% 35%

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

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

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

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

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

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

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

TYR VAL PRO ILE ARG THR PRO ALA ARG LYS LEU THR ALA THR PRO THR PRO LEU GLY GLY MET THR GLY PHE HIS MET GLN THR GLU ASP ARG THR MET LYS SER VAL N457 D458 Q459 P460 S461 G462 N463 L464 P465 F466 L467 K468 P469 D470 D471 I472 Q473 Y474 F475 D476 K477 L478 L479 V480

D481 V482 D483 E484 S485 T486 L487 S488 P489 E490 E491 Q492 K493 E494 R495 K496 I497 M498 K499 L500 L501 L502 K503 I504 K505 N506 G507 T508 P509 P510 M511 R512 K513 A514 A515 L516 R517 Q518 I519 T520 D521 K522 A523 R524 E525 F526 G527 A528 G529 P530 L531 F532 N533 Q534 I535 L536 P537 L538 L539 M540

S541 P542 T543 L544 E545 D546 Q547 E548 R549 H550 L551 L552 V553 K554 V555 I556 D557 R558 I559 L560 Y561 K562 L563 D564 D565 L566 V567 R568 P569 Y570 V571 H572 K573 I574 L575 V576 V577 I578 E579 P580 L581 L582 I583 D584 E585 D586 Y587 Y588 A589 R590 V591 E592 G593 R594 E595 I596 I597 S598 N599 L600

A601 K602 A603 A604 G605 L606 A607 T608 M609 I610 S611 T612 M613 R614 P615 D616 I617 D618 N619 M620 D621 E622 Y623 V624 R625 N626 T627 T628 A629 R630 A631 F632 A633 V634 V635 A636 S637 A638 L639 G640 I641 P642 S643 L644 L645 P646 F647 L648 K649 A650 V651 C652 K653 S654 K655 K656 S657 W658 Q659 A660

R661 H662 T663 G664 I665 K666 I667 V668 Q669 Q670 I671 A672 I673 L674 M675 G676 G677 A678 I679 L680 P681 H682 L683 R684 S685 L686 V687 E688 I689 I690 E691 H692 G693 L694 V695 D696 E697 Q698 Q699 K700 V701 R702 T703 I704 S705 A706 L707 A708 I709 A710 A711 L712 A713 E714 A715 A716 T717 P718 Y719 G720

I721 E722 S723 F724 D725 S726 V727 L728 K729 P730 L731 W732 K733 G734 I735 R736 Q737 H738 R739 G740 K741 G742 L743 A744 A745 F746 L747 K748 A749 I750 G751 Y752 L753 I754 P755 L756 M757 D758 A759 E760 Y761 A762 N763 Y764 Y765 T766 R767 E768 V769 M770 L771 I772 L773 I774 R775 E776 F777 Q778 S779 P780

D781 E782 E783 M784 K785 K786 I787 V788 L789 K790 V791 V792 K793 Q794 C795 C796 G797 T798 D799 G800 V801 E802 A803 N804 Y805 I806 K807 T808 E809 I810 L811 P812 P813 F814 F815 K816 H817 F818 W819 Q820 H821 R822 M823 A824 L825 D826 R827 R828 N829 Y830 R831 Q832 L833 V834 D835 T836 T837 V838 E839 L840

A841 N842 K843 V844 G845 A846 A847 E848 I849 I850 S851 R852 I853 V854 D855 D856 L857 K858 D859 E860 A861 E862 Q863 Y864 R865 K866 M867 V868 M869 E870 T871 I872 E873 K874 I875 M876 G877 N878 L879 G880 A881 A882 D883 I884 D885 H886 K887 L888 E889 E890 Q891 L892 I893 D894 G895 I896 L897 Y898 A899 F900

Q901 E902 Q903 T904 T905 E906 D907 S908 V909 M910 L911 N912 G913 F914 G915 T916 V917 V918 N919 A920 L921 G922 K923 R924 V925 K926 P927 Y928 L929 P930 Q931 I932 C933 G934 T935 V936 L937 W938 R939 L940 N941 N942 K943 S944 A945 K946 V947 R948 Q949 Q950 A951 A952 D953 L954 I955 S956 R957 T958 A959 V960

V961 M962 K963 T964 C965 Q966 E967 E968 K969 L970 M971 G972 H973 L974 G975 V976 V977 L978 Y979 E980 Y981 L982 G983 E984 E985 Y986 P987 E988 V989 L990 G991 S992 I993 L994 G995 A996 L997 K998 A999 I1000 V1001 N1002 V1003 I1004 G1005 M1006 H1007 K1008 M1009 T1010 P1011 P1012 I1013 K1014 D1015 L1016 L1017 P1018 R1019 L1020

T1021 P1022 I1023 L1024 K1025 N1026 R1027 H1028 E1029 K1030 V1031 Q1032 E1033 N1034 C1035 I1036 D1037 L1038 V1039 G1040 R1041 I1042 A1043 D1044 R1045 G1046 A1047 E1048 Y1049 V1050 S1051 A1052 R1053 E1054 W1055 M1056 R1057 I1058 C1059 F1060 E1061 L1062 L1063 E1064 L1065 L1066 K1067 A1068 H1069 K1070 K1071 A1072 I1073 R1074 R1075 A1076 T1077 V1078 N1079 T1080

F1081 G1082 Y1083 I1084 A1085 K1086 A1087 I1088 G1089 P1090 H1091 D1092 V1093 L1094 A1095 T1096 L1097 L1098 N1099 N1100 L1101 K1102 V1103 Q1104 E1105 R1106 Q1107 N1108 R1109 V1110 C1111 T1112 T1113 V1114 A1115 I1116 A1117 I1118 V1119 A1120 E1121 T1122 C1123 S1124 P1125 F1126 T1127 V1128 L1129 P1130 A1131 L1132 M1133 N1134 E1135 Y1136 R1137 V1138 P1139 E1140

L1141 N1142 V1143 Q1144 N1145 G1146 V1147 L1148 K1149 S1150 L1151 S1152 F1153 L1154 F1155 E1156 Y1157 I1158 G1159 E1160 M1161 G1162 K1163 D1164 Y1165 I1166 Y1167 A1168 V1169 T1170 P1171 L1172 L1173 E1174 D1175 A1176 L1177 M1178 D1179 R1180 D1181 L1182 V1183 H1184 R1185 Q1186 T1187 A1188 S1189 A1190 V1191 V1192 Q1193 H1194 M1195 S1196 L1197 G1198 V1199 Y1200

G1201 F1202 G1203 C1204 E1205 D1206 S1207 L1208 N1209 H1210 L1211 L1212 N1213 Y1214 V1215 W1216 P1217 N1218 V1219 F1220 E1221 T1222 S1223 P1224 H1225 V1226 I1227 Q1228 A1229 V1230 M1231 G1232 A1233 L1234 E1235 G1236 L1237 R1238 V1239 A1240 I1241 G1242 P1243 C1244 R1245 M1246 L1247 Q1248 Y1249 C1250 L1251 Q1252 G1253 L1254 F1255 H1256 P1257 A1258 R1259 K1260

V1261 R1262 D1263 V1264 Y1265 W1266 K1267 I1268 Y1269 N1270 S1271 I1272 Y1273 I1274 G1275 S1276 Q1277 D1278 A1279 L1280 I1281 A1282 H1283 Y1284 P1285 R1286 I1287 Y1288 N1289 D1290 D1291 K1292 N1293 T1294 Y1295 I1296 R1297 Y1298 E1299 L1300 D1301 Y1302 I1303 L1304

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	86146	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 ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.097	Depositor
Minimum map value	-0.048	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.01	Depositor
Map size (Å)	429.24, 429.24, 429.24	wwPDB
Map dimensions	420, 420, 420	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.022, 1.022, 1.022	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: IHP, GTP, MG, M7M, ZN

