wwPDB EM Validation Summary Report (i)

Nov 3, 2024 – 11:18 pm GMT

PDB ID	:	7B9V
EMDB ID	:	EMD-12106
Title	:	Yeast C complex spliceosome at 2.8 Angstrom resolution with Prp18/Slu7 bound
Authors	:	Wilkinson, M.E.; Fica, S.M.; Galej, W.P.; Nagai, K.
Deposited on	:	2020-12-14
Resolution	:	2.80 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 (1)) were used in the production of this report:

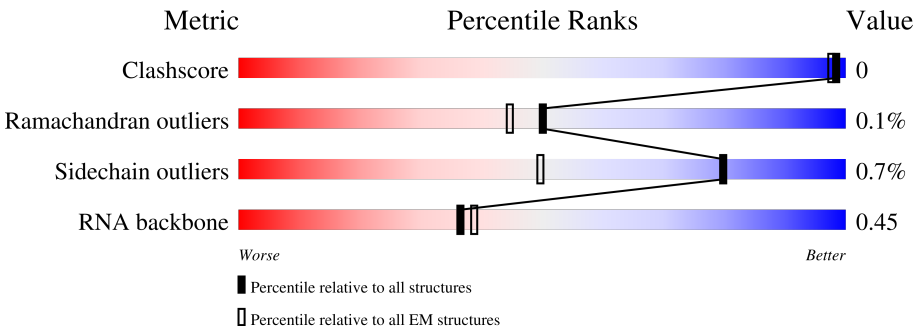
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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

Mol	Chain	Length	Quality of chain
1	2	1175	6%; 9%; 5%; •; 83%
2	5	214	39%; 47%; 30%; 6%; 17%
3	6	112	16%; 67%; 20%; •; 9%
4	A	2413	9%; 90%; •; 9%
5	B	2163	25%; 77%; •; 21%
6	C	1008	•; 88%; •; 11%
7	D	291	9%; 66%; 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	E	47	21% 6% 70%
9	F	179	9% 36% 64%
10	G	235	6% 58% 41%
11	H	577	11% 77% 21%
12	I	95	24% 26% 19% 13% 42%
13	J	451	80% 18%
14	K	379	6% 47% 52%
15	L	157	99%
16	M	339	74% 25%
17	N	364	13% 72% 27%
18	O	590	7% 47% 53%
19	P	175	41% 58%
20	Q	1071	10% 55% 42%
21	R	135	37% 59% 40%
22	S	687	6% 78% 20%
23	T	859	5% 81% 16%
24	W	238	87% 10%
25	X	240	5% 30% 70%
26	Y	111	77% 21%
27	Z	140	39% 61%
28	a	251	55% 55% 45%
29	b	196	5% 46% 53%
29	k	196	52% 48%
30	c	382	8% 10% 90%
31	d	101	81% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	n	101	80% 19%
32	e	94	14% 80% 18%
32	p	94	80% 18%
33	f	86	14% 87% 13%
33	q	86	86% 13%
34	g	77	13% 88% 5% 6%
34	r	77	92% 6%
35	h	146	14% 71% 27%
35	l	146	71% 27%
36	j	110	20% 85% 12%
36	m	110	85% 12%
37	o	455	48% 70% 27%
38	s	175	6% 31% 65%
39	t	503	12% 21% 79%
39	u	503	8% 22% 77%
39	v	503	22% 78%
39	w	503	50% 84% 13%
40	y	215	9% 61% 38%

2 Entry composition (i)

There are 45 unique types of molecules in this entry. The entry contains 109105 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	2	196	4120	1846	681	1399	194	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	5	178	3777	1691	660	1249	177	0	0

- Molecule 3 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	6	102	2170	972	386	710	102	0	0

- Molecule 4 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	2191	18036	11598	3079	3295	64	0	0

- Molecule 5 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	B	1707	13675	8758	2279	2583	55	0	0

- Molecule 6 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	C	898	7139	4614	1189	1309	27	0	0

- Molecule 7 is a protein called Splicing factor YJU2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	D	194	1547	956	280	298	13	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	218	UNK	-	insertion	UNP A0A6L1B9A1
D	219	UNK	-	insertion	UNP A0A6L1B9A1
D	220	UNK	-	insertion	UNP A0A6L1B9A1
D	221	UNK	-	insertion	UNP A0A6L1B9A1
D	222	UNK	-	insertion	UNP A0A6L1B9A1
D	223	UNK	-	insertion	UNP A0A6L1B9A1
D	224	UNK	-	insertion	UNP A0A6L1B9A1
D	225	UNK	-	insertion	UNP A0A6L1B9A1
D	226	UNK	-	insertion	UNP A0A6L1B9A1
D	227	UNK	-	insertion	UNP A0A6L1B9A1
D	228	UNK	-	insertion	UNP A0A6L1B9A1
D	229	UNK	-	insertion	UNP A0A6L1B9A1
D	230	UNK	-	insertion	UNP A0A6L1B9A1
D	231	UNK	-	insertion	UNP A0A6L1B9A1
D	232	UNK	-	insertion	UNP A0A6L1B9A1
D	233	UNK	-	insertion	UNP A0A6L1B9A1
D	234	UNK	-	insertion	UNP A0A6L1B9A1
D	?	-	ASP	deletion	UNP A0A6L1B9A1
D	?	-	ASN	deletion	UNP A0A6L1B9A1
D	?	-	ASN	deletion	UNP A0A6L1B9A1
D	?	-	ASP	deletion	UNP A0A6L1B9A1

- Molecule 8 is a RNA chain called 5' exon of UBC4 mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	E	14	304	136	59	95	14	0	0

- Molecule 9 is a protein called Pre-mRNA-splicing factor CWC25.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	F	64	505	314	92	98	1	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	72	UNK	LYS	conflict	UNP A0A6A5Q526
F	73	UNK	LYS	conflict	UNP A0A6A5Q526
F	74	UNK	SER	conflict	UNP A0A6A5Q526
F	75	UNK	GLY	conflict	UNP A0A6A5Q526
F	76	UNK	LEU	conflict	UNP A0A6A5Q526
F	77	UNK	GLU	conflict	UNP A0A6A5Q526
F	78	UNK	TRP	conflict	UNP A0A6A5Q526
F	79	UNK	MET	conflict	UNP A0A6A5Q526
F	80	UNK	TYR	conflict	UNP A0A6A5Q526
F	81	UNK	GLN	conflict	UNP A0A6A5Q526

- Molecule 10 is a protein called Pre-mRNA-splicing factor ISY1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	G	138	1090	679	209	199	3	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	163	UNK	ARG	conflict	UNP A0A6V8S636
G	164	UNK	ASN	conflict	UNP A0A6V8S636
G	165	UNK	ASP	conflict	UNP A0A6V8S636
G	166	UNK	PHE	conflict	UNP A0A6V8S636
G	167	UNK	TYR	conflict	UNP A0A6V8S636
G	168	UNK	TYR	conflict	UNP A0A6V8S636
G	169	UNK	HIS	conflict	UNP A0A6V8S636
G	170	UNK	GLY	conflict	UNP A0A6V8S636
G	171	UNK	LYS	conflict	UNP A0A6V8S636
G	172	UNK	VAL	conflict	UNP A0A6V8S636
G	173	UNK	THR	conflict	UNP A0A6V8S636

- Molecule 11 is a protein called CWC22 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	H	453	3705	2383	612	692	18	0	0

- Molecule 12 is a RNA chain called Branched intron and 3' exon of UBC4 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
12	I	55	1068	476	159	378	55	0	0

- Molecule 13 is a protein called BJ4_G0054360.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	J	370	2926	1849	513	553	11	0	0

- Molecule 14 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	K	182	1455	911	268	270	6	0	0

- Molecule 15 is a protein called BUD31 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	L	156	1283	803	239	231	10	0	0

- Molecule 16 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	M	255	2048	1297	362	378	11	0	0

- Molecule 17 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	N	264	2092	1331	364	382	15	0	0

- Molecule 18 is a protein called Y55_G0042700.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	O	280	2143	1347	390	399	7	0	0

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	125	ALA	VAL	conflict	UNP A0A6L0ZW46
O	304	UNK	THR	conflict	UNP A0A6L0ZW46
O	305	UNK	LYS	conflict	UNP A0A6L0ZW46
O	306	UNK	GLN	conflict	UNP A0A6L0ZW46

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	307	UNK	GLY	conflict	UNP A0A6L0ZW46
O	308	UNK	LYS	conflict	UNP A0A6L0ZW46
O	309	UNK	VAL	conflict	UNP A0A6L0ZW46
O	310	UNK	THR	conflict	UNP A0A6L0ZW46
O	311	UNK	TYR	conflict	UNP A0A6L0ZW46
O	312	UNK	LYS	conflict	UNP A0A6L0ZW46
O	313	UNK	LYS	conflict	UNP A0A6L0ZW46
O	314	UNK	LYS	conflict	UNP A0A6L0ZW46
O	315	UNK	LEU	conflict	UNP A0A6L0ZW46
O	316	UNK	GLU	conflict	UNP A0A6L0ZW46
O	317	UNK	SER	conflict	UNP A0A6L0ZW46
O	318	UNK	LYS	conflict	UNP A0A6L0ZW46
O	319	UNK	ARG	conflict	UNP A0A6L0ZW46
O	320	UNK	GLN	conflict	UNP A0A6L0ZW46
O	321	UNK	LYS	conflict	UNP A0A6L0ZW46
O	322	UNK	LEU	conflict	UNP A0A6L0ZW46
O	323	UNK	ILE	conflict	UNP A0A6L0ZW46
O	324	UNK	GLU	conflict	UNP A0A6L0ZW46
O	325	UNK	ALA	conflict	UNP A0A6L0ZW46
O	326	UNK	GLN	conflict	UNP A0A6L0ZW46
O	327	UNK	ALA	conflict	UNP A0A6L0ZW46

- Molecule 19 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	P	74	607	382	120	104	1	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	49	ARG	LYS	conflict	UNP A0A6L0Y8G8
P	68	MET	VAL	conflict	UNP A0A6L0Y8G8
P	99	VAL	ILE	conflict	UNP A0A6L0Y8G8

- Molecule 20 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	Q	624	4959	3184	833	921	21	0	0

- Molecule 21 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	R	81	555	336	109	109	1	0	0

- Molecule 22 is a protein called CLF1 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	S	549	4170	2663	732	762	13	0	0

- Molecule 23 is a protein called SYF1 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	T	720	5387	3441	914	1015	17	0	0

- Molecule 24 is a protein called HLJ1_G0053790.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	W	214	1734	1084	317	324	9	0	0

- Molecule 25 is a protein called Unassigned structure.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
25	X	71	355	213	71	71	0	0

- Molecule 26 is a protein called BJ4_G0027490.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	Y	88	713	457	124	129	3	0	0

- Molecule 27 is a protein called NTC20 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	Z	55	446	276	85	83	2	0	0

- Molecule 28 is a protein called Pre-mRNA-splicing factor 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	a	138	1132	729	195	205	3	0	0

- Molecule 29 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	b	93	752	477	138	134	3	0	0
			Total	C	N	O	S		
29	k	102	830	526	155	146	3	0	0

- Molecule 30 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	c	37	293	180	52	56	5	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	9	GLU	LYS	conflict	UNP A0A6V8RGB0
c	154	ASN	THR	conflict	UNP A0A6V8RGB0

- Molecule 31 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	d	82	633	404	109	118	2	0	0
			Total	C	N	O	S		
31	n	82	633	404	109	118	2	0	0

- Molecule 32 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	e	77	606	398	94	111	3	0	0
			Total	C	N	O	S		
32	p	77	606	398	94	111	3	0	0

- Molecule 33 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
33	f	75	601	385	105	110	1	0	0
33	q	75	601	385	105	110	1	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
34	g	72	557	351	97	107	2	0	0
34	r	72	557	351	97	107	2	0	0

- Molecule 35 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
35	h	106	819	516	144	156	3	0	0
35	l	106	819	516	144	156	3	0	0

- Molecule 36 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
36	j	97	795	506	144	141	4	0	0
36	m	97	795	506	144	141	4	0	0

- Molecule 37 is a protein called CDC40 isoform 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
37	o	330	2673	1696	475	493	1	8	0	0

- Molecule 38 is a protein called SNT309 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
38	s	61	426	277	76	72	1	0	0

- Molecule 39 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	t	107	Total 794	C 511	N 128	O 152	S 3	0	0
39	u	117	Total 856	C 551	N 138	O 164	S 3	0	0
39	v	113	Total 827	C 532	N 134	O 158	S 3	0	0
39	w	438	Total 3405	C 2167	N 551	O 670	S 17	0	0

- Molecule 40 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	y	134	Total 1003	C 618	N 187	O 197	S 1	0	0

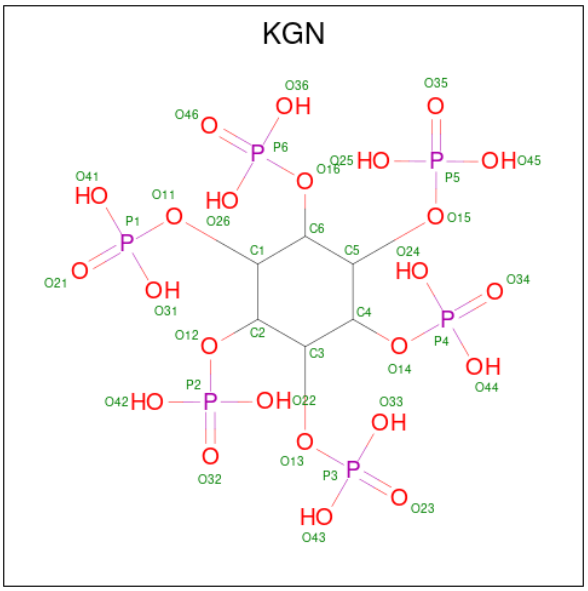
- Molecule 41 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
41	6	5	Total 5	Mg 5	0
41	C	1	Total 1	Mg 1	0

- Molecule 42 is POTASSIUM ION (three-letter code: K) (formula: K).