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	1	0.63	3/3891 (0.1%)	0.78	9/6061 (0.1%)
2	6	0.51	5/1264 (0.4%)	0.45	0/1961
3	5O	0.26	0/2448	0.60	0/3316
4	B4	0.91	0/632	1.39	4/855 (0.5%)
5	13	0.72	0/645	1.22	4/870 (0.5%)
5	23	0.47	0/660	0.71	0/889
5	43	0.32	0/660	0.71	0/889
5	53	0.36	0/665	0.55	0/896
6	4B	0.45	0/2921	0.74	1/3966 (0.0%)
7	1e	0.82	0/646	1.42	6/867 (0.7%)
7	2e	0.47	0/677	0.69	0/908
7	4e	0.34	0/639	0.87	2/857 (0.2%)
7	5e	0.34	0/646	0.75	0/867
8	I	0.54	0/590	0.84	2/916 (0.2%)
9	1K	1.47	13/1695 (0.8%)	1.33	11/2288 (0.5%)
10	4C	0.31	0/2406	0.65	2/3232 (0.1%)
11	11	0.93	0/649	1.35	6/878 (0.7%)
11	21	0.41	0/642	0.60	0/867
11	41	0.36	0/649	0.78	0/878
11	51	0.36	0/649	0.78	1/878 (0.1%)
12	R	0.34	0/891	1.01	2/1188 (0.2%)
13	1f	0.82	1/588 (0.2%)	1.34	5/795 (0.6%)
13	2f	0.43	0/574	0.65	0/775
13	4f	0.39	0/574	0.74	0/775
13	5f	0.37	0/579	0.82	0/783
14	66	1.12	1/575 (0.2%)	1.55	5/776 (0.6%)
15	X	0.38	0/398	0.84	2/524 (0.4%)
16	12	0.95	0/786	1.33	0/1055
16	22	0.40	0/784	0.65	0/1053
16	42	0.38	0/747	0.71	0/1000
16	52	0.33	0/805	0.79	0/1081
17	5	0.32	0/2444	0.56	1/3798 (0.0%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
18	67	1.11	1/611 (0.2%)	1.53	6/824 (0.7%)
19	62	1.00	0/773	1.59	7/1043 (0.7%)
20	2B	0.34	0/759	0.50	0/1016
21	A2	0.72	0/1254	1.18	4/1682 (0.2%)
22	B2	0.82	5/1747 (0.3%)	1.22	24/2356 (1.0%)
23	5C	0.48	0/6879	0.63	6/9344 (0.1%)
24	5X	0.60	1/4859 (0.0%)	0.85	5/6522 (0.1%)
25	1b	0.88	1/702 (0.1%)	1.32	5/936 (0.5%)
25	2b	0.44	0/674	0.59	0/899
25	4b	0.30	0/679	0.66	0/905
25	5b	0.33	0/602	0.55	0/801
26	B5	0.53	0/584	0.62	0/789
27	1A	1.32	1/801 (0.1%)	1.22	2/1074 (0.2%)
28	S	0.40	1/925 (0.1%)	0.79	5/1229 (0.4%)
29	5J	0.30	0/6430	0.69	13/8681 (0.1%)
30	4D	0.43	0/967	0.58	0/1305
31	63	1.05	0/709	1.41	7/959 (0.7%)
32	2	0.44	5/2209 (0.2%)	0.48	1/3429 (0.0%)
33	B3	0.42	0/9485	0.64	4/12870 (0.0%)
34	1g	0.81	0/575	1.40	6/768 (0.8%)
34	2g	0.47	0/575	0.67	0/768
34	4g	0.38	0/584	0.78	0/779
34	5g	0.38	0/584	0.78	0/779
35	68	1.09	0/728	1.67	12/987 (1.2%)
36	5A	0.43	1/18874 (0.0%)	0.65	16/25606 (0.1%)
37	A3	0.83	0/3294	1.47	18/4423 (0.4%)
38	U	0.49	0/3846	0.67	4/5208 (0.1%)
39	5D	0.31	0/1198	0.58	2/1620 (0.1%)
40	64	1.05	0/609	1.53	3/824 (0.4%)
41	BP	0.50	0/779	0.61	0/1047
42	1C	0.70	0/437	1.33	2/587 (0.3%)
43	K	0.43	0/2673	0.82	2/3593 (0.1%)
44	4A	0.39	1/1983 (0.1%)	0.69	2/2657 (0.1%)
45	4	0.38	1/2967 (0.0%)	0.50	3/4610 (0.1%)
46	2A	0.28	0/1299	0.64	0/1761
47	A1	0.83	0/1234	1.45	4/1657 (0.2%)
48	65	1.12	0/593	1.66	7/800 (0.9%)
49	5B	0.39	0/16393	0.60	2/22174 (0.0%)
50	B1	0.48	0/6878	0.68	0/9315
All	All	0.54	41/141171 (0.0%)	0.81	235/193369 (0.1%)

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	4B	0	3
7	4e	0	1
9	1K	0	1
10	4C	0	2
13	4f	0	1
13	5f	0	1
14	66	0	1
16	12	0	1
16	52	0	3
18	67	0	1
19	62	0	1
21	A2	0	4
23	5C	0	2
25	4b	0	1
29	5J	0	2
31	63	0	2
33	B3	0	1
35	68	0	3
36	5A	0	2
37	A3	0	6
38	U	0	3
40	64	0	1
44	4A	0	2
45	4	0	1
47	A1	0	4
49	5B	0	2
50	B1	0	4
All	All	0	56

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	1	78	U	O3'-P	-15.64	1.37	1.61
1	1	2	U	O3'-P	-14.60	1.39	1.61
22	B2	629	PRO	CA-C	-13.89	1.39	1.52
14	66	70	VAL	C-O	-10.01	1.13	1.24
18	67	61	GLN	C-O	-9.74	1.11	1.23

The worst 5 of 235 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
36	5A	853	LYS	N-CA-C	-16.73	92.94	113.18
29	5J	674	THR	N-CA-C	-15.87	83.54	109.40
22	B2	629	PRO	CB-CA-C	-12.37	95.83	110.92
48	65	70	ASP	CA-CB-CG	12.35	124.95	112.60
36	5A	1550	GLY	N-CA-C	11.62	127.29	111.03

There are no chirality outliers.

5 of 56 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
9	1K	56	PRO	Peptide
6	4B	420	TYR	Peptide
6	4B	459	PRO	Peptide
6	4B	470	TYR	Peptide
10	4C	350	GLN	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	1	3485	0	1756	389	0
2	6	1133	0	573	46	0
3	5O	2394	0	2326	56	0
4	B4	618	0	613	15	0
5	13	637	0	652	32	0
5	23	652	0	670	22	0
5	43	652	0	670	27	0
5	53	657	0	675	16	0
6	4B	2842	0	2746	88	0
7	1e	638	0	657	43	0
7	2e	669	0	692	20	0
7	4e	631	0	648	29	0
7	5e	638	0	657	23	0
8	I	530	0	271	11	0
9	1K	1649	0	1607	166	0
10	4C	2375	0	2375	39	0
11	11	641	0	680	18	0
11	21	634	0	679	59	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	41	641	0	681	18	0
11	51	641	0	681	16	0
12	R	874	0	883	44	0
13	1f	576	0	589	19	0
13	2f	562	0	574	11	0
13	4f	562	0	574	36	0
13	5f	567	0	575	19	0
14	66	567	0	571	15	0
15	X	394	0	406	79	0
16	12	777	0	798	31	0
16	22	774	0	802	42	0
16	42	737	0	780	44	0
16	52	796	0	821	27	0
17	5	2192	0	1110	36	0
18	67	604	0	623	8	0
19	62	761	0	777	4	0
20	2B	745	0	767	11	0
21	A2	1221	0	1197	69	0
22	B2	1699	0	1736	60	0
23	5C	6727	0	6735	138	0
24	5X	4780	0	4845	298	0
25	1b	692	0	717	25	0
25	2b	664	0	690	19	0
25	4b	669	0	697	19	0
25	5b	594	0	615	6	0
26	B5	567	0	533	14	0
27	1A	787	0	797	10	0
28	S	917	0	921	67	0
29	5J	6316	0	6227	163	0
30	4D	955	0	1008	30	0
31	63	699	0	702	13	0
32	2	1984	0	1005	23	0
33	B3	9296	0	9216	172	0
34	1g	568	0	590	40	0
34	2g	568	0	590	12	0
34	4g	577	0	603	20	0
34	5g	577	0	603	20	0
35	68	722	0	712	19	0
36	5A	18366	0	18268	574	0
37	A3	3227	0	3117	89	0
38	U	3750	0	3766	81	0
39	5D	1169	0	1141	66	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
40	64	596	0	591	9	0
41	BP	766	0	737	9	0
42	1C	425	0	392	20	0
43	K	2626	0	2698	192	0
44	4A	1946	0	2011	105	0
45	4	2690	0	1368	63	0
46	2A	1282	0	1305	27	0
47	A1	1339	0	1234	44	0
48	65	587	0	605	12	0
49	5B	16077	0	16192	256	0
50	B1	6749	0	6948	89	0
51	1C	1	0	0	0	0
51	A2	1	0	0	0	0
51	BP	3	0	0	0	0
51	U	1	0	0	0	0
52	5C	1	0	0	0	0
53	5C	32	0	12	0	0
54	5A	36	0	6	1	0
All	All	137494	0	132089	3475	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 13.

The worst 5 of 3475 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)
1:1:7:A:C5'	24:5X:478:ARG:H	1.04	1.64
38:U:134:TYR:CD1	38:U:145:GLY:HA3	1.13	1.64
24:5X:362:LEU:HD23	24:5X:391:ARG:CD	1.21	1.62
21:A2:115:PRO:CG	21:A2:176:TYR:HD1	1.00	1.59
29:5J:839:HIS:HD2	43:K:942:LYS:CE	0.95	1.59

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	
3	5O	304/357 (85%)	283 (93%)	19 (6%)	2 (1%)	18	48
4	B4	76/424 (18%)	76 (100%)	0	0	100	100
5	13	79/126 (63%)	75 (95%)	4 (5%)	0	100	100
5	23	81/126 (64%)	76 (94%)	5 (6%)	0	100	100
5	43	81/126 (64%)	76 (94%)	5 (6%)	0	100	100
5	53	82/126 (65%)	77 (94%)	5 (6%)	0	100	100
6	4B	357/522 (68%)	330 (92%)	25 (7%)	2 (1%)	21	51
7	1e	75/92 (82%)	70 (93%)	5 (7%)	0	100	100
7	2e	79/92 (86%)	77 (98%)	2 (2%)	0	100	100
7	4e	74/92 (80%)	71 (96%)	3 (4%)	0	100	100
7	5e	75/92 (82%)	72 (96%)	3 (4%)	0	100	100
9	1K	199/437 (46%)	184 (92%)	12 (6%)	3 (2%)	8	33
10	4C	293/499 (59%)	275 (94%)	18 (6%)	0	100	100
11	11	79/119 (66%)	77 (98%)	2 (2%)	0	100	100
11	21	78/119 (66%)	75 (96%)	3 (4%)	0	100	100
11	41	79/119 (66%)	75 (95%)	4 (5%)	0	100	100
11	51	79/119 (66%)	75 (95%)	4 (5%)	0	100	100
12	R	104/480 (22%)	91 (88%)	13 (12%)	0	100	100
13	1f	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
13	2f	70/86 (81%)	69 (99%)	1 (1%)	0	100	100
13	4f	70/86 (81%)	69 (99%)	1 (1%)	0	100	100
13	5f	71/86 (83%)	64 (90%)	7 (10%)	0	100	100
14	66	70/80 (88%)	69 (99%)	0	1 (1%)	9	34
15	X	45/155 (29%)	39 (87%)	6 (13%)	0	100	100
16	12	91/118 (77%)	85 (93%)	6 (7%)	0	100	100
16	22	91/118 (77%)	86 (94%)	5 (6%)	0	100	100
16	42	90/118 (76%)	84 (93%)	6 (7%)	0	100	100
16	52	94/118 (80%)	87 (93%)	7 (7%)	0	100	100
18	67	75/103 (73%)	72 (96%)	2 (3%)	1 (1%)	9	35
19	62	93/95 (98%)	84 (90%)	6 (6%)	3 (3%)	3	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	2B	90/225 (40%)	88 (98%)	2 (2%)	0	100	100
21	A2	138/209 (66%)	123 (89%)	11 (8%)	4 (3%)	3	20
22	B2	204/895 (23%)	181 (89%)	22 (11%)	1 (0%)	24	55
23	5C	850/854 (100%)	817 (96%)	31 (4%)	2 (0%)	43	71
24	5X	574/820 (70%)	561 (98%)	13 (2%)	0	100	100
25	1b	84/240 (35%)	82 (98%)	2 (2%)	0	100	100
25	2b	80/240 (33%)	74 (92%)	6 (8%)	0	100	100
25	4b	80/240 (33%)	71 (89%)	9 (11%)	0	100	100
25	5b	69/240 (29%)	67 (97%)	2 (3%)	0	100	100
26	B5	67/86 (78%)	61 (91%)	6 (9%)	0	100	100
27	1A	96/282 (34%)	94 (98%)	2 (2%)	0	100	100
28	S	110/800 (14%)	101 (92%)	8 (7%)	1 (1%)	14	43
29	5J	793/850 (93%)	748 (94%)	44 (6%)	1 (0%)	48	76
30	4D	121/128 (94%)	119 (98%)	2 (2%)	0	100	100
31	63	83/102 (81%)	79 (95%)	3 (4%)	1 (1%)	10	37
33	B3	1176/1217 (97%)	1082 (92%)	92 (8%)	2 (0%)	43	71
34	1g	71/76 (93%)	69 (97%)	2 (3%)	0	100	100
34	2g	71/76 (93%)	69 (97%)	2 (3%)	0	100	100
34	4g	72/76 (95%)	66 (92%)	6 (8%)	0	100	100
34	5g	72/76 (95%)	66 (92%)	6 (8%)	0	100	100
35	68	93/96 (97%)	81 (87%)	6 (6%)	6 (6%)	1	7
36	5A	2198/2311 (95%)	2094 (95%)	102 (5%)	2 (0%)	48	76
37	A3	377/501 (75%)	346 (92%)	26 (7%)	5 (1%)	9	35
38	U	454/555 (82%)	424 (93%)	27 (6%)	3 (1%)	18	48
39	5D	139/142 (98%)	131 (94%)	8 (6%)	0	100	100
40	64	71/139 (51%)	66 (93%)	3 (4%)	2 (3%)	4	21
41	BP	98/104 (94%)	92 (94%)	6 (6%)	0	100	100
42	1C	48/159 (30%)	47 (98%)	1 (2%)	0	100	100
43	K	316/1007 (31%)	294 (93%)	18 (6%)	4 (1%)	9	35
44	4A	229/683 (34%)	210 (92%)	18 (8%)	1 (0%)	30	60
46	2A	160/255 (63%)	147 (92%)	13 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
47	A1	138/647 (21%)	129 (94%)	7 (5%)	2 (1%)	9 34
48	65	74/91 (81%)	70 (95%)	2 (3%)	2 (3%)	4 21
49	5B	1989/2136 (93%)	1885 (95%)	103 (5%)	1 (0%)	48 76
50	B1	846/1304 (65%)	792 (94%)	53 (6%)	1 (0%)	48 76
All	All	15337/23178 (66%)	14438 (94%)	846 (6%)	53 (0%)	37 66

5 of 53 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	5O	59	ILE
9	1K	59	ARG
9	1K	63	ARG
19	62	47	ASP
19	62	53	HIS

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
3	5O	263/300 (88%)	262 (100%)	1 (0%)	84 85
4	B4	66/336 (20%)	66 (100%)	0	100 100
5	13	71/101 (70%)	61 (86%)	10 (14%)	3 16
5	23	73/101 (72%)	73 (100%)	0	100 100
5	43	73/101 (72%)	73 (100%)	0	100 100
5	53	73/101 (72%)	73 (100%)	0	100 100
6	4B	306/442 (69%)	301 (98%)	5 (2%)	55 72
7	1e	72/84 (86%)	61 (85%)	11 (15%)	3 13
7	2e	76/84 (90%)	76 (100%)	0	100 100
7	4e	71/84 (84%)	71 (100%)	0	100 100
7	5e	72/84 (86%)	72 (100%)	0	100 100
9	1K	170/373 (46%)	154 (91%)	16 (9%)	8 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	4C	255/424 (60%)	253 (99%)	2 (1%)	73	79
11	11	76/101 (75%)	72 (95%)	4 (5%)	20	49
11	21	75/101 (74%)	75 (100%)	0	100	100
11	41	76/101 (75%)	53 (70%)	23 (30%)	0	1
11	51	76/101 (75%)	53 (70%)	23 (30%)	0	1
12	R	94/369 (26%)	93 (99%)	1 (1%)	65	76
13	1f	63/74 (85%)	56 (89%)	7 (11%)	6	23
13	2f	61/74 (82%)	61 (100%)	0	100	100
13	4f	61/74 (82%)	61 (100%)	0	100	100
13	5f	61/74 (82%)	60 (98%)	1 (2%)	55	72
14	66	62/70 (89%)	62 (100%)	0	100	100
15	X	42/144 (29%)	42 (100%)	0	100	100
16	12	91/110 (83%)	82 (90%)	9 (10%)	7	27
16	22	91/110 (83%)	91 (100%)	0	100	100
16	42	86/110 (78%)	86 (100%)	0	100	100
16	52	93/110 (84%)	92 (99%)	1 (1%)	65	76
18	67	69/91 (76%)	67 (97%)	2 (3%)	37	62
19	62	88/88 (100%)	86 (98%)	2 (2%)	44	66
20	2B	81/195 (42%)	81 (100%)	0	100	100
21	A2	129/180 (72%)	127 (98%)	2 (2%)	55	72
22	B2	187/776 (24%)	183 (98%)	4 (2%)	47	67
23	5C	754/756 (100%)	750 (100%)	4 (0%)	81	83
24	5X	517/721 (72%)	500 (97%)	17 (3%)	33	59
25	1b	78/177 (44%)	69 (88%)	9 (12%)	5	22
25	2b	75/177 (42%)	75 (100%)	0	100	100
25	4b	75/177 (42%)	75 (100%)	0	100	100
25	5b	67/177 (38%)	67 (100%)	0	100	100
26	B5	60/77 (78%)	60 (100%)	0	100	100
27	1A	85/240 (35%)	81 (95%)	4 (5%)	23	52
28	S	91/681 (13%)	90 (99%)	1 (1%)	65	76
29	5J	636/715 (89%)	632 (99%)	4 (1%)	78	82

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
30	4D	107/111 (96%)	107 (100%)	0	100	100
31	63	79/94 (84%)	76 (96%)	3 (4%)	29	56
33	B3	1027/1051 (98%)	1018 (99%)	9 (1%)	70	78
34	1g	63/66 (96%)	54 (86%)	9 (14%)	3	15
34	2g	63/66 (96%)	61 (97%)	2 (3%)	34	60
34	4g	64/66 (97%)	43 (67%)	21 (33%)	0	1
34	5g	64/66 (97%)	43 (67%)	21 (33%)	0	1
35	68	81/82 (99%)	77 (95%)	4 (5%)	22	51
36	5A	2002/2090 (96%)	1989 (99%)	13 (1%)	78	82
37	A3	345/446 (77%)	337 (98%)	8 (2%)	44	66
38	U	418/503 (83%)	418 (100%)	0	100	100
39	5D	129/130 (99%)	129 (100%)	0	100	100
40	64	68/111 (61%)	67 (98%)	1 (2%)	57	73
41	BP	86/90 (96%)	86 (100%)	0	100	100
42	1C	48/135 (36%)	41 (85%)	7 (15%)	3	15
43	K	291/919 (32%)	257 (88%)	34 (12%)	5	21
44	4A	210/599 (35%)	209 (100%)	1 (0%)	81	83
46	2A	139/218 (64%)	138 (99%)	1 (1%)	76	80
47	A1	130/550 (24%)	126 (97%)	4 (3%)	35	61
48	65	68/80 (85%)	67 (98%)	1 (2%)	57	73
49	5B	1779/1908 (93%)	1769 (99%)	10 (1%)	78	82
50	B1	733/1104 (66%)	731 (100%)	2 (0%)	86	86
All	All	13735/20051 (68%)	13421 (98%)	314 (2%)	44	66