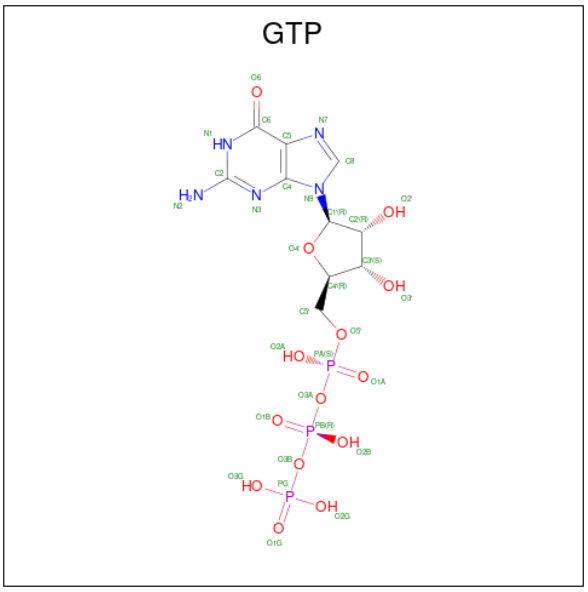
Mol	Chain	Residues	Atoms		AltConf
42	6	1	Total 1	K 1	0

- Molecule 43 is D-chiro inositol hexakisphosphate (three-letter code: KGN) (formula: $C_6H_{18}O_{24}P_6$).



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

- Molecule 44 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



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

- Molecule 45 is ZINC ION (three-letter code: ZN) (formula: Zn).

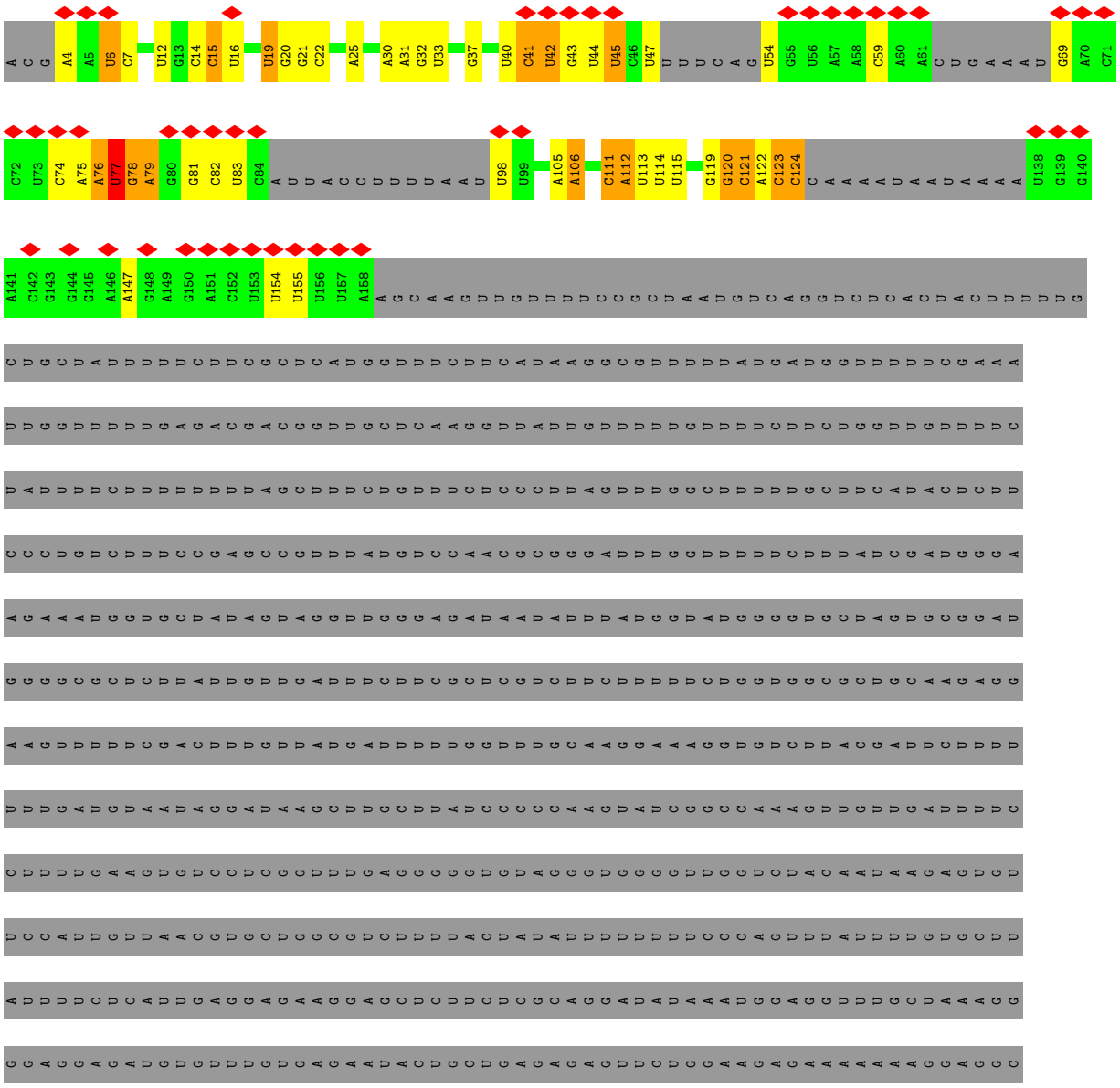
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
45	D	1	1	1	0
45	L	3	3	3	0
45	M	1	1	1	0
45	N	2	2	2	0
45	c	1	1	1	0

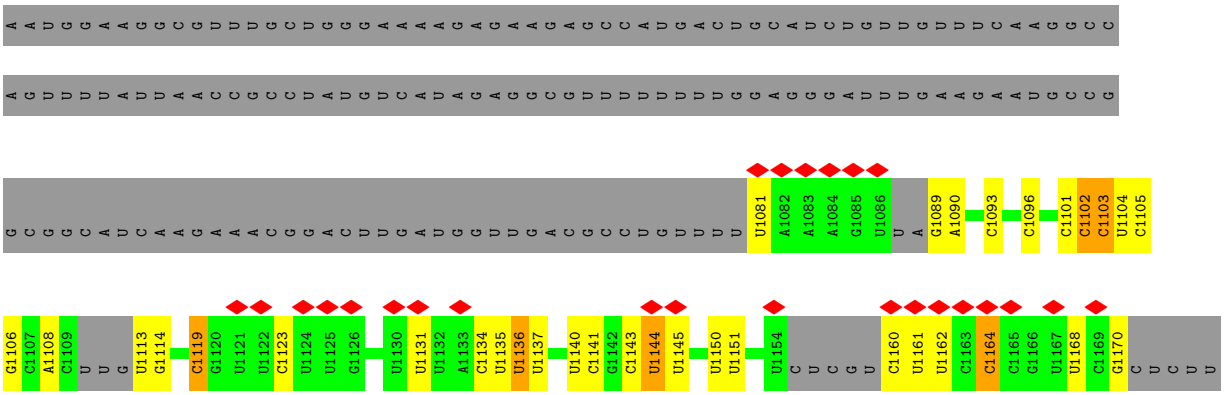
3 Residue-property plots (i)

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: U2 snRNA

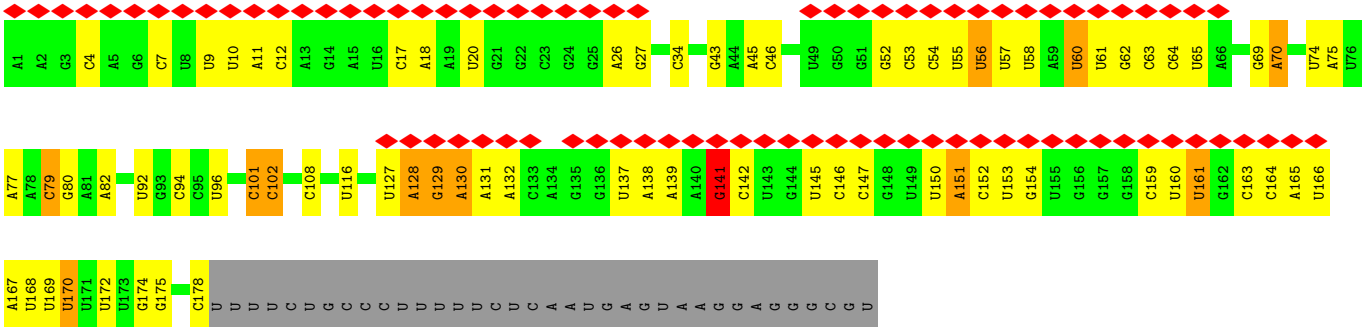
Chain 2:





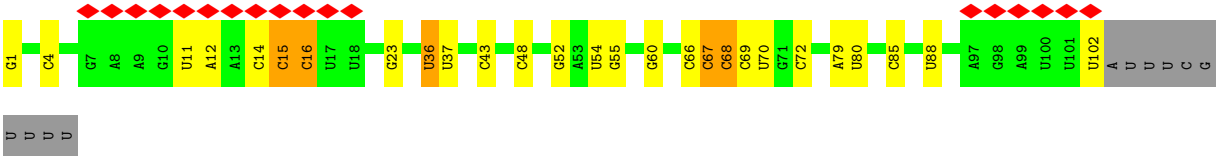
- Molecule 2: U5 snRNA

Chain 5:



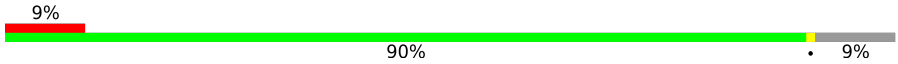
- Molecule 3: U6 snRNA

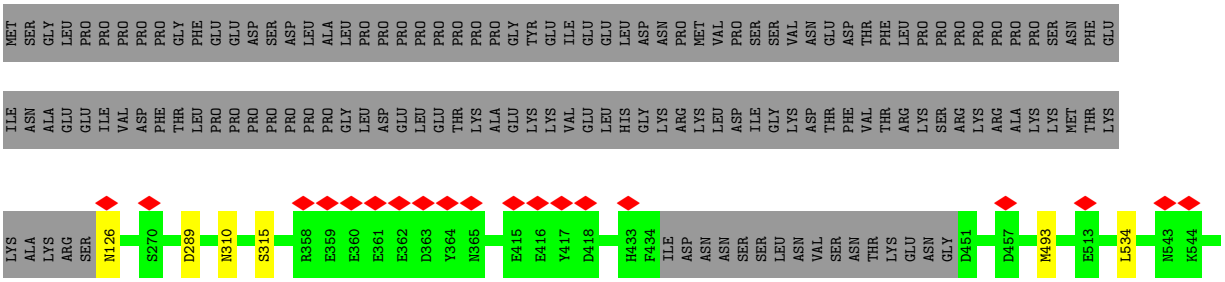
Chain 6:



- Molecule 4: Pre-mRNA-splicing factor 8

Chain A:





R709 E772 S773 I774 R775 Q776 K777 R780 L903 D908 R928 E933 E941 D945 L1035 S1036 E1037 N1140 D1166 D1192 F1195 E1196 E1276 W1286 R1292 L1302 R1347 L1412 E1432 D1433 E1467 A1468 I1469 Q1470 Q1471 N1472 R1473 R1474 L1475

A1476 F1477 E1478 E1481 F1495 Q1496 R1497 D1498 R1499 E1576 K1577 A1578 S1579 T1591 H1592 A1593 Q1594 R1595 T1596 G1597 L1598 H1707 D1722 K1725 D1785 A1786 V1830 Q1831 E1832 L1835 N1836 S1837 S1838 N1839 Y1840 A1841 E1842 N1845 N1846 D1847 I1848 D1853 D1854 T1855 N1856 R1859 V1860

T1861 V1862 H1863 K1864 T1865 F1866 E1867 G1868 N1869 V1870 A1871 T1872 K1873 A1874 I1875 N1876 G1877 F1880 P1884 K1885 H1888 L1889 K1892 I1893 I1894 H1895 T1896 S1897 V1898 W1899 A1900 G1901 Q1902 K1903 R1904 L1905 S1906 Q1907 L1908 E1915 R1922 S1923 L1924 P1925 K1926 E1927 E1928 Q1929 P1930 K1931 R1937 D1942

D1950 N1953 R1957 P1958 T1959 E1960 L1961 R1962 M1969 M1979 K1980 A1981 T1982 E1983 P1984 Q1985 M1986 V1987 L1988 Y1992 L1996 S2000 S2001 Y2002 F2005 D2057 Y2063 G2064 R2065 K2066 Y2067 N2068 V2069 N2070 I2071 S2072 A2073 L2074 T2075 Q2076 T2077 E2078 I2079 K2080 D2081 I2082 I2083 L2084 G2085

Q2086 N2087 ILE LYS ALA PRO SER VAL LYS ARG GLN LYS MET ALA GLU LEU GLU ALA ALA ARG SER GLU LYS GLN ASN ASP GLU GLU ALA ALA GLY ALA SER THR VAL MET LYS THR LYS THR ILE ASN ALA GLN GLY GLU GLU ILE VAL VAL VAL ALA SER ALA ASP TYR GLU SER GLN THR

PHE SER SER LYS ASN E2151 A2173 D2174 D2175 E2179 L2189 L2190 K2191 K2192 F2193 V2203 K2213 D2214 K2222 T2223 V2224 V2233 D2243 I2244 G2245 E2251 G2252 E2264 E2265 L2266 D2287 K2319 D2320 I2321 M2322 N2323 V2324 L2325 S2326 T2332 F2333 S2334 L2339 L2340 D2343