5 of 314 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
34	5g	53	GLN
34	4g	51	SER
34	5g	73	LEU
43	K	862	ILE
47	A1	238	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 266 such sidechains are listed below:

Mol	Chain	Res	Type
49	5B	1113	ASN
49	5B	1402	ASN
50	B1	1107	GLN
30	4D	111	GLN
11	51	12	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	1	163/164 (99%)	49 (30%)	6 (3%)
17	5	101/117 (86%)	42 (41%)	4 (3%)
2	6	50/106 (47%)	5 (10%)	2 (4%)
32	2	90/188 (47%)	23 (25%)	4 (4%)
45	4	120/146 (82%)	25 (20%)	3 (2%)
8	I	23/62 (37%)	5 (21%)	0
All	All	547/783 (69%)	149 (27%)	19 (3%)

5 of 149 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	1	10	U
1	1	12	G
1	1	15	G
1	1	16	G
1	1	17	G

5 of 19 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
32	2	103	U
45	4	68	A
45	4	114	U
45	4	1	A
17	5	57	G

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
53	GTP	5C	1002	52	33,34,34	0.99	2 (6%)	50,54,54	1.64	9 (18%)
54	IHP	5A	2401	-	36,36,36	0.71	0	60,60,60	0.76	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
53	GTP	5C	1002	52	-	4/22/38/38	0/3/3/3
54	IHP	5A	2401	-	-	6/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
53	5C	1002	GTP	O4'-C4'	-2.12	1.40	1.45
53	5C	1002	GTP	C5-N7	-2.01	1.35	1.39

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
53	5C	1002	GTP	C5-C4-N3	-4.53	121.18	128.39
53	5C	1002	GTP	C2-N3-C4	4.47	119.99	112.30
53	5C	1002	GTP	C2-N1-C6	-3.05	119.58	125.11
53	5C	1002	GTP	N9-C8-N7	-3.00	107.85	113.40
53	5C	1002	GTP	N9-C4-N3	2.82	131.60	125.95

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

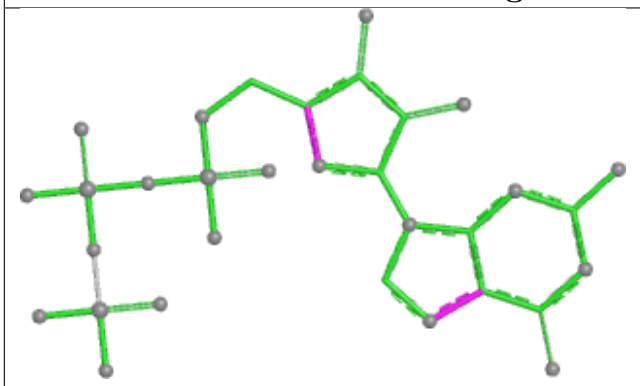
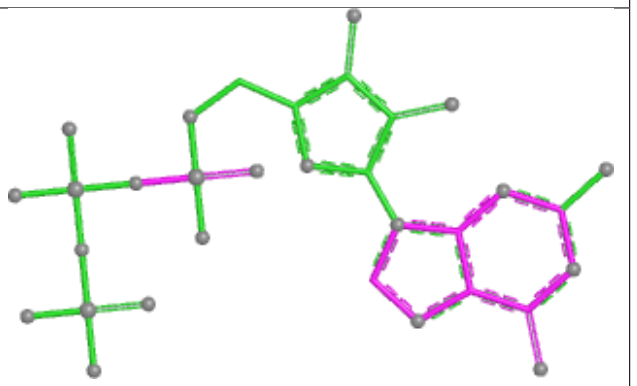
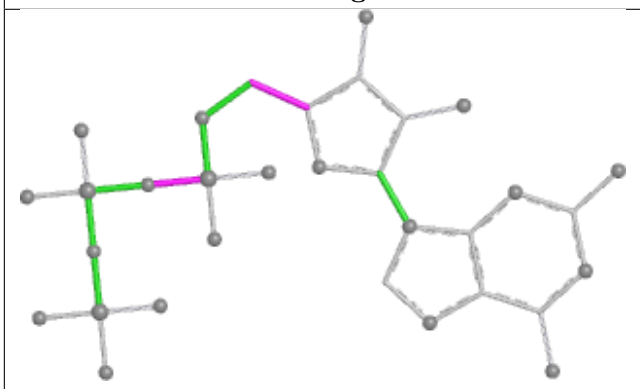
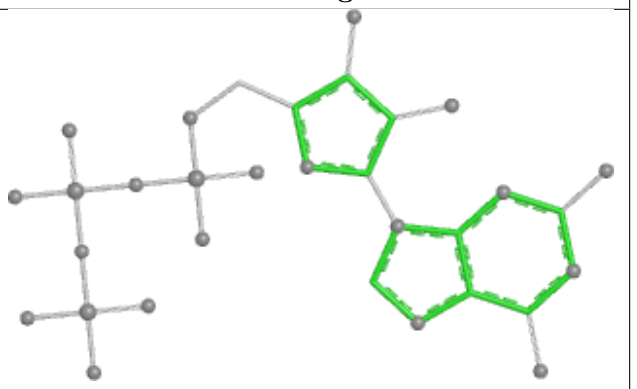
Mol	Chain	Res	Type	Atoms
53	5C	1002	GTP	O4'-C4'-C5'-O5'
53	5C	1002	GTP	PB-O3A-PA-O2A
53	5C	1002	GTP	C3'-C4'-C5'-O5'
54	5A	2401	IHP	C1-O11-P1-O41
54	5A	2401	IHP	C3-O13-P3-O23

There are no ring outliers.

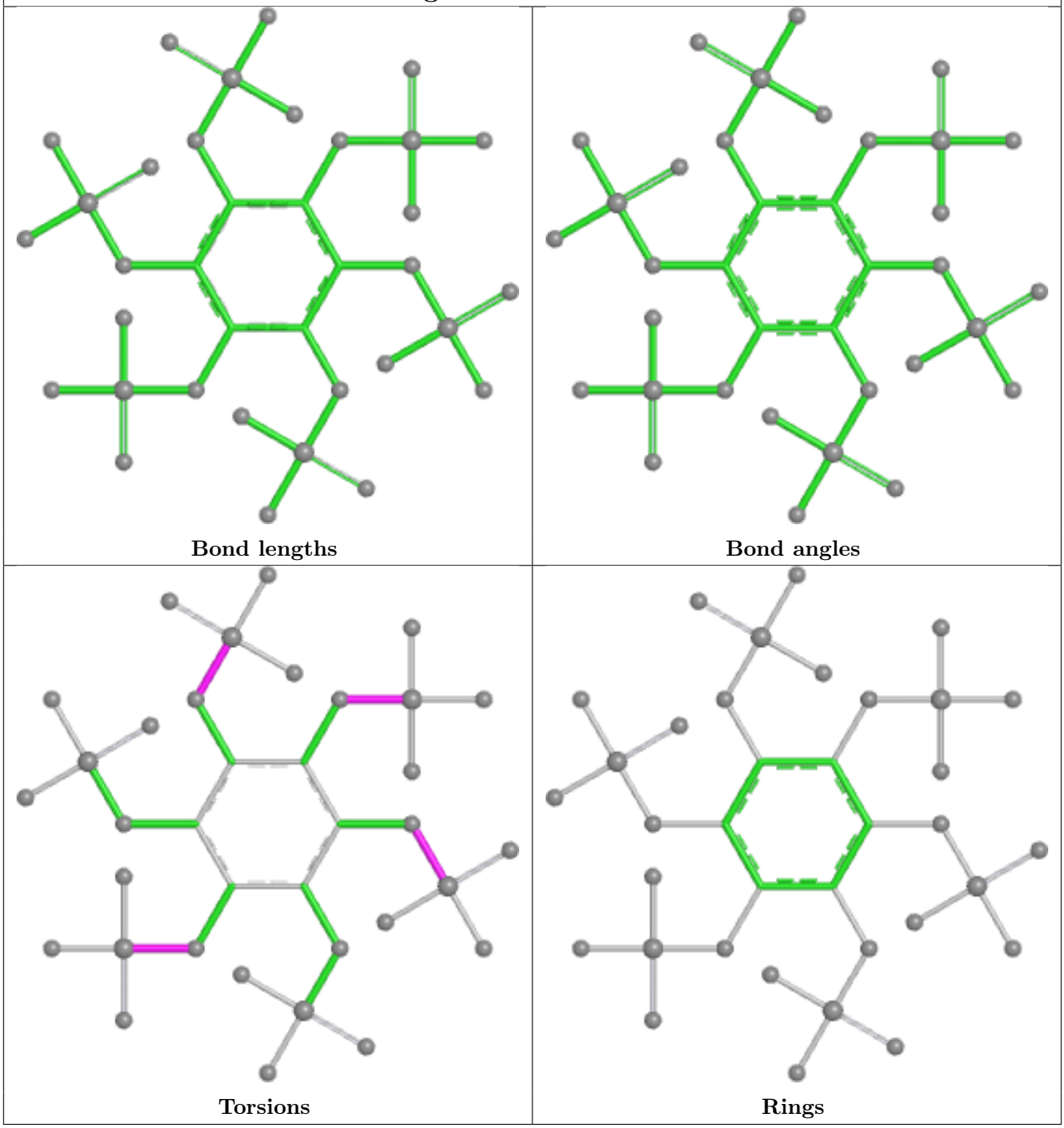
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
54	5A	2401	IHP	1	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 then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand GTP 5C 1002	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand IHP 5A 2401



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
47	A1	2
1	1	2
29	5J	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A1	282:VAL	C	422:UNK	N	108.26
1	5J	165:ASP	C	236:GLY	N	34.19
1	A1	447:UNK	C	455:VAL	N	4.52
1	1	2:U	O3'	3:A	P	1.39
1	1	78:U	O3'	79:G	P	1.37

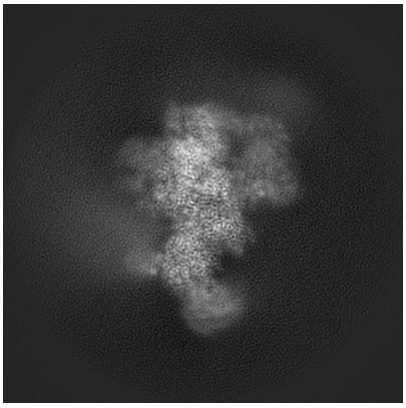
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-4665. These allow visual inspection of the internal detail of the map and identification of artifacts.

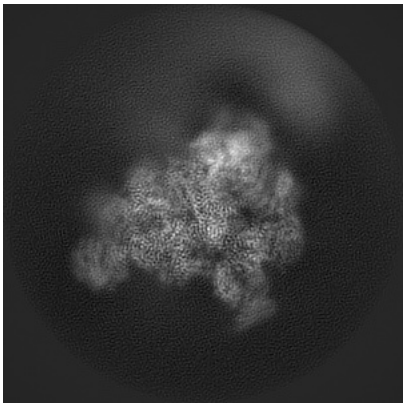
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

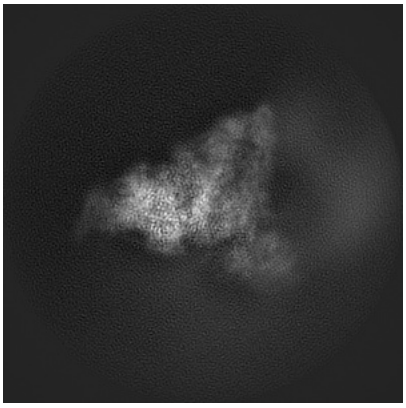
6.1.1 Primary map



X



Y

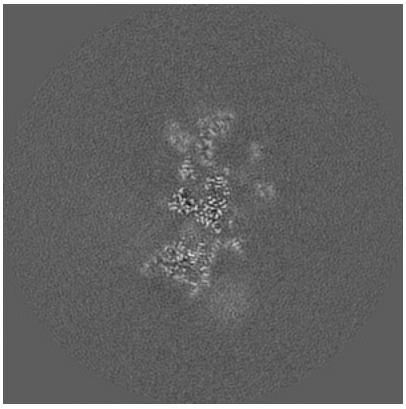


Z

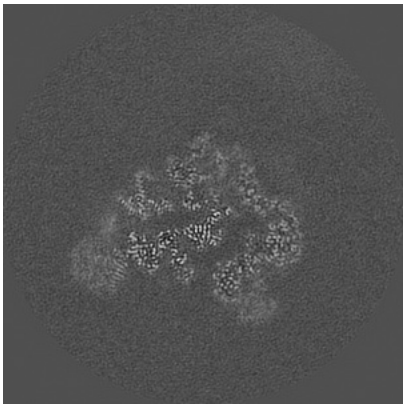
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

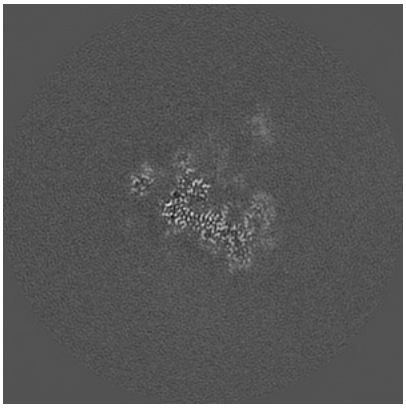
6.2.1 Primary map



X Index: 210



Y Index: 210

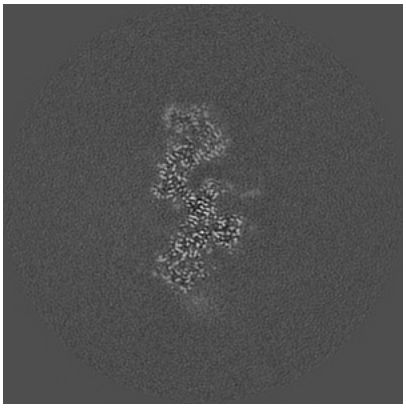


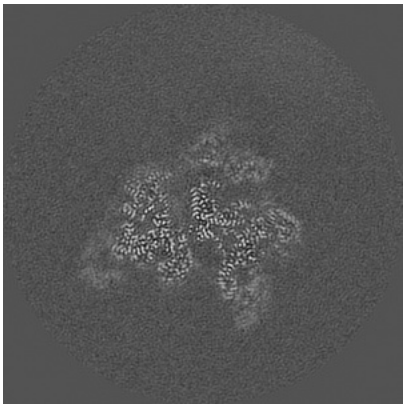
Z Index: 210

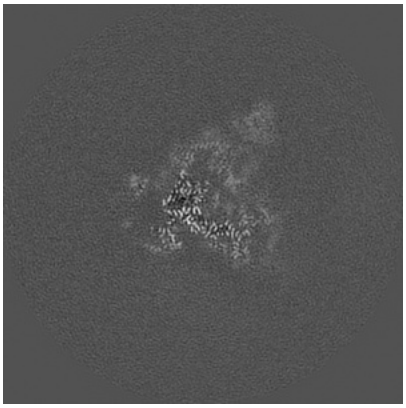
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map


X Index: 177

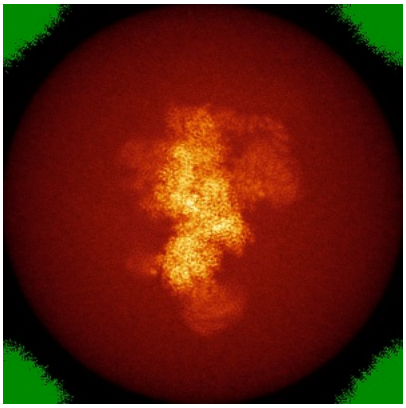

Y Index: 193

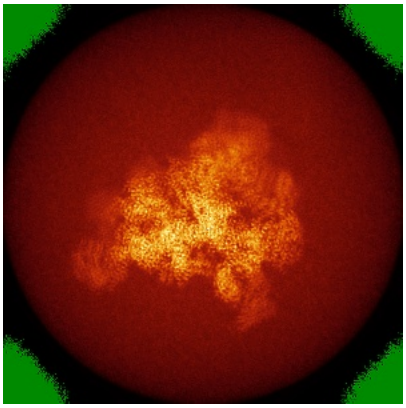

Z Index: 218

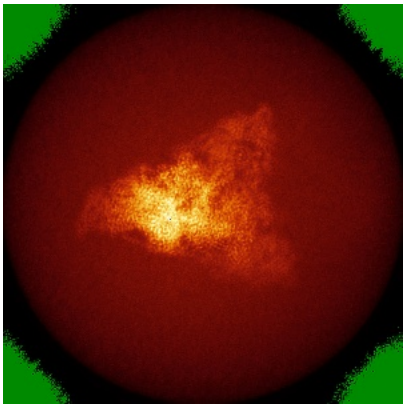
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map


X

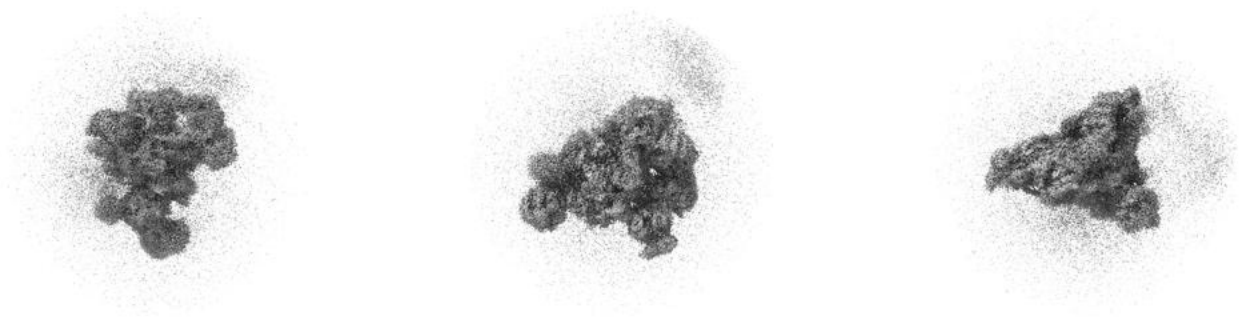

Y


Z

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

6.5 Orthogonal surface views (i)

6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.01. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