R2344 E2369 G2370 D2371 Y2372 G2377 I2378 P2379 L2380 E2381 Q2394 F2395 SER GLU LEU ALA GLY ASP GLU GLU LEU GLU ALA GLU GLN ILE ASP VAL PHE SER

- Molecule 5: Pre-mRNA-splicing helicase BRR2

Chain B: 25% 77% • 21%

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

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

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

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

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

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

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

PRO ALA ILE PRO PRO VAL ILE ASP LEU GLU LYS ILE LYS PHE ASP GLU SER SER LYS LEU MET T442 V443 T444 K445 E450 G451 S452 F453 K454 R455 V456 K457 P458 Q459 Y460 D461 E462 I463 H464 I465 P466 A467 P468 S469 K470 P471 V472 I473 D474 Y475 K478 E479 I480 T481 S482 L483 P484

E489 A490 F491 P492 S493 S494 E495 T496 T497 S498 L499 N500 P501 F507 E512 M517 L518 I519 C520 A521 P522 T523 G524 S525 G526 K527 T528 N529 I530 A531 L532 L533 T534 V535 L536 K537 A538 L539 S540 H541 L550 N551 L552 S553 A554 I560 A561 P562 L563 K564 A565 L566 R572

Q575 R576 R577 G582 A586 E587 L588 T589 G590 D591 S592 R593 L594 S595 R596 K597 Q598 I599 D600 L605 V606 P609 D613 T616 R617 N618 S619 C620 N621 L622 V628 R629 L630 L631 I632 I633 D634 E635 I636 D641 D642 E648 R653 A657 S658 K659 Y660 G661

Q662 E663 I667 I668 G669 L670 S671 A672 V680 G681 R682 F683 L684 R685 K688 E689 G690 L691 S696 S697 F698 P702 L703 S704 Q705 Q706 F707 C708 G709 I710 K711 E712 R713 N714 S715 L716 A725 C726 E736 G737 F743 A754 T755 W756 A762 E763 E764 T767 H768

K769 L770 T771 K772 N773 D774 A775 G776 N787 V788 L789 D790 L793 R794 K795 L796 I797 E798 S799 G800 I801 G802 T803 G807 L808 T809 R810 S816 E817 A821 D822 G823 L824 L825 Q826 T833 L834 A835 W836 G837 L840 V845 I846 I847 Y853 S854 P855 E856 E861

L872 E885 G886 I887 I888 I889 S910 Q911 F912 V913 S914 D918 N919 L920 N921 R947 M953 K956 V957 P958 D959 I960 S961 S962 Q965 L966 K967 K968 L973 V974 H975 S976 A977 L978 C979 E983 D990 E997 A998 F1008 R1020 E1021 L1022 D1023 R1033 I1034

F1035 E1039 E1040 V1044 S1045 V1046 R1047 R1052 Q1056 L1057 L1058 E1059 K1060 A1061 I1065 R1066 E1067 D1068 I1069 D1070 D1071 P1072 Y1083 L1087 K1088 F1089 E1090 G1091 H1102 A1105 G1106 R1107 K1118 K1138 K1150 V1154 R1159 L1160 T1164 D1169 E1174 R1181

R1184 K1187 Y1188 G1189 M1202 N1207 A1224 D1229 M1230 N1231 V1232 D1245 T1246 D1247 G1248 D1249 S1250 Y1253 D1262 G1265 Q1281 I1293 S1294 E1295 H1299 S1300 E1301 S1307 F1308 N1309 G1310 F1311 K1312 K1316 F1317 P1318 P1319 P1320 L1323 L1324 E1325 N1326 I1327 S1328

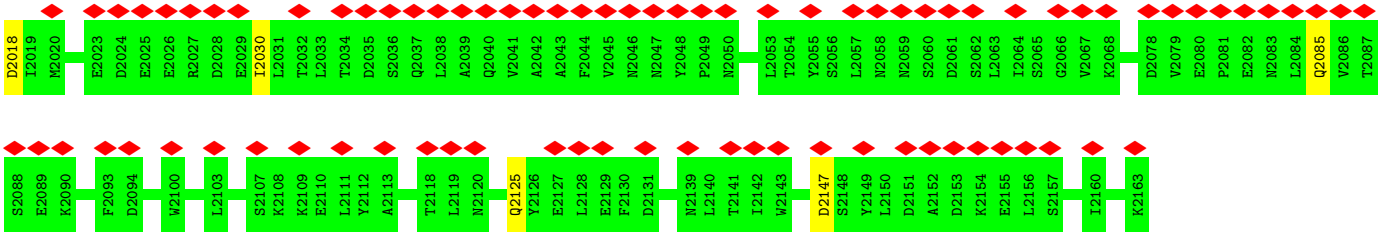
I1329 G1335 N1336 D1337 D1338 F1339 S1340 E1341 K1350 Q1354 V1355 F1356 N1361 S1362 N1363 S1370 G1371 K1372 A1380 A1383 R1389 A1395 V1396 Y1397 I1405 D1406 F1407 W1412 L1428 G1429 N1430 D1431 P1432 S1433 L1434 L1455 M1471 I1472 Y1473 D1474 D1475 A1476 Y1488 E1503

K1504 K1505 I1506 A1518 R1519 T1528 K1529 P1537 S1538 E1539 R1540 I1541 E1542 P1543 L1544 E1545 I1546 K1552 D1553 F1559 N1560 F1561 A1567 A1570 S1571 A1572 A1573 A1574 N1577 R1578 F1584 L1585 P1586 S1587 R1588 M1592 E1593 F1601 A1604 I1605 E1606 W1607 D1608 N1611 V1612 E1613 E1614

E1615 K1623 L1624 T1625 D1626 R1630 L1633 I1639 G1643 R1654 Y1658 G1659 A1660 S1667 K1668 D1669 A1672 F1673 D1678 G1684 D1689 G1690 A1691 E1692 H1693 N1701 E1702 L1703 M1706 A1710 S1711 G1712 N1713 D1714 S1715 M1716 A1717 G1718 K1719 N1727 M1728 K1729 K1733

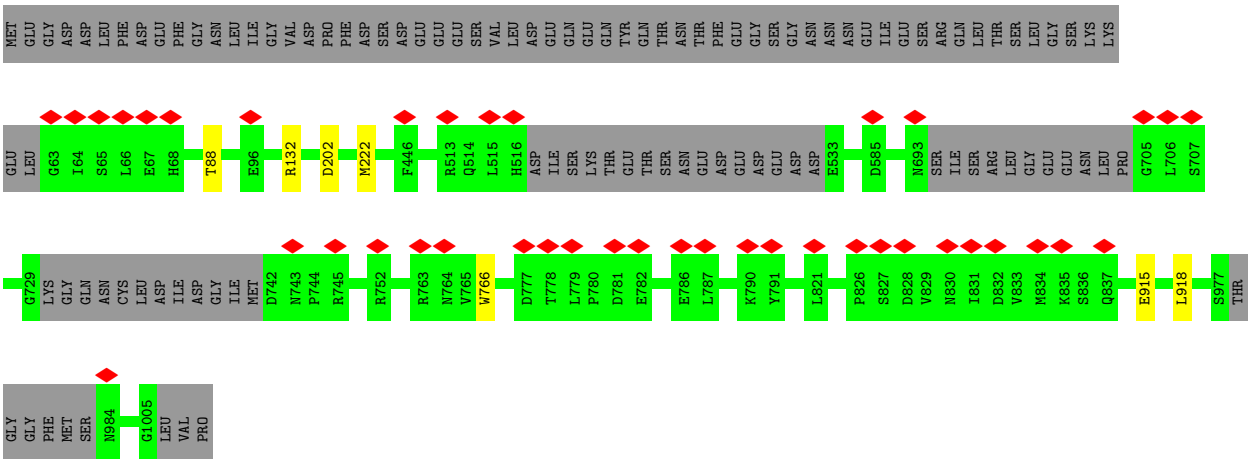
I1737 E1738 P1739 L1740 Y1745 A1759 N1760 I1763 S1787 Y1788 Y1789 G1790 V1791 R1792 D1793 T1794 F1819 D1824 T1825 GLU ALA GLU VAL THR ALA GLU VAL ASN GLY GLY ASP ASP GLU ALA T1841 T1846 N1849 A1853 G1857 V1858 F1861 F1866 S1871 S1874 T1875

L1876 L1883 A1886 V1887 E1888 F1889 E1890 S1891 R1899 R1908 R1912 F1913 P1914 E1915 H1916 T1917 S1918 S1919 G1920 A1932 D1942 F1943 Q1944 N1945 D1946 L1947 K1948 D1949 I1950 L1951 E1952 K1953 I1964 L1965 S1966 A1967 N1968 G1969 Y1970 I1984 V1990 H1999 K2009 E2010 I2011 N2012 V2013



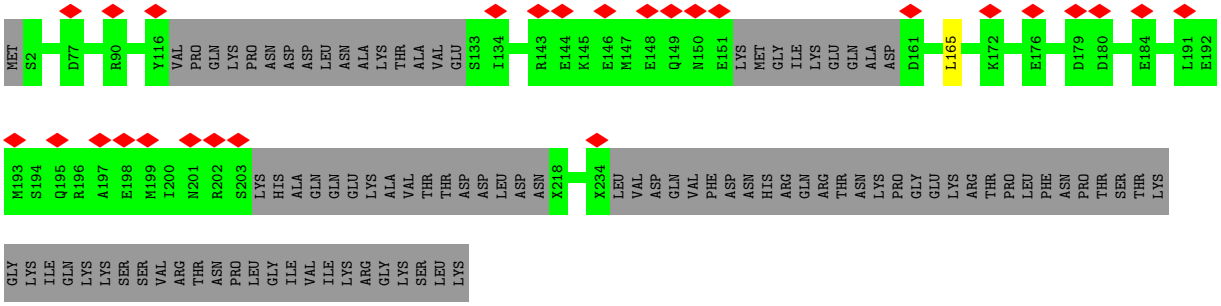
- Molecule 6: Pre-mRNA-splicing factor SNU114

Chain C:



- Molecule 7: Splicing factor YJU2

Chain D:

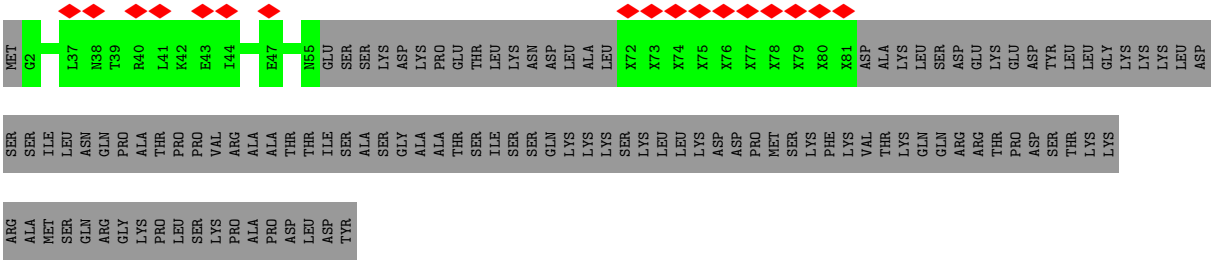


- Molecule 8: 5' exon of UBC4 mRNA

Chain E:

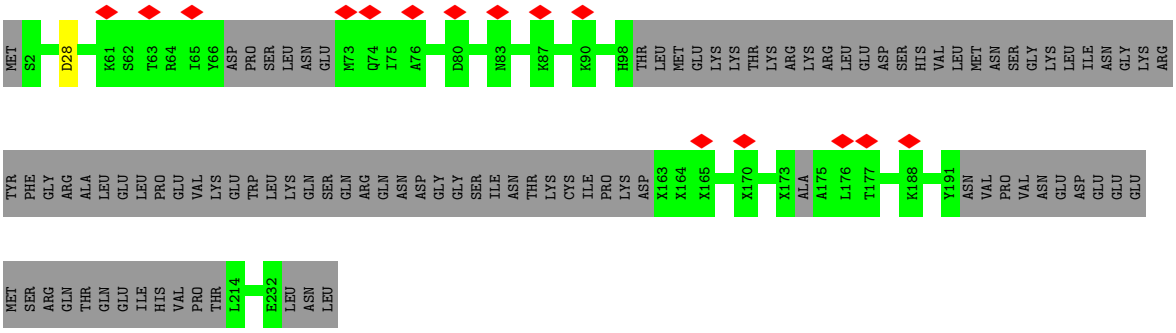
- Molecule 9: Pre-mRNA-splicing factor CWC25

Chain F:



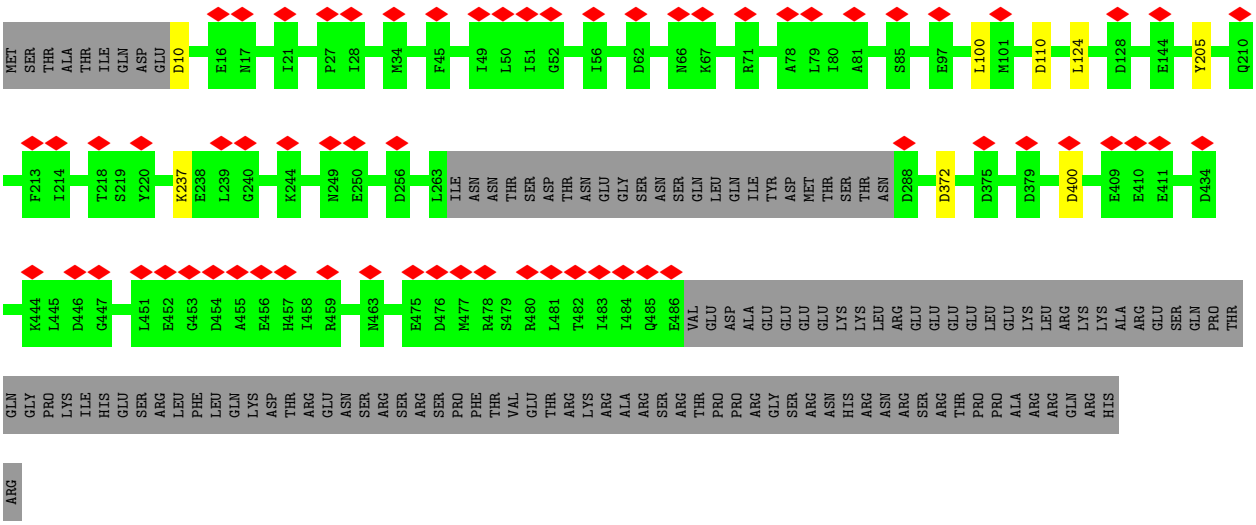
- Molecule 10: Pre-mRNA-splicing factor ISY1