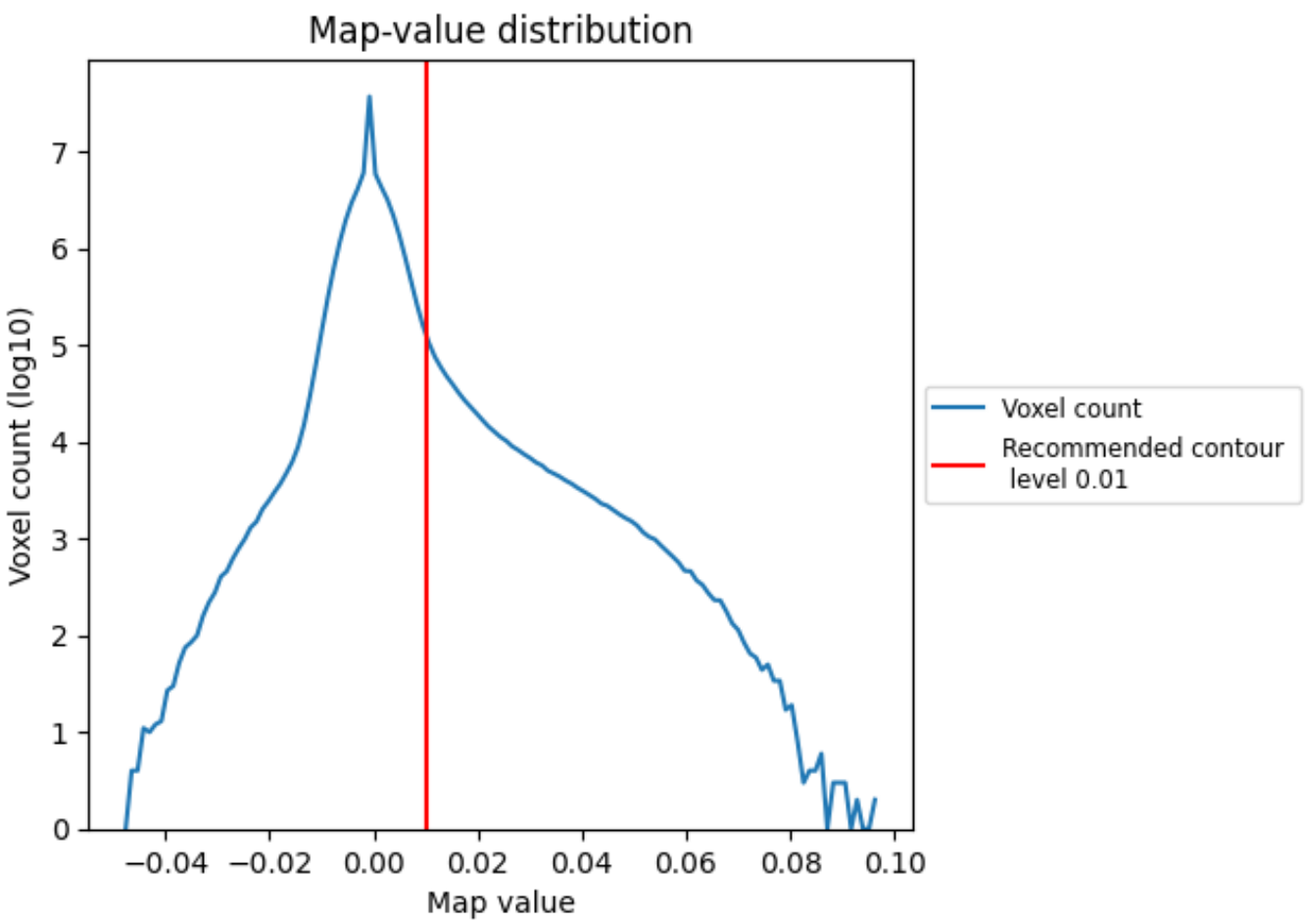
6.6 Mask visualisation (i)

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

7 Map analysis (i)

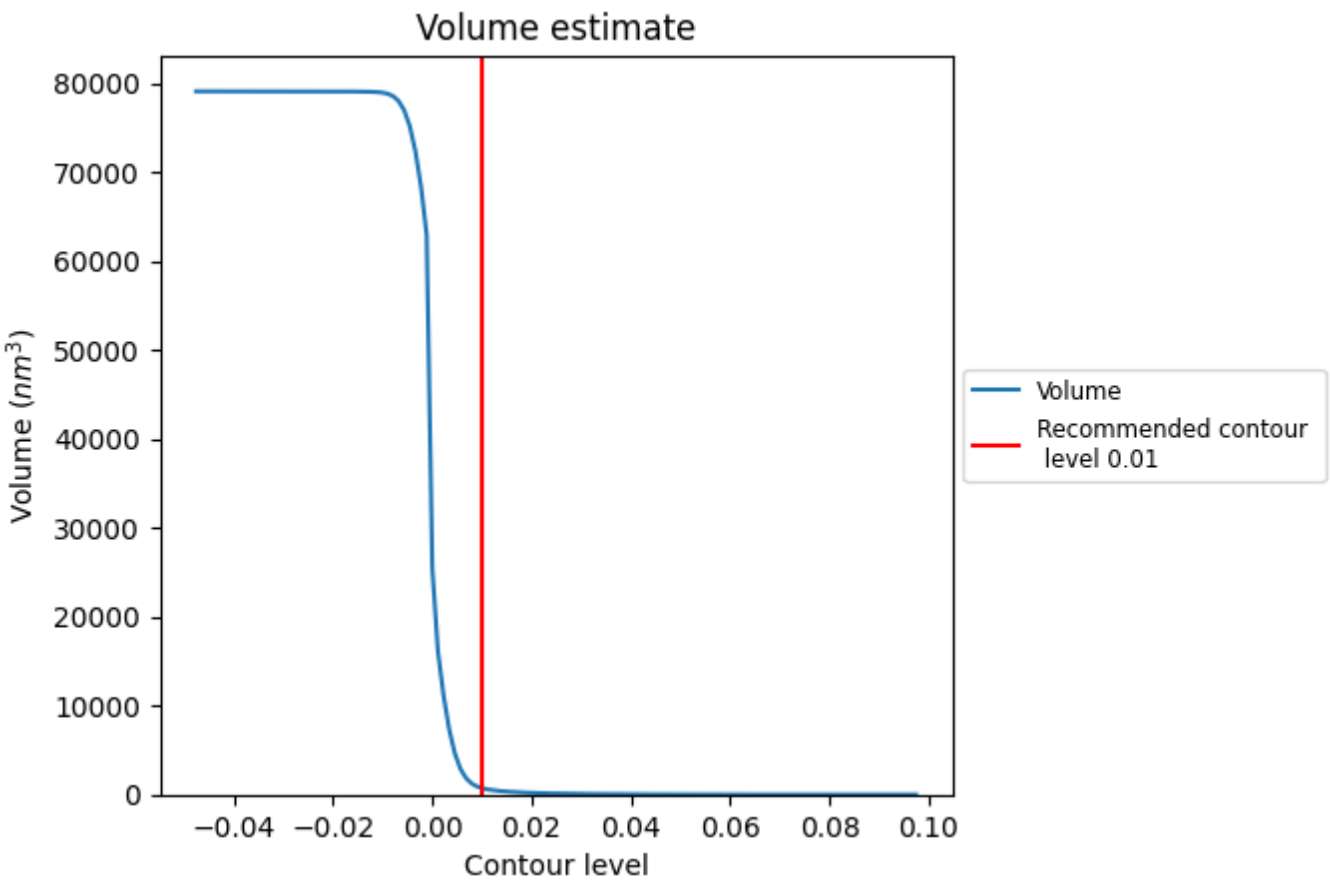
This section contains the results of statistical analysis of the map.

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.