Chain G:



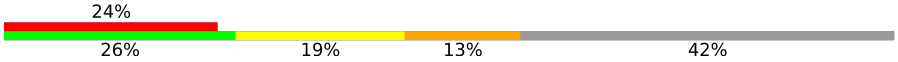
- Molecule 11: CWC22 isoform 1

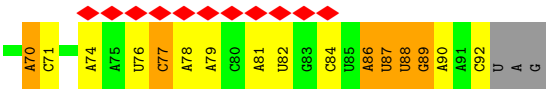
Chain H:



- Molecule 12: Branched intron and 3' exon of UBC4 pre-mRNA

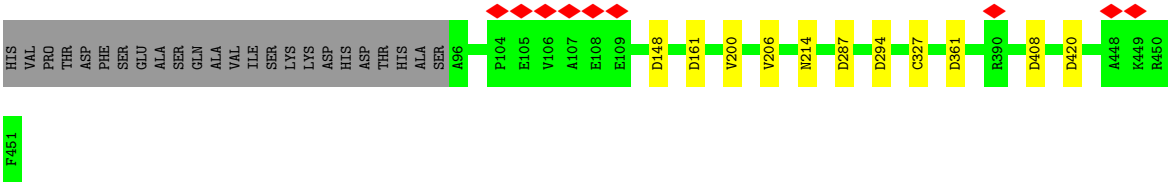
Chain I:





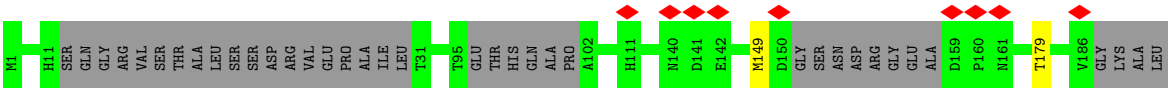
• Molecule 13: BJ4_G0054360.mRNA.1.CDS.1

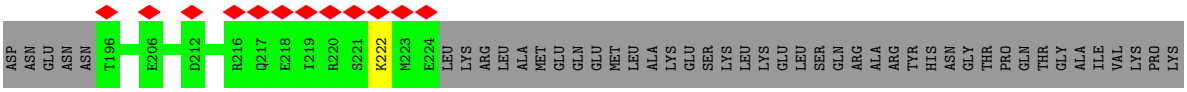
Chain J:



• Molecule 14: Pre-mRNA-processing protein 45

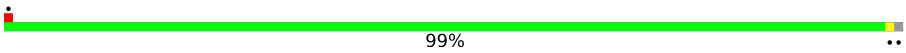
Chain K:

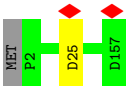




• Molecule 15: BUD31 isoform 1

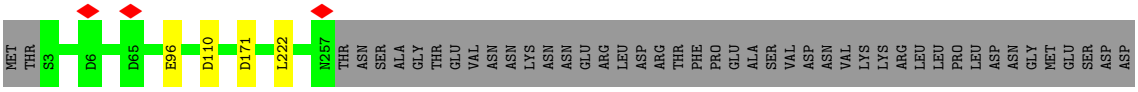
Chain L:





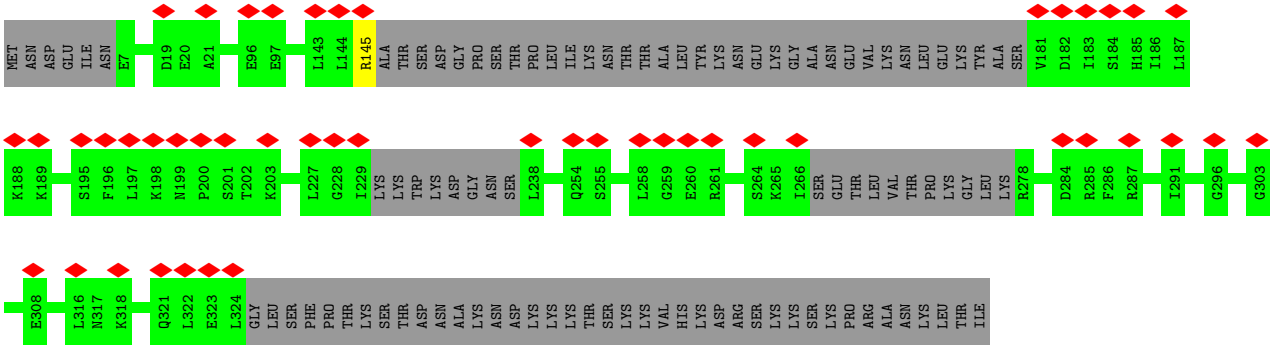
• Molecule 16: Pre-mRNA-splicing factor CWC2

Chain M:



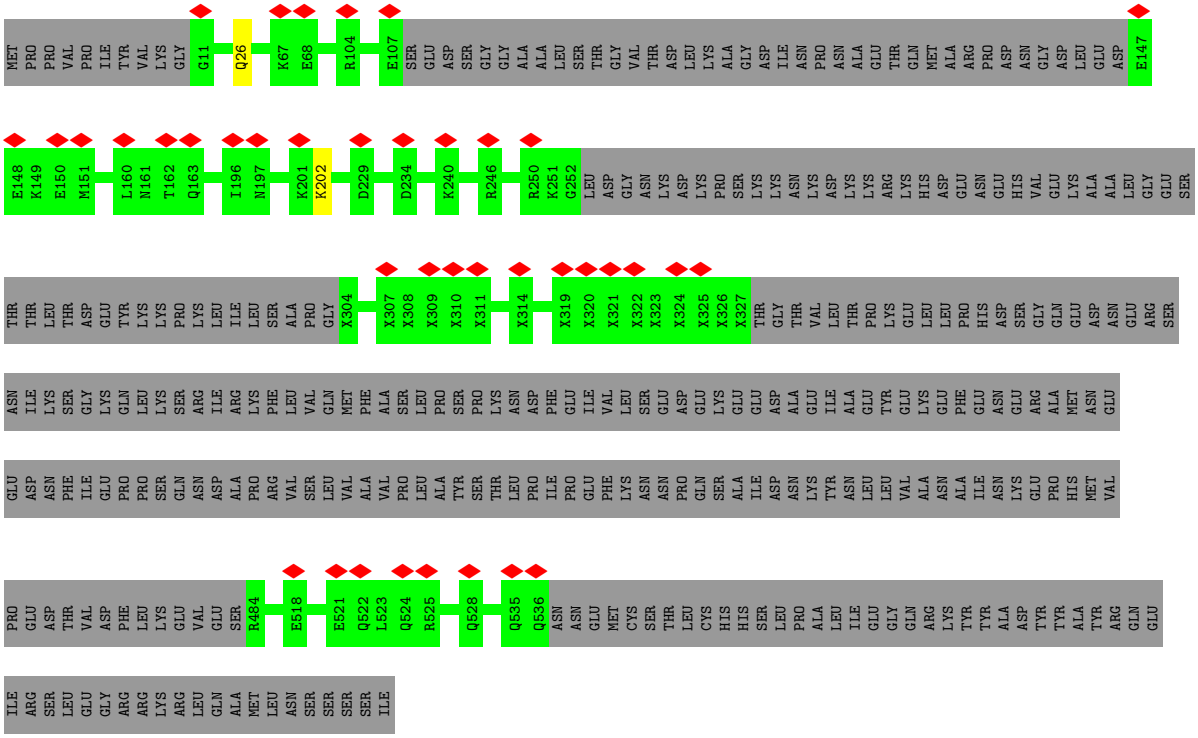
- Molecule 17: Pre-mRNA-splicing factor SLT11

Chain N:



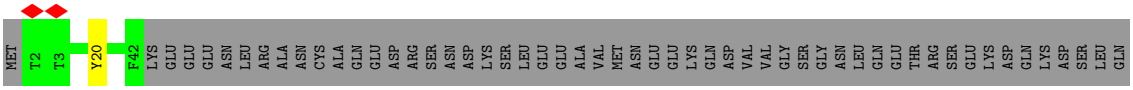
- Molecule 18: Y55_G0042700.mRNA.1.CDS.1

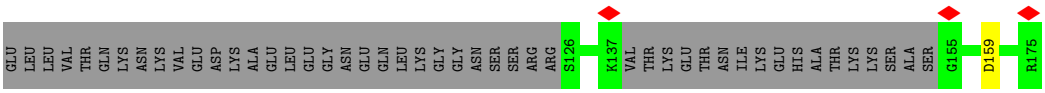
Chain O:



- Molecule 19: Pre-mRNA-splicing factor CWC15

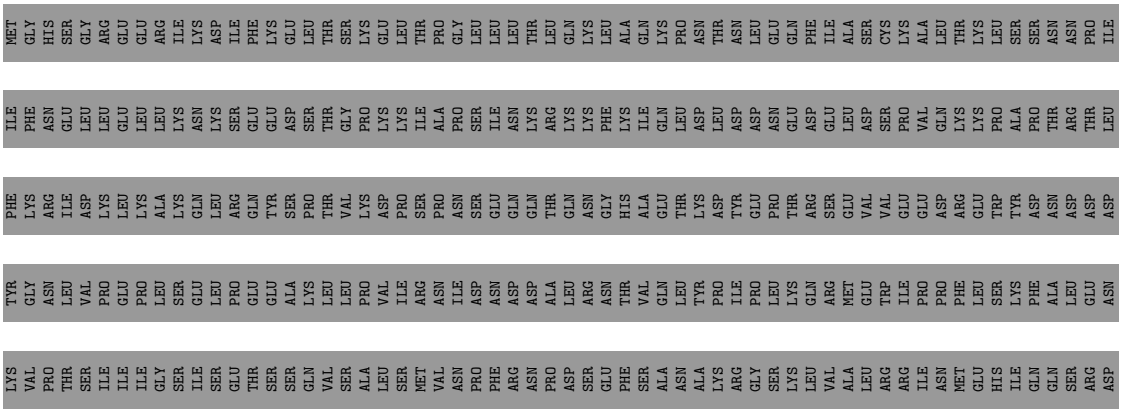
Chain P:

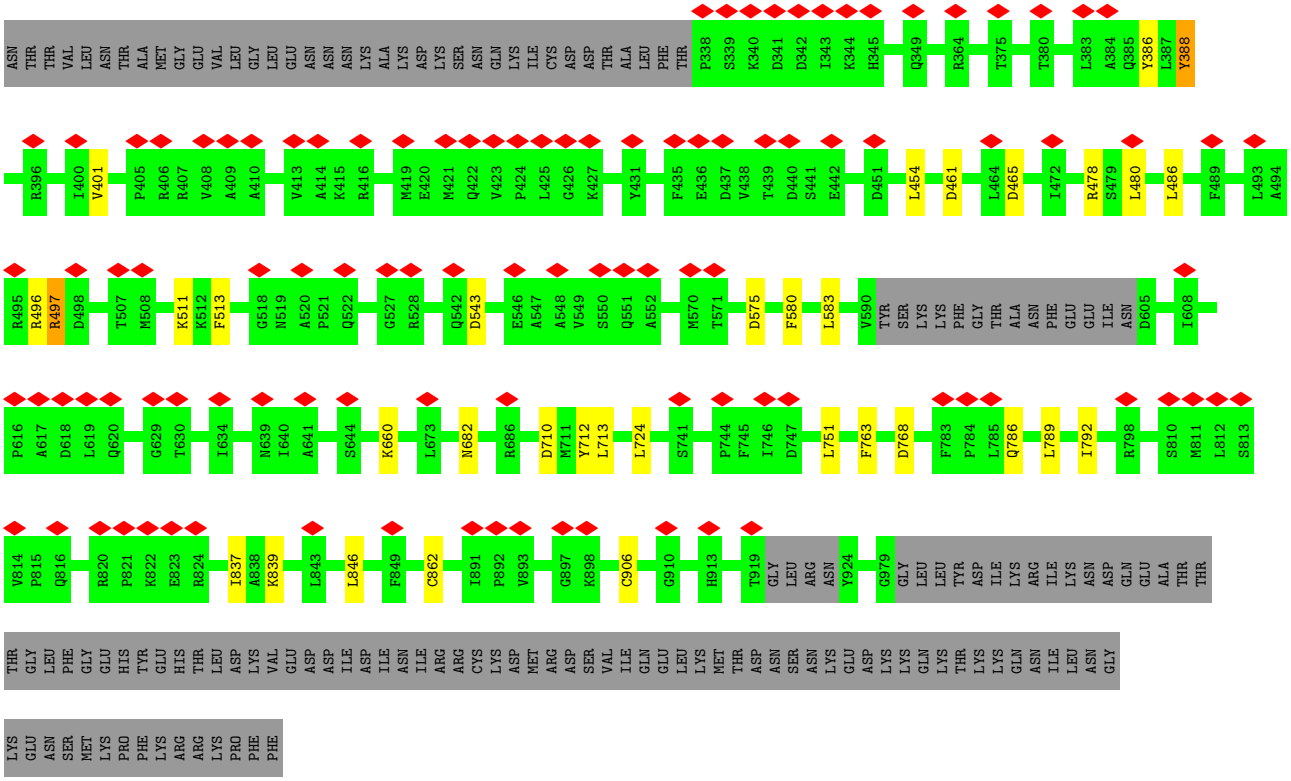




• Molecule 20: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP16

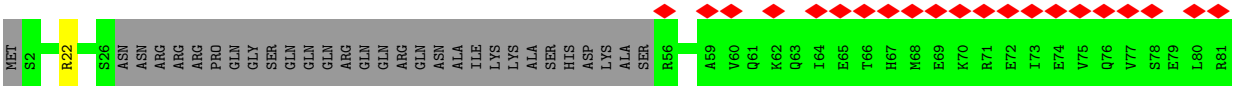
Chain Q:

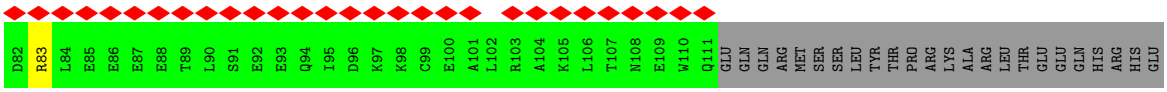




• Molecule 21: Pre-mRNA-splicing factor CWC21

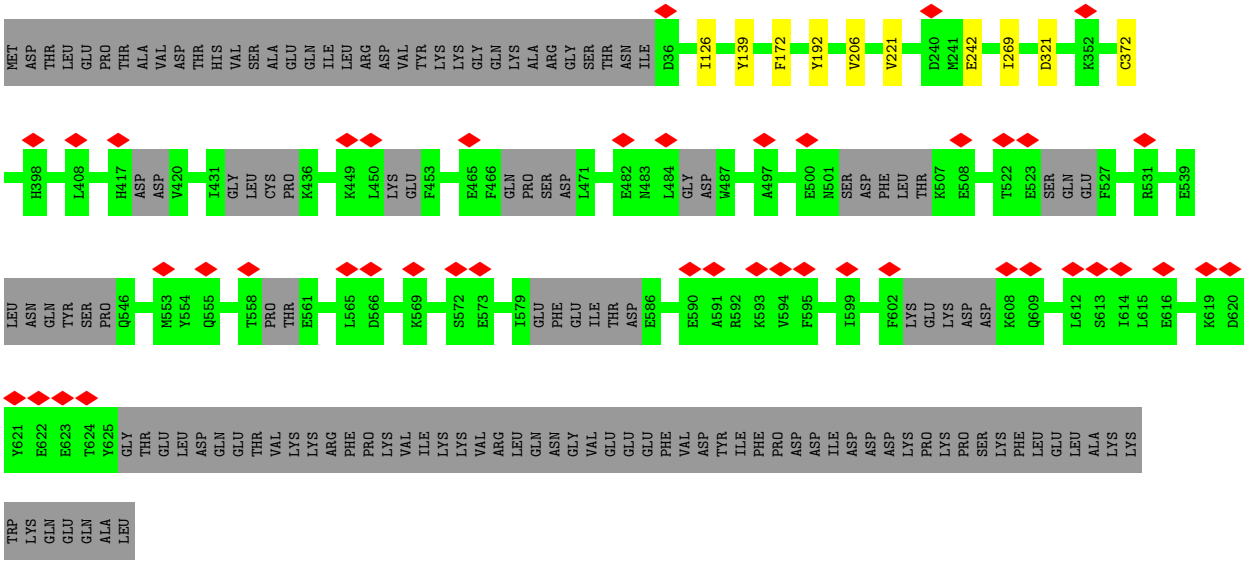
Chain R:





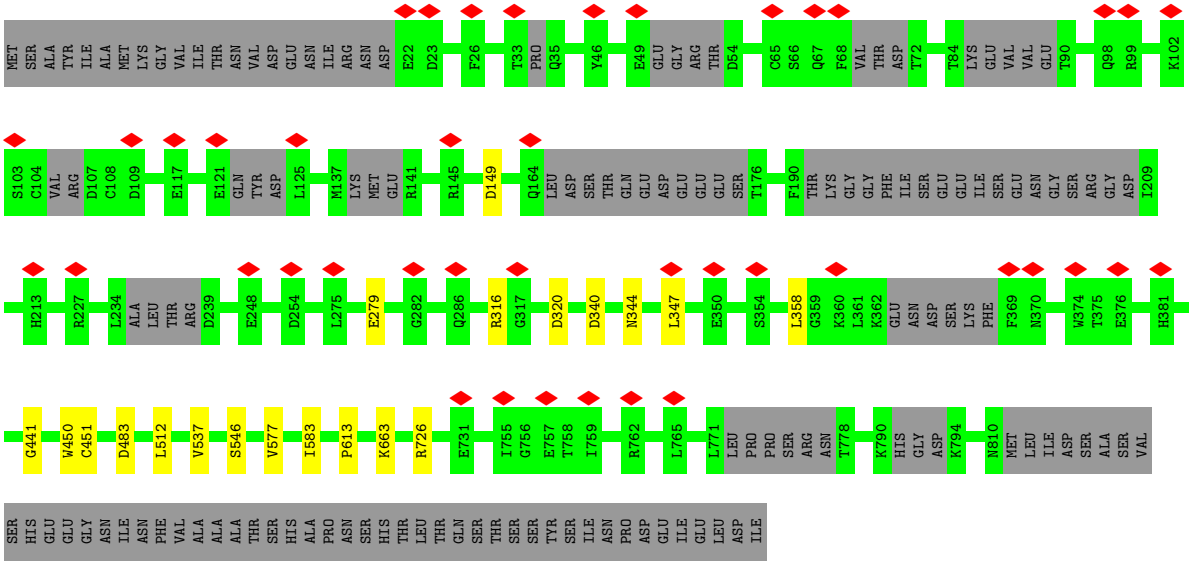
• Molecule 22: CLF1 isoform 1

Chain S:



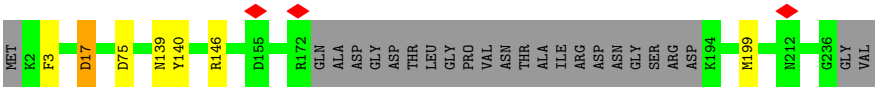
• Molecule 23: SYF1 isoform 1

Chain T:



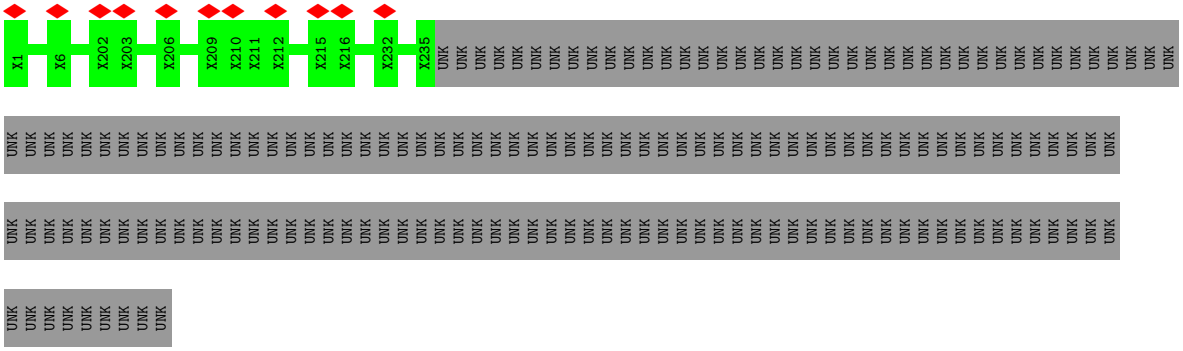
• Molecule 24: HLJ1_G0053790.mRNA.1.CDS.1

Chain W:



• Molecule 25: Unassigned structure

Chain X:

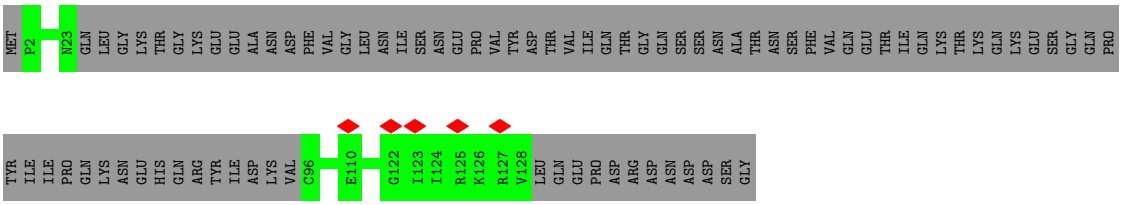


• Molecule 26: BJ4_G0027490.mRNA.1.CDS.1

Chain Y:

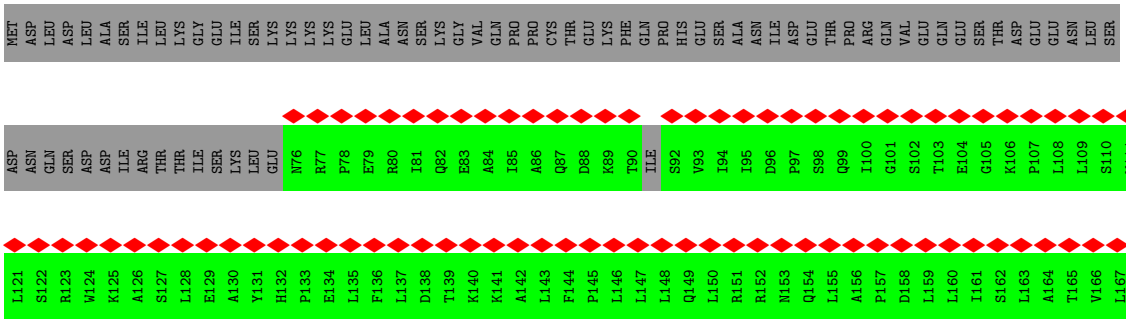
• Molecule 27: NTC20 isoform 1

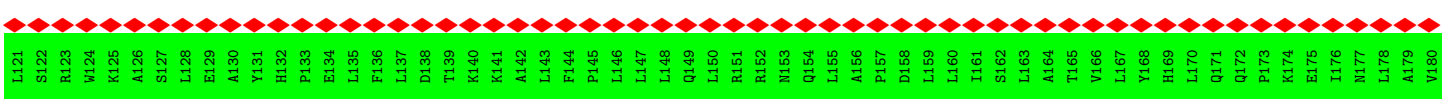
Chain Z:

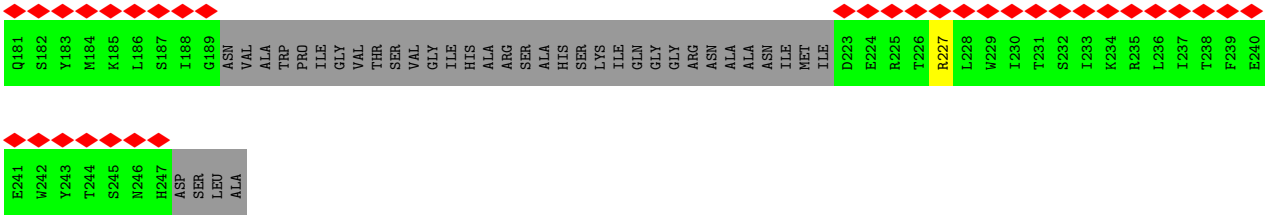


• Molecule 28: Pre-mRNA-splicing factor 18

Chain a:

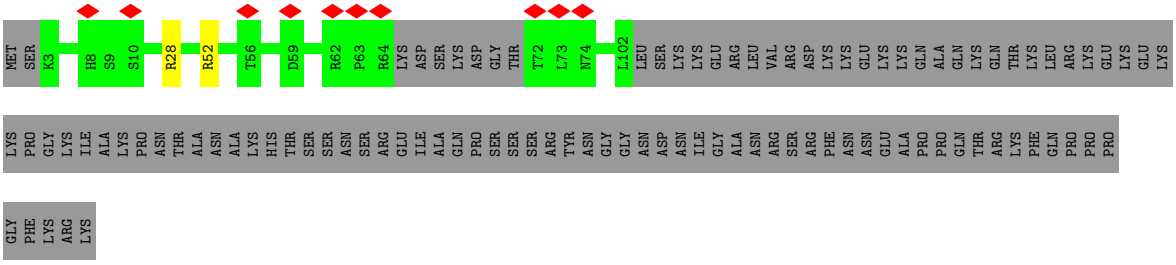






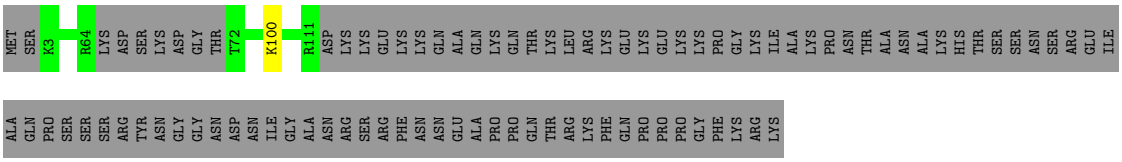
• Molecule 29: Small nuclear ribonucleoprotein-associated protein B

Chain b:



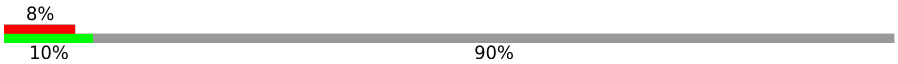
• Molecule 29: Small nuclear ribonucleoprotein-associated protein B

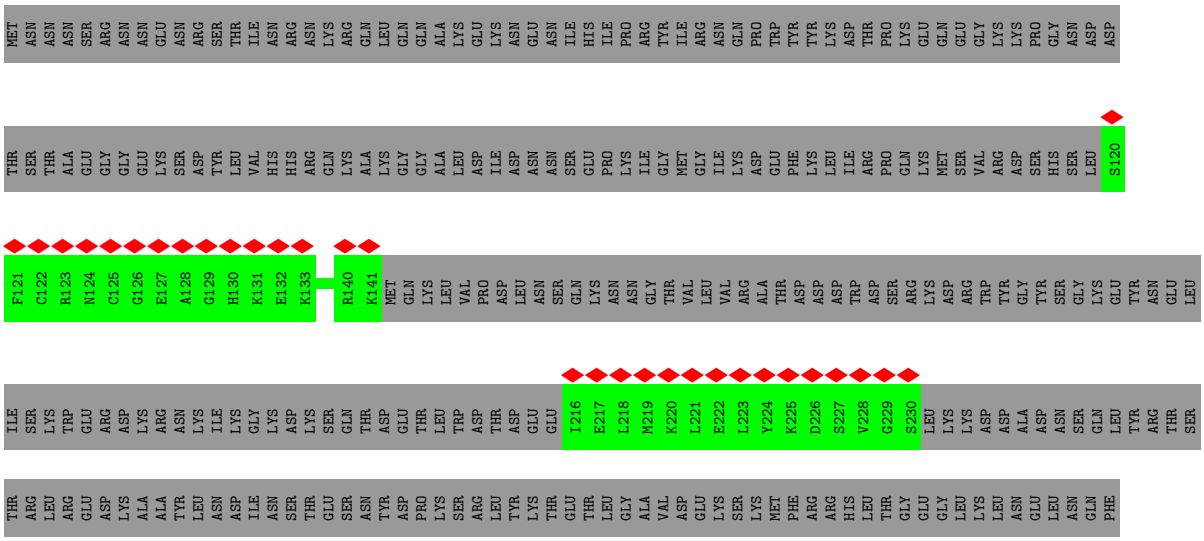
Chain k:

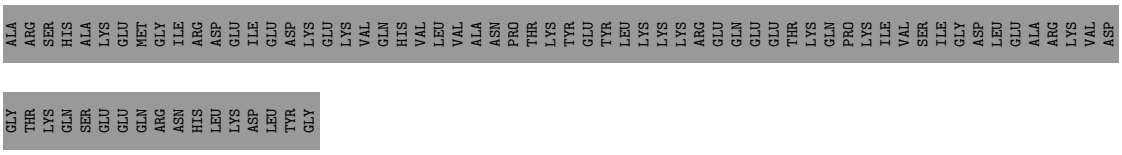


• Molecule 30: Pre-mRNA-splicing factor SLU7

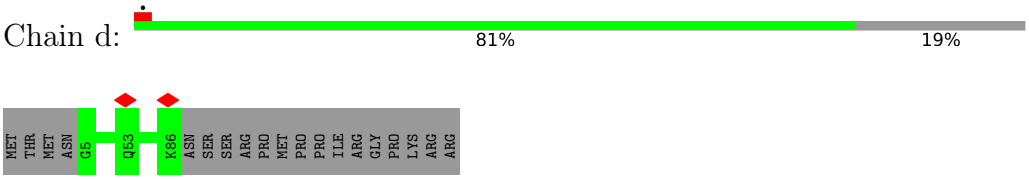
Chain c:



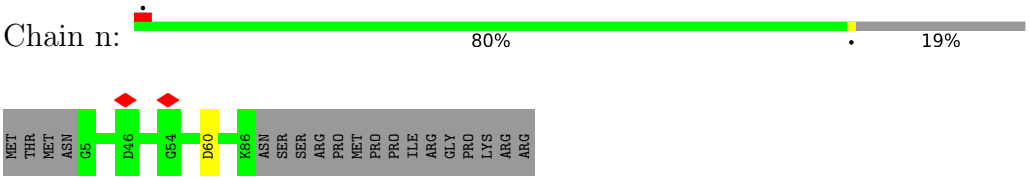




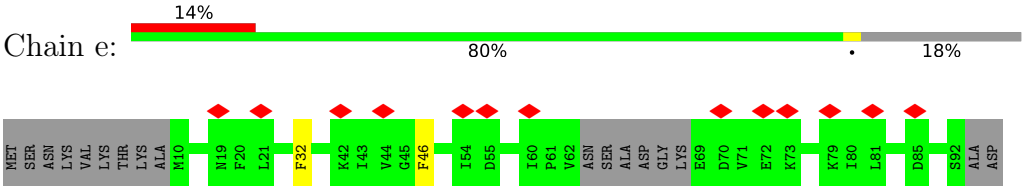
• Molecule 31: Small nuclear ribonucleoprotein Sm D3



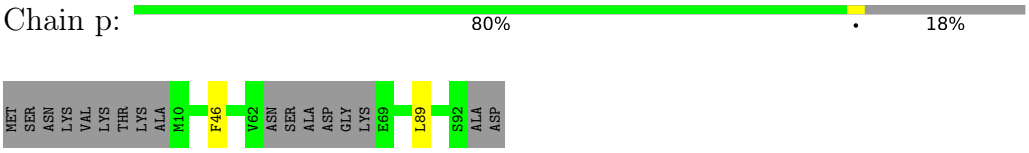
• Molecule 31: Small nuclear ribonucleoprotein Sm D3



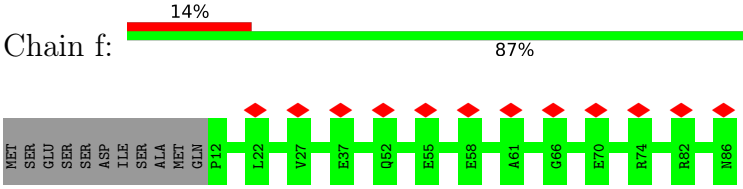
• Molecule 32: Small nuclear ribonucleoprotein E



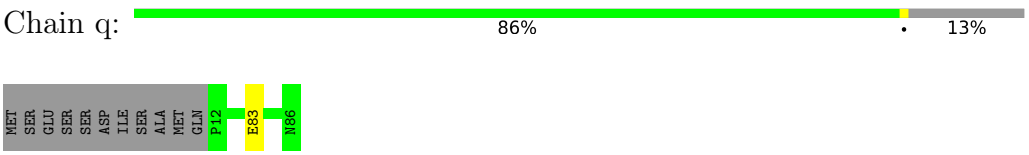
• Molecule 32: Small nuclear ribonucleoprotein E



• Molecule 33: Small nuclear ribonucleoprotein F

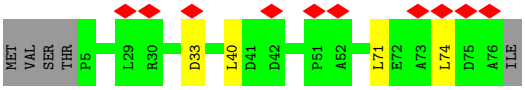


• Molecule 33: Small nuclear ribonucleoprotein F



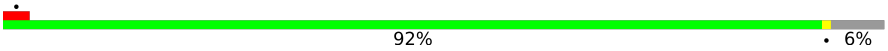
• Molecule 34: Small nuclear ribonucleoprotein G

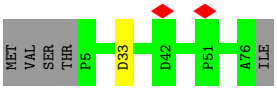
Chain g:



• Molecule 34: Small nuclear ribonucleoprotein G

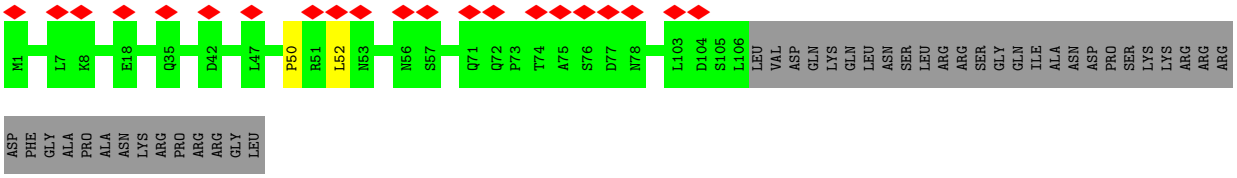
Chain r:





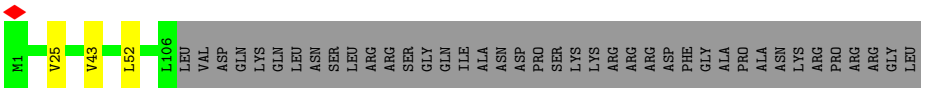
• Molecule 35: Small nuclear ribonucleoprotein Sm D1

Chain h:



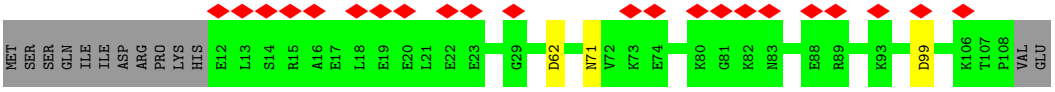
• Molecule 35: Small nuclear ribonucleoprotein Sm D1

Chain l:



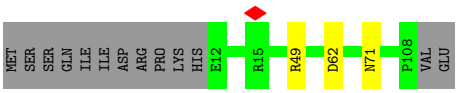
• Molecule 36: Small nuclear ribonucleoprotein Sm D2

Chain j:



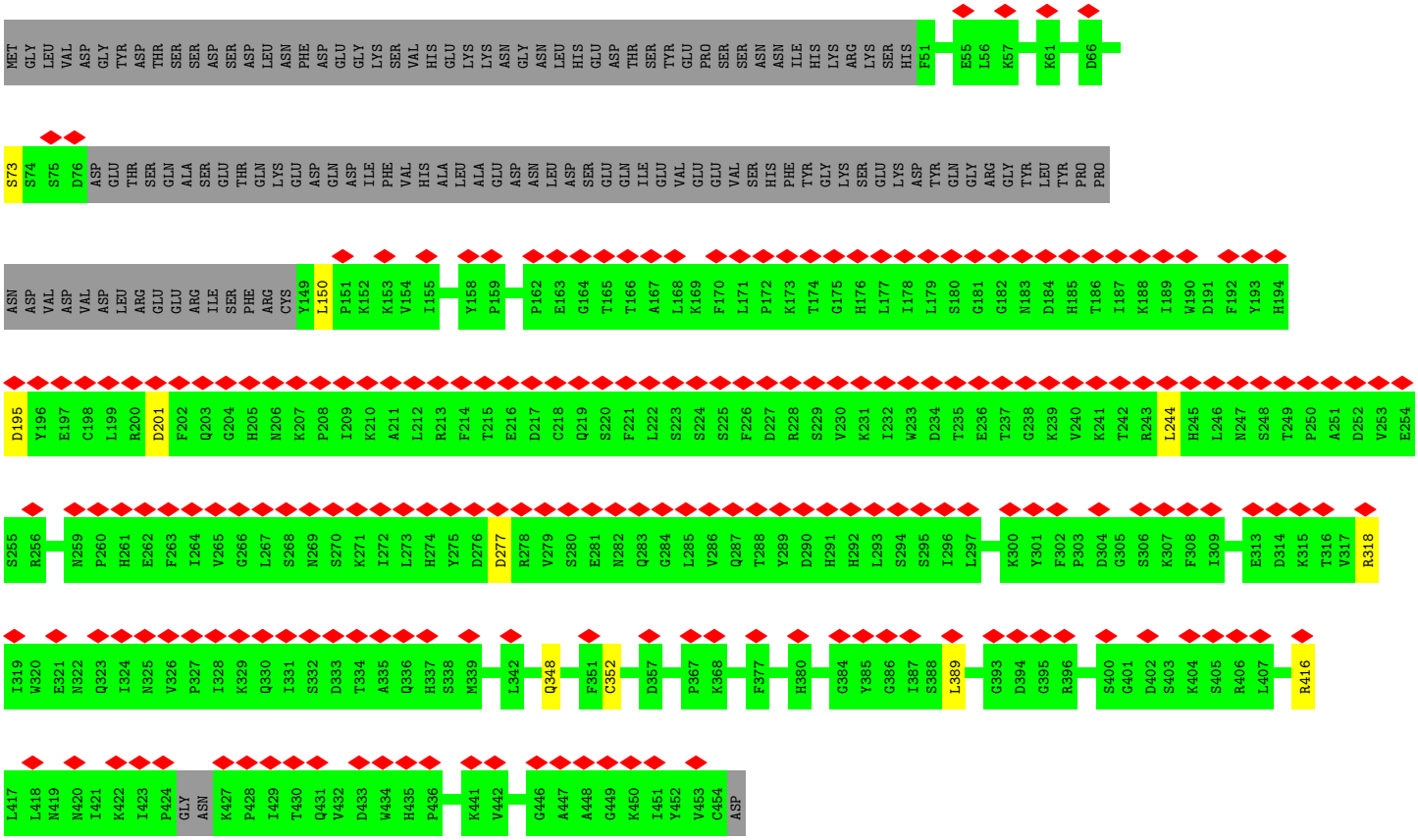
• Molecule 36: Small nuclear ribonucleoprotein Sm D2

Chain m:



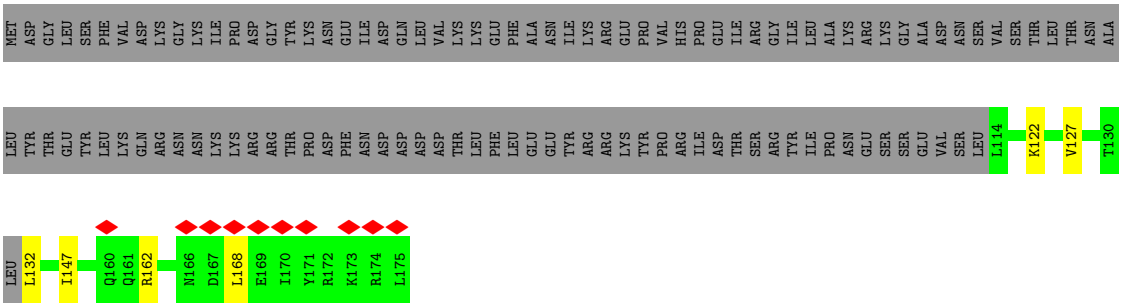
• Molecule 37: CDC40 isoform 1

Chain o:



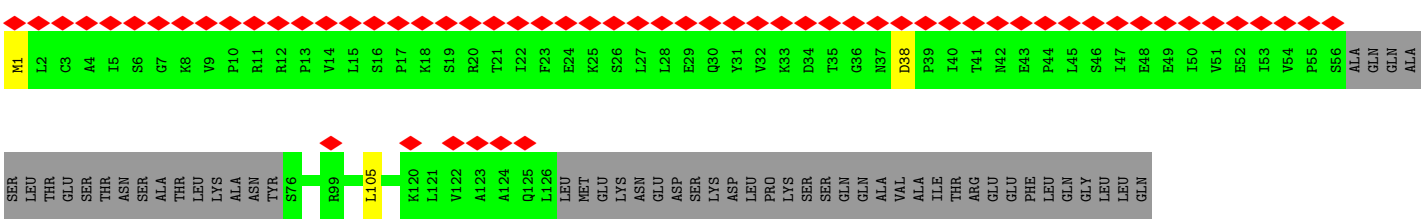
- Molecule 38: SNT309 isoform 1

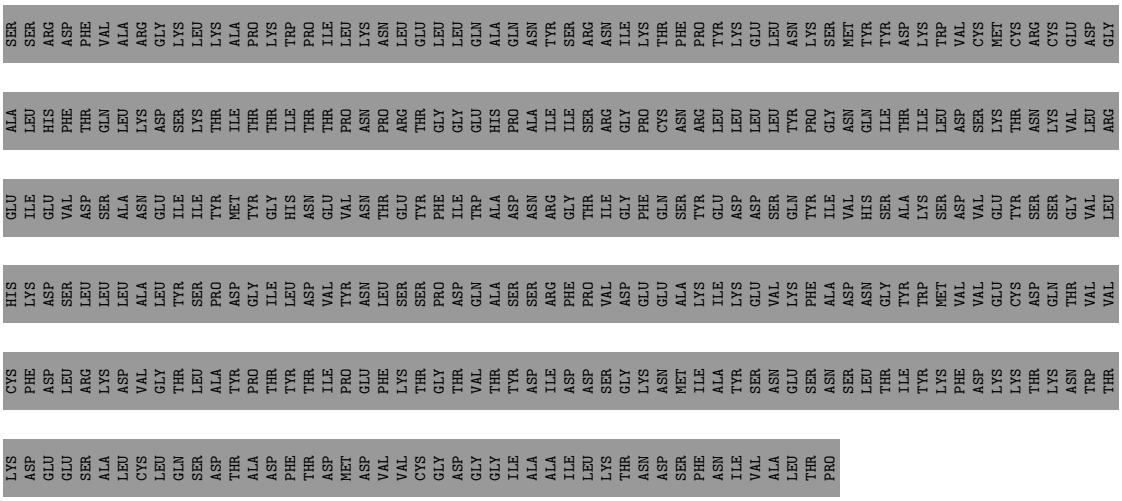
Chain s:



- Molecule 39: Pre-mRNA-processing factor 19

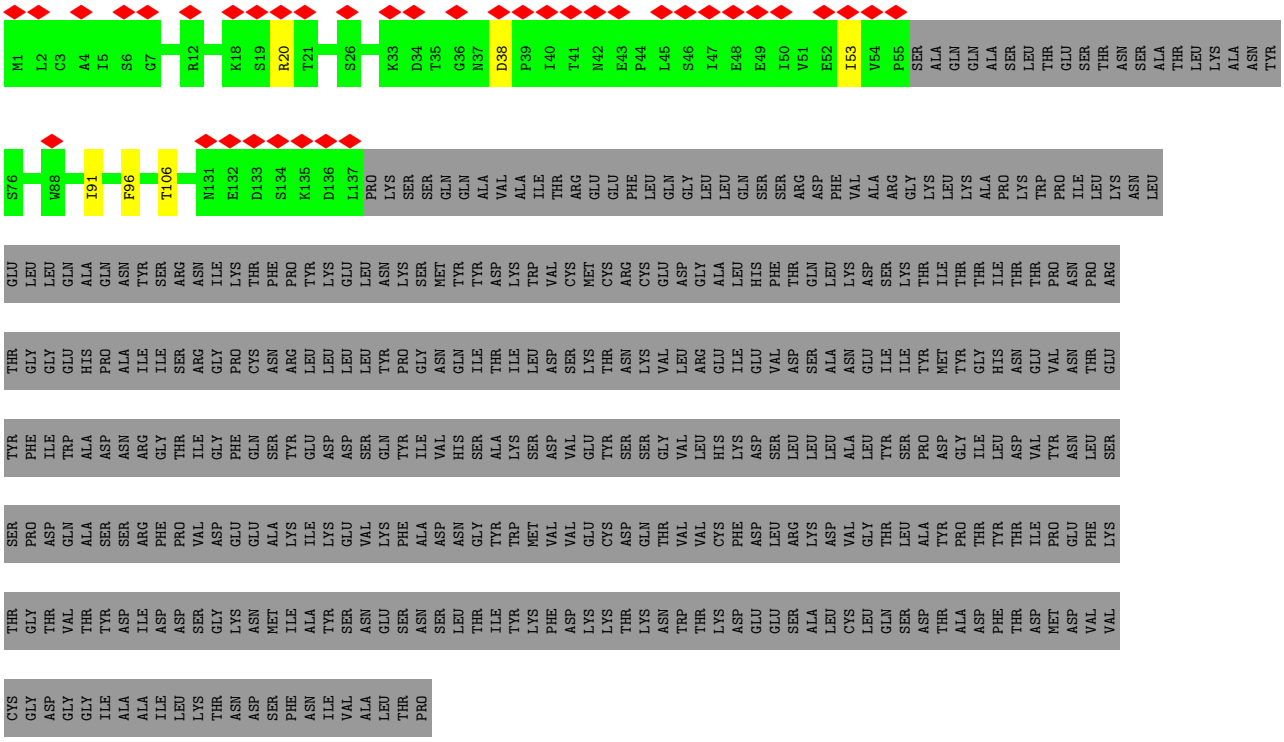
Chain t:





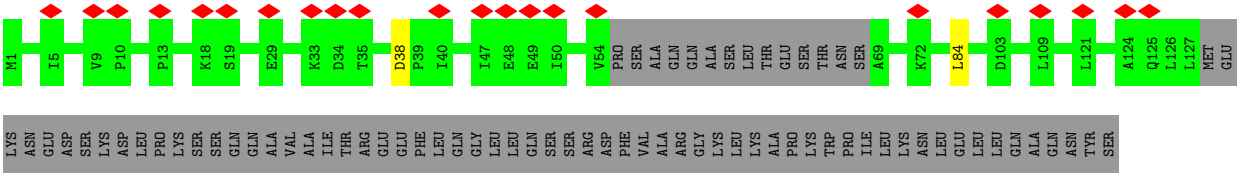
• Molecule 39: Pre-mRNA-processing factor 19

Chain u:



• Molecule 39: Pre-mRNA-processing factor 19

Chain v:



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

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

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

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

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

LEU LYS THR ASN ASP SER PHE ASN ILE VAL ALA LEU THR PRO

- Molecule 39: Pre-mRNA-processing factor 19

Chain w: 50% 84% • 13%

M1 K8 V9 P10 R11 R12 P13 S16 P17 I22 S26 L27 V32 K33 D34 T35 G36 P39 I40 T41 N42 E43 P44 L45 S46 I47 E48 E49 P55 SER ALA GLN GLN ALA SER LEU THR GLU SER THR ASN SER ALA THR LEU LYS ALA ASN TYR S76 N92

A124 Q125 L126 L127 K130 N131 E132 K135 D136 L137 PRO LYS SER SER GLN GLN ALA VAL ALA ILE THR ARG GLU GLU PHE LEU GLN GLY LEU LEU GLN SER SER ARG ASP PHE VAL ALA ARG GLY LYS LEU LYS A171 P172 K173 W174 P175 I176 L177 K178 N179 L180 E181 L182 L183 Q184 A185

Q186 N187 Y188 S189 R190 N191 E199 L200 N201 K202 S203 M204 Y205 Y206 V210 C211 M212 C213 R214 C215 E216 D217 G218 A219 F222 L225 K226 D227 S228 K229 R240 T241 G242 G243 E244 H245 P246 A247 I248 I249 P253 C254 N255 R256 L257 L258 L259 L260 Y261 P262 D270 S271 K272

T273 N274 K275 V276 L277 R278 D283 S284 A285 N286 E287 I288 I289 H294 N295 GLU VAL ASN T299 E300 Y301 A305 D306 N307 R308 G309 T310 I311 G312 F313 Q314 S315 Y316 E317 D318 D319 S320 Q321 Y322 I323 V324 H325 S326 A327 K328 S329 D330 V331 E332 Y333 S334 S335 G336 H339 K340 D341

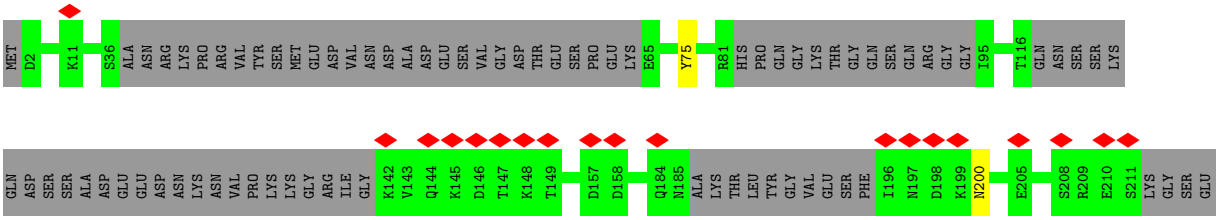
S342 L343 L344 L345 A346 L347 Y348 S349 P350 D351 G352 I353 L354 D355 V356 Y357 N358 L359 S360 S361 P362 D363 Q364 A365 S366 S367 R368 F369 P370 E373 E374 A375 K376 I377 K378 E379 F382 N385 G386 Y387 W388 M389 V397 V398 C399 F400 D401 L402 R403 K404 D405 V406 G407 T408 L409 A410

Y411 P412 T413 Y414 T415 I416 PRO GLU PHE LYS THR G422 T423 V424 T425 D429 D430 S431 G432 K433 A437 Y438 S439 N440 E441 S442 N443 S444 L445 T446 I447 D451 K452 K453 T454 K455 K459 D460 E461 E462 S463 A464 L465 C466 L467 GLN SER ASP THR A472 D473 F474 T475 D476 V479

V480 C481 G482 D483 G484 A488 I489 L490 K491 T492 N493 D494 S495 F496 N497 I498 V499 A500 L501 T502 P503

- Molecule 40: Pre-mRNA-splicing factor SYF2

Chain y: 9% 61% • 38%


MET D2 K11 S36 ALA ASN ARG LYS PRO ARG VAL TYR SER MET GLU ASP VAL ASN ASP ALA ASP GLU SER VAL GLY ASP THR GLU SER PRO GLU LYS E65 Y75 R81 HIS PRO GLN GLY LYS THR GLY GLN SER GLN ARG GLY GLY I95 T116 GLN ASN SER SER LYS
GLN ASP SER SER ALA ASP GLU GLU ASP ASN LYS ASN VAL PRO LYS LYS GLY ARG ILE GLY K142 V143 Q144 K145 D146 T147 K148 T149 D157 D158 Q184 N185 ALA LYS THR LEU TYR GLY VAL GLU SER PHE T196 N197 D198 K199 N200 E205 S208 R209 E210 S211 LYS GLY SER GLU

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	403474	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 ($e^-/Å^2$)	40	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	140.475	Depositor
Minimum map value	-91.484	Depositor
Average map value	0.011	Depositor
Map value standard deviation	1.810	Depositor
Recommended contour level	5	Depositor
Map size (Å)	458.0, 458.0, 458.0	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.145, 1.145, 1.145	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: MG, ZN, K, GTP, SEP, KGN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	2	0.68	1/4585 (0.0%)	1.65	135/7114 (1.9%)
2	5	0.66	2/4221 (0.0%)	1.60	118/6573 (1.8%)
3	6	0.52	1/2427 (0.0%)	1.35	38/3778 (1.0%)
4	A	0.45	0/18498	0.71	7/25078 (0.0%)
5	B	0.55	2/13971 (0.0%)	0.86	28/18941 (0.1%)
6	C	0.41	0/7291	0.69	1/9875 (0.0%)
7	D	0.44	0/1478	0.73	1/1967 (0.1%)
8	E	0.54	0/341	1.40	8/530 (1.5%)
9	F	0.42	0/459	0.63	0/613
10	G	0.41	0/1051	0.64	1/1406 (0.1%)
11	H	0.44	0/3767	0.73	3/5076 (0.1%)
12	I	0.68	0/1187	1.67	38/1839 (2.1%)
13	J	0.44	1/2989 (0.0%)	0.80	6/4055 (0.1%)
14	K	0.40	0/1479	0.68	0/1995
15	L	0.39	0/1307	0.70	1/1748 (0.1%)
16	M	0.40	0/2094	0.76	4/2815 (0.1%)
17	N	0.43	0/2124	0.73	0/2860
18	O	0.44	0/2049	0.72	0/2748
19	P	0.42	0/623	0.75	1/832 (0.1%)
20	Q	0.61	4/5056 (0.1%)	0.94	22/6846 (0.3%)
21	R	0.42	0/557	0.69	0/750
22	S	0.45	1/4248 (0.0%)	0.69	1/5759 (0.0%)
23	T	0.49	1/5482 (0.0%)	0.77	6/7438 (0.1%)
24	W	0.44	0/1757	0.80	0/2372
26	Y	0.45	0/722	0.86	2/963 (0.2%)
27	Z	0.39	0/446	0.69	0/591
28	a	0.48	0/1154	0.72	0/1561
29	b	0.44	0/758	0.86	0/1018
29	k	0.50	1/836 (0.1%)	0.96	0/1120
30	c	0.39	0/295	0.65	0/386
31	d	0.50	0/642	0.81	0/868
31	n	0.56	0/642	0.90	0/868