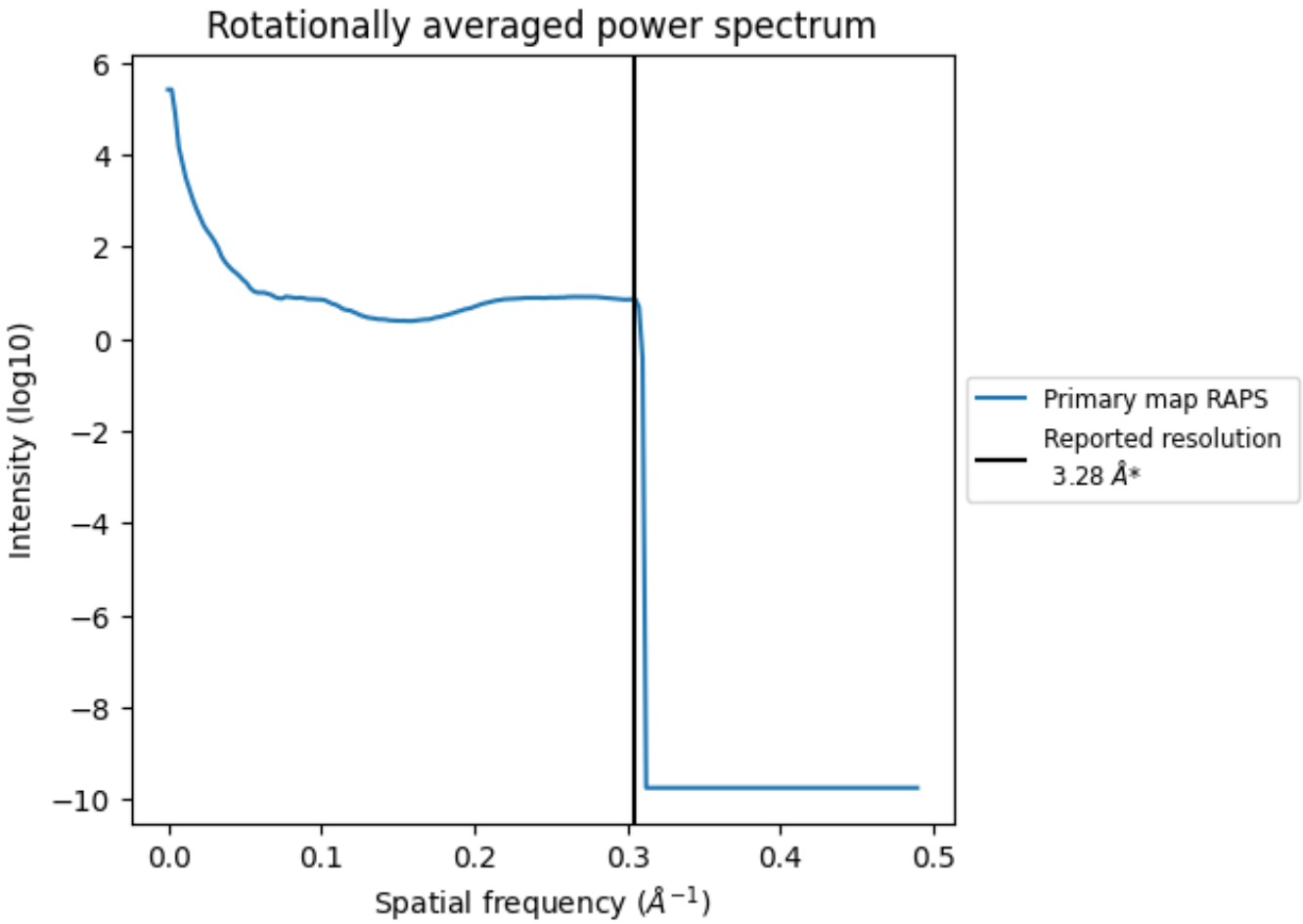
7.2 Volume estimate (i)



The volume at the recommended contour level is 726 nm^3; this corresponds to an approximate mass of 656 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)

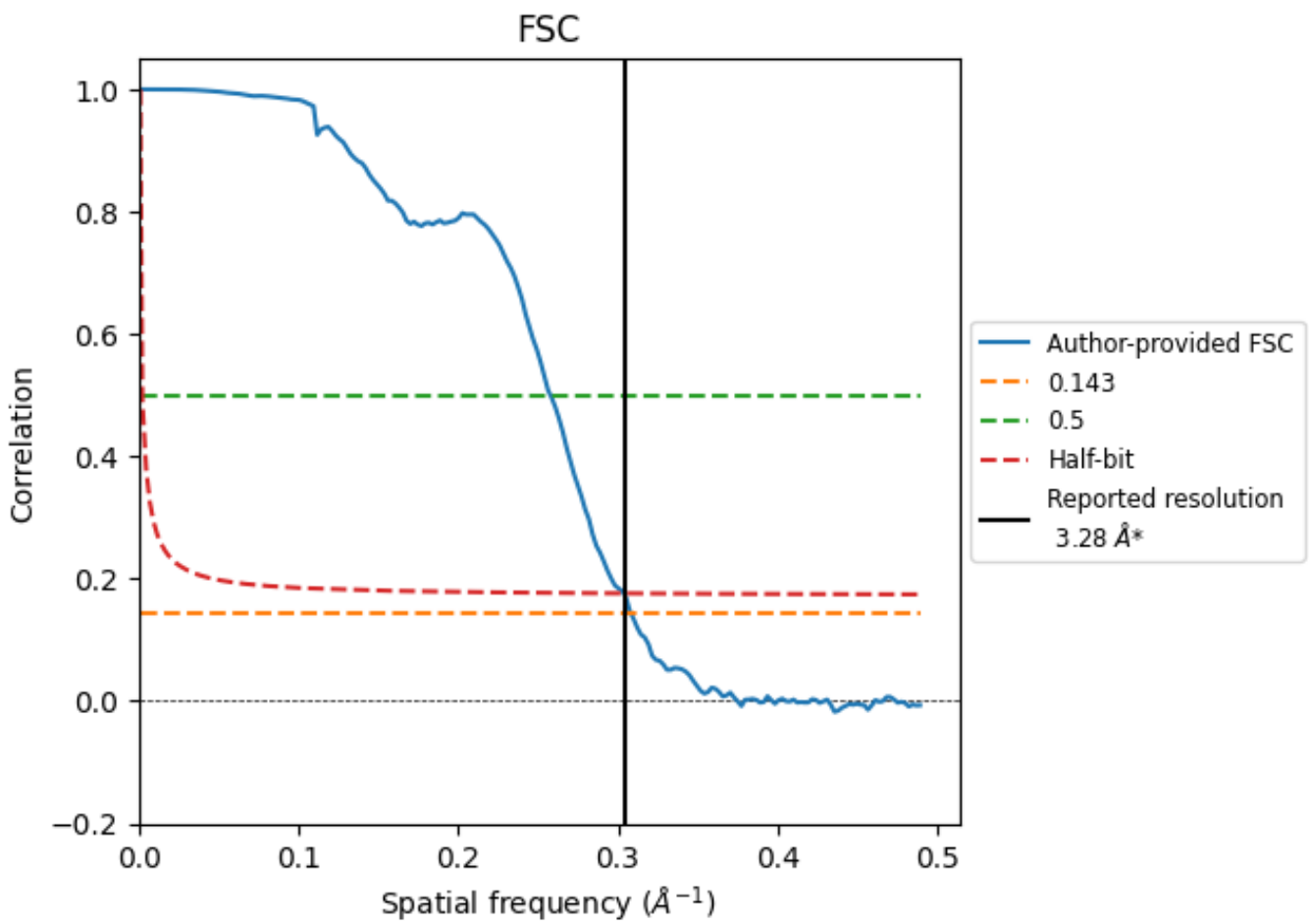


*Reported resolution corresponds to spatial frequency of 0.305 Å^{-1}

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 0.305 Å^{-1}

8.2 Resolution estimates (i)

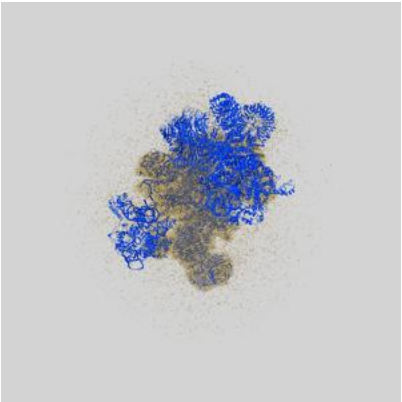
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.28	-	-
Author-provided FSC curve	3.24	3.88	3.29
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

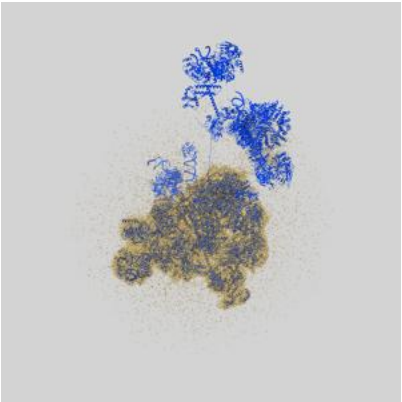
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-4665 and PDB model 6QX9. Per-residue inclusion information can be found in section 3 on page 16.

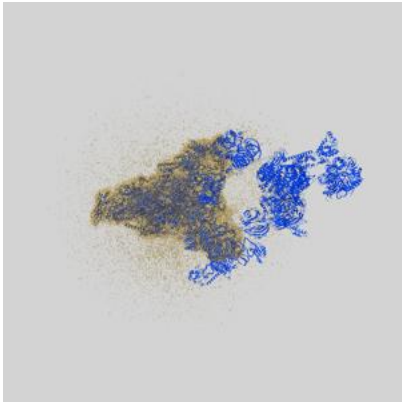
9.1 Map-model overlay (i)



X



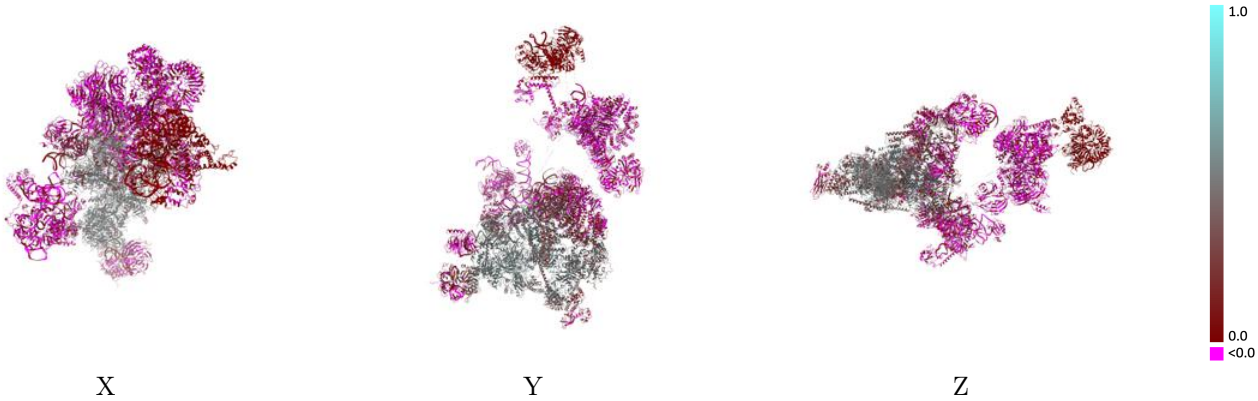
Y



Z

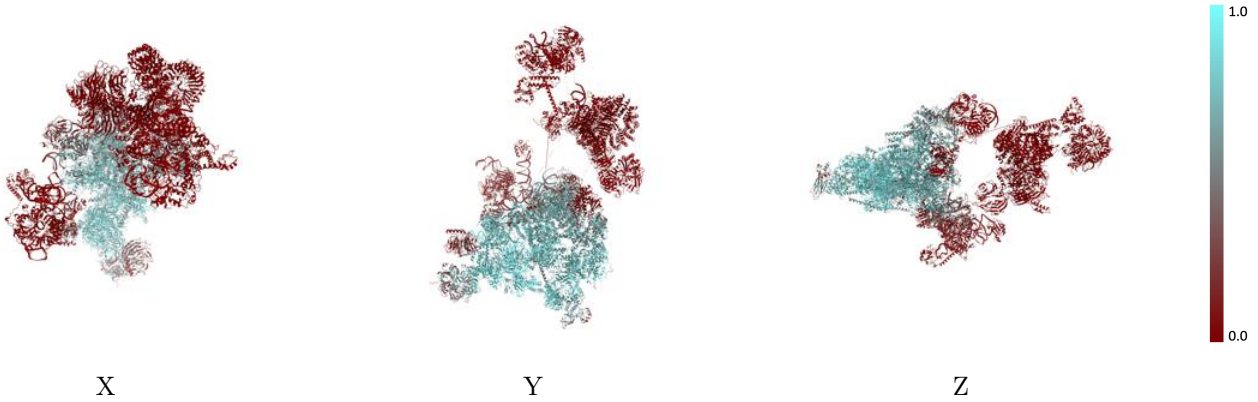
The images above show the 3D surface view of the map at the recommended contour level 0.01 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



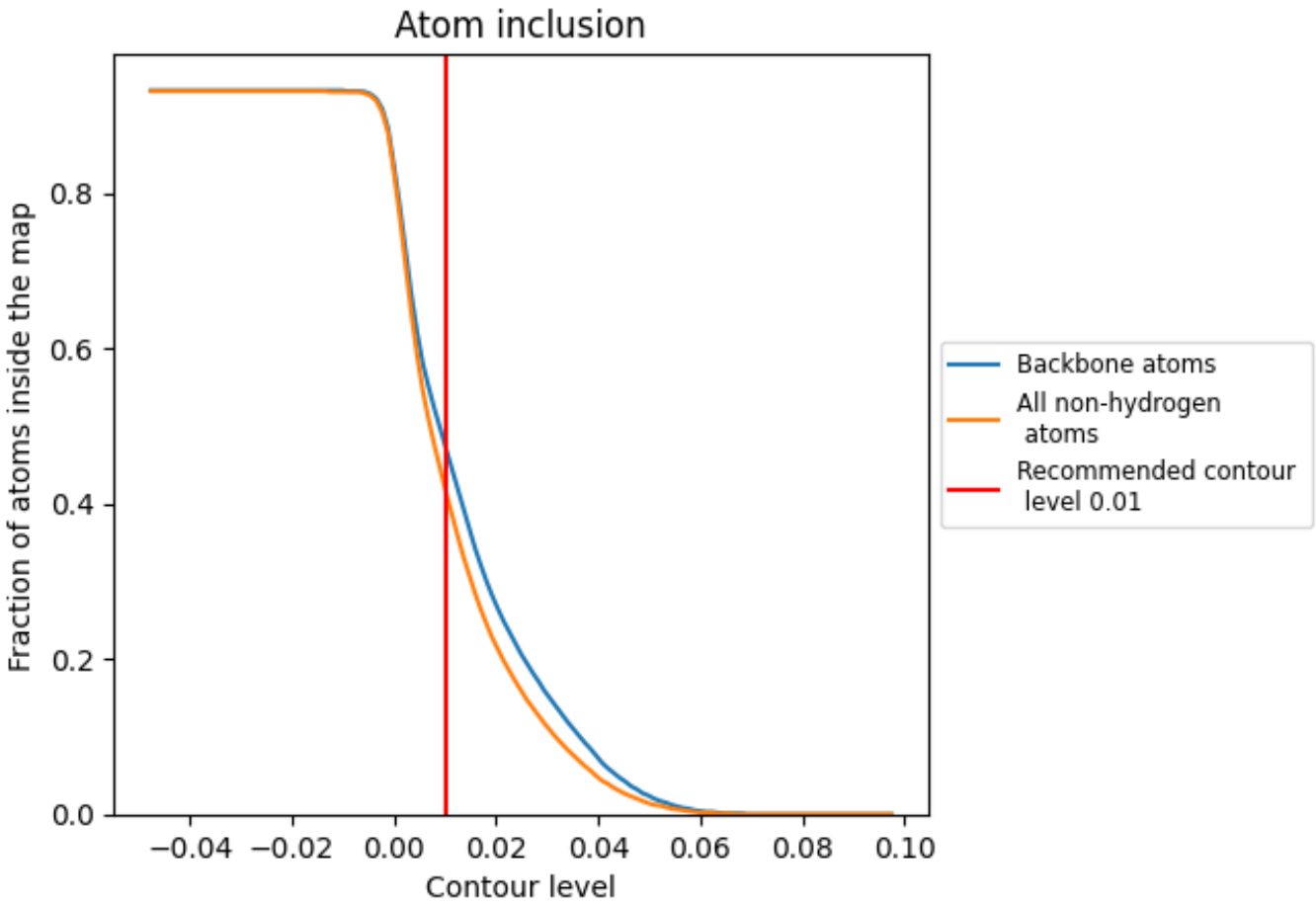
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.01).

9.4 Atom inclusion (i)



At the recommended contour level, 48% of all backbone atoms, 42% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.01) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.4190	0.2070
1	0.0020	0.0030
11	0.0000	-0.0020
12	0.0000	0.0170
13	0.0000	-0.0140
1A	0.0000	0.0280
1C	0.0000	0.0100
1K	0.0010	-0.0000
1b	0.0030	0.0140
1e	0.0000	0.0140
1f	0.0020	0.0020
1g	0.0000	0.0360
2	0.0010	0.0120
21	0.0000	0.0000
22	0.0000	0.0000
23	0.0000	0.0000
2A	0.0000	0.0000
2B	0.0000	0.0000
2b	0.0000	0.0000
2e	0.0000	0.0000
2f	0.0000	0.0000
2g	0.0000	0.0000
4	0.6370	0.2230
41	0.4710	0.2010
42	0.4720	0.2320
43	0.2390	0.0480
4A	0.4880	0.1430
4B	0.5560	0.1550
4C	0.6880	0.3260
4D	0.6860	0.3020
4b	0.2840	0.0870
4e	0.2410	0.0760
4f	0.3330	0.1270
4g	0.2040	0.0230
5	0.7530	0.2980

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
51	0.4680	0.1340
52	0.4130	0.1040
53	0.7280	0.3830
5A	0.8230	0.4690
5B	0.7650	0.3870
5C	0.8730	0.5100
5D	0.8170	0.4650
5J	0.6500	0.2450
5O	0.1710	0.0160
5X	0.3060	0.1470
5b	0.5770	0.2660
5e	0.4730	0.1160
5f	0.3570	0.0930
5g	0.5680	0.2690
6	0.5610	0.1990
62	0.0050	0.0140
63	0.0100	-0.0020
64	0.0000	-0.0470
65	0.0020	0.0290
66	0.0000	-0.0070
67	0.0020	-0.0100
68	0.0010	0.0250
A1	0.1030	0.0580
A2	0.0000	-0.0020
A3	0.0000	0.0040
B1	0.0010	0.0020
B2	0.0020	0.0090
B3	0.0010	-0.0030
B4	0.0000	-0.0010
B5	0.0040	-0.0090
BP	0.0010	-0.0110
I	0.0040	-0.0000
K	0.0120	0.0130
R	0.6000	0.2310
S	0.6350	0.3130
U	0.8800	0.5060
X	0.5430	0.2760