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
32	e	0.45	0/616	0.84	0/835
32	p	0.51	0/616	0.88	0/835
33	f	0.46	0/614	0.74	0/830
33	q	0.46	0/614	0.79	0/830
34	g	0.49	0/562	0.95	2/756 (0.3%)
34	r	0.49	0/562	0.93	1/756 (0.1%)
35	h	0.48	0/828	0.80	2/1124 (0.2%)
35	l	0.59	2/828 (0.2%)	0.98	2/1124 (0.2%)
36	j	0.46	0/807	0.86	2/1083 (0.2%)
36	m	0.49	0/807	0.93	1/1083 (0.1%)
37	o	0.50	1/2737 (0.0%)	0.88	2/3696 (0.1%)
38	s	0.67	2/428 (0.5%)	0.95	3/577 (0.5%)
39	t	0.53	0/805	0.92	3/1094 (0.3%)
39	u	0.57	0/867	0.98	3/1178 (0.3%)
39	v	0.56	0/837	0.92	2/1137 (0.2%)
39	w	0.48	0/3469	0.88	11/4707 (0.2%)
40	y	0.38	0/1008	0.63	0/1350
All	All	0.50	19/111541 (0.0%)	0.93	455/153378 (0.3%)

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
5	B	0	1
6	C	0	1
20	Q	0	2
23	T	0	2
29	b	0	1
36	m	0	1
37	o	0	2
38	s	0	1
39	w	0	1
All	All	0	12

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	5	128	A	N9-C4	9.45	1.43	1.37
5	B	756	TRP	CB-CG	8.09	1.64	1.50
38	s	122	LYS	CB-CG	-7.31	1.32	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
20	Q	906	CYS	CB-SG	-6.53	1.71	1.82
37	o	352	CYS	CB-SG	-6.51	1.71	1.82

The worst 5 of 455 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	2	77	U	N1-C2-O2	15.45	133.62	122.80
1	2	1103	C	N1-C2-O2	15.07	127.94	118.90
1	2	77	U	N3-C4-O4	-14.99	108.91	119.40
3	6	54	U	N1-C2-O2	14.34	132.84	122.80
1	2	77	U	N3-C2-O2	-14.29	112.20	122.20

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	B	1519	ARG	Sidechain
6	C	132	ARG	Sidechain
20	Q	478	ARG	Sidechain
20	Q	496	ARG	Sidechain
23	T	546	SER	Mainchain

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	2	4120	0	2095	2	0
2	5	3777	0	1908	3	0
3	6	2170	0	1095	0	0
4	A	18036	0	17967	5	0
5	B	13675	0	13679	8	0
6	C	7139	0	7304	1	0
7	D	1547	0	1503	0	0
8	E	304	0	151	0	0
9	F	505	0	485	0	0
10	G	1090	0	1000	0	0
11	H	3705	0	3777	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	I	1068	0	538	0	0
13	J	2926	0	2918	2	0
14	K	1455	0	1482	1	0
15	L	1283	0	1301	0	0
16	M	2048	0	2011	0	0
17	N	2092	0	2162	0	0
18	O	2143	0	2032	1	0
19	P	607	0	596	0	0
20	Q	4959	0	5070	3	0
21	R	555	0	491	0	0
22	S	4170	0	3705	3	0
23	T	5387	0	4887	4	0
24	W	1734	0	1787	2	0
25	X	355	0	79	0	0
26	Y	713	0	746	0	0
27	Z	446	0	486	0	0
28	a	1132	0	1166	0	0
29	b	752	0	811	0	0
29	k	830	0	905	0	0
30	c	293	0	290	0	0
31	d	633	0	660	0	0
31	n	633	0	660	0	0
32	e	606	0	630	0	0
32	p	606	0	630	0	0
33	f	601	0	600	0	0
33	q	601	0	600	0	0
34	g	557	0	575	0	0
34	r	557	0	575	0	0
35	h	819	0	866	0	0
35	l	819	0	866	0	0
36	j	795	0	830	0	0
36	m	795	0	830	0	0
37	o	2673	0	2614	0	0
38	s	426	0	385	0	0
39	t	794	0	784	0	0
39	u	856	0	827	0	0
39	v	827	0	807	0	0
39	w	3405	0	3303	0	0
40	y	1003	0	909	0	0
41	6	5	0	0	0	0
41	C	1	0	0	0	0
42	6	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
43	A	36	0	0	0	0
44	C	32	0	12	0	0
45	D	1	0	0	0	0
45	L	3	0	0	0	0
45	M	1	0	0	0	0
45	N	2	0	0	0	0
45	c	1	0	0	0	0
All	All	109105	0	102390	33	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 0.

The worst 5 of 33 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)
24:W:17:ASP:OD1	24:W:17:ASP:N	2.38	0.52
5:B:2011:ILE:HD12	5:B:2030:ILE:HD11	1.94	0.50
5:B:599:ILE:HG23	5:B:605:LEU:HD11	1.96	0.47
5:B:793:LEU:HD21	5:B:808:LEU:HD11	1.98	0.46
5:B:1759:ALA:HB2	5:B:1846:THR:HG21	1.99	0.45

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	2185/2413 (91%)	2146 (98%)	38 (2%)	1 (0%)	100	100
5	B	1703/2163 (79%)	1657 (97%)	46 (3%)	0	100	100
6	C	888/1008 (88%)	873 (98%)	15 (2%)	0	100	100
7	D	171/291 (59%)	167 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	F	52/179 (29%)	52 (100%)	0	0	100	100
10	G	119/235 (51%)	119 (100%)	0	0	100	100
11	H	449/577 (78%)	440 (98%)	9 (2%)	0	100	100
13	J	366/451 (81%)	354 (97%)	12 (3%)	0	100	100
14	K	172/379 (45%)	170 (99%)	2 (1%)	0	100	100
15	L	154/157 (98%)	151 (98%)	3 (2%)	0	100	100
16	M	253/339 (75%)	249 (98%)	4 (2%)	0	100	100
17	N	256/364 (70%)	251 (98%)	5 (2%)	0	100	100
18	O	250/590 (42%)	245 (98%)	5 (2%)	0	100	100
19	P	68/175 (39%)	67 (98%)	1 (2%)	0	100	100
20	Q	618/1071 (58%)	589 (95%)	29 (5%)	0	100	100
21	R	77/135 (57%)	77 (100%)	0	0	100	100
22	S	525/687 (76%)	518 (99%)	7 (1%)	0	100	100
23	T	692/859 (81%)	663 (96%)	28 (4%)	1 (0%)	48	77
24	W	210/238 (88%)	199 (95%)	11 (5%)	0	100	100
26	Y	86/111 (78%)	83 (96%)	3 (4%)	0	100	100
27	Z	51/140 (36%)	51 (100%)	0	0	100	100
28	a	132/251 (53%)	131 (99%)	1 (1%)	0	100	100
29	b	89/196 (45%)	88 (99%)	1 (1%)	0	100	100
29	k	98/196 (50%)	94 (96%)	4 (4%)	0	100	100
30	c	33/382 (9%)	31 (94%)	2 (6%)	0	100	100
31	d	80/101 (79%)	80 (100%)	0	0	100	100
31	n	80/101 (79%)	77 (96%)	3 (4%)	0	100	100
32	e	73/94 (78%)	71 (97%)	1 (1%)	1 (1%)	9	30
32	p	73/94 (78%)	71 (97%)	2 (3%)	0	100	100
33	f	73/86 (85%)	71 (97%)	2 (3%)	0	100	100
33	q	73/86 (85%)	72 (99%)	1 (1%)	0	100	100
34	g	70/77 (91%)	64 (91%)	6 (9%)	0	100	100
34	r	70/77 (91%)	65 (93%)	5 (7%)	0	100	100
35	h	104/146 (71%)	98 (94%)	6 (6%)	0	100	100
35	l	104/146 (71%)	98 (94%)	6 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
36	j	95/110 (86%)	89 (94%)	6 (6%)	0	100	100
36	m	95/110 (86%)	90 (95%)	5 (5%)	0	100	100
37	o	323/455 (71%)	308 (95%)	14 (4%)	1 (0%)	37	67
38	s	57/175 (33%)	55 (96%)	2 (4%)	0	100	100
39	t	103/503 (20%)	103 (100%)	0	0	100	100
39	u	113/503 (22%)	111 (98%)	2 (2%)	0	100	100
39	v	109/503 (22%)	105 (96%)	4 (4%)	0	100	100
39	w	426/503 (85%)	419 (98%)	5 (1%)	2 (0%)	25	56
40	y	124/215 (58%)	124 (100%)	0	0	100	100
All	All	11942/17672 (68%)	11636 (97%)	300 (2%)	6 (0%)	50	77

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
32	e	32	PHE
39	w	376	LYS
4	A	1347	ARG
37	o	348	GLN
39	w	351	ASP

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	1987/2182 (91%)	1978 (100%)	9 (0%)	86	95
5	B	1536/1955 (79%)	1528 (100%)	8 (0%)	86	95
6	C	799/910 (88%)	795 (100%)	4 (0%)	86	95
7	D	165/252 (66%)	165 (100%)	0	100	100
9	F	52/154 (34%)	52 (100%)	0	100	100
10	G	102/206 (50%)	102 (100%)	0	100	100
11	H	422/538 (78%)	419 (99%)	3 (1%)	81	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
13	J	325/397 (82%)	323 (99%)	2 (1%)	84	95
14	K	164/328 (50%)	162 (99%)	2 (1%)	67	89
15	L	140/141 (99%)	140 (100%)	0	100	100
16	M	219/296 (74%)	218 (100%)	1 (0%)	86	95
17	N	243/332 (73%)	242 (100%)	1 (0%)	89	96
18	O	199/504 (40%)	198 (100%)	1 (0%)	86	95
19	P	61/151 (40%)	60 (98%)	1 (2%)	58	85
20	Q	557/969 (58%)	551 (99%)	6 (1%)	70	90
21	R	47/121 (39%)	45 (96%)	2 (4%)	25	57
22	S	369/633 (58%)	365 (99%)	4 (1%)	70	90
23	T	505/786 (64%)	502 (99%)	3 (1%)	84	95
24	W	201/219 (92%)	195 (97%)	6 (3%)	36	70
26	Y	79/100 (79%)	79 (100%)	0	100	100
27	Z	51/128 (40%)	51 (100%)	0	100	100
28	a	127/225 (56%)	126 (99%)	1 (1%)	79	93
29	b	86/176 (49%)	85 (99%)	1 (1%)	67	89
29	k	95/176 (54%)	95 (100%)	0	100	100
30	c	33/346 (10%)	33 (100%)	0	100	100
31	d	71/89 (80%)	71 (100%)	0	100	100
31	n	71/89 (80%)	70 (99%)	1 (1%)	62	87
32	e	70/83 (84%)	69 (99%)	1 (1%)	62	87
32	p	70/83 (84%)	68 (97%)	2 (3%)	37	71
33	f	67/77 (87%)	67 (100%)	0	100	100
33	q	67/77 (87%)	66 (98%)	1 (2%)	60	86
34	g	61/66 (92%)	59 (97%)	2 (3%)	33	67
34	r	61/66 (92%)	61 (100%)	0	100	100
35	h	95/129 (74%)	95 (100%)	0	100	100
35	l	95/129 (74%)	95 (100%)	0	100	100
36	j	90/103 (87%)	89 (99%)	1 (1%)	70	90
36	m	90/103 (87%)	89 (99%)	1 (1%)	70	90
37	o	298/412 (72%)	294 (99%)	4 (1%)	65	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
38	s	30/165 (18%)	30 (100%)	0	100	100
39	t	81/451 (18%)	81 (100%)	0	100	100
39	u	83/451 (18%)	80 (96%)	3 (4%)	30	64
39	v	80/451 (18%)	80 (100%)	0	100	100
39	w	369/451 (82%)	368 (100%)	1 (0%)	91	97
40	y	91/193 (47%)	89 (98%)	2 (2%)	47	79
All	All	10504/15893 (66%)	10430 (99%)	74 (1%)	80	94

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
34	g	74	LEU
39	w	245	HIS
36	m	71	ASN
32	p	46	PHE
13	J	361	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
20	Q	942	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	187/1175 (15%)	42 (22%)	7 (3%)
12	I	53/95 (55%)	21 (39%)	4 (7%)
2	5	177/214 (82%)	33 (18%)	8 (4%)
3	6	101/112 (90%)	16 (15%)	1 (0%)
8	E	13/47 (27%)	1 (7%)	0
All	All	531/1643 (32%)	113 (21%)	20 (3%)

5 of 113 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	6	U
1	2	16	U
1	2	19	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	2	20	G
1	2	21	G

5 of 20 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	5	166	U
12	I	62	A
12	I	88	U
12	I	76	U
1	2	1101	C

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
37	SEP	o	73	37	8,9,10	1.53	1 (12%)	8,12,14	1.53	2 (25%)

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
37	SEP	o	73	37	-	0/5/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
37	o	73	SEP	P-O1P	3.36	1.61	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	o	73	SEP	P-OG-CB	-2.75	110.73	118.30
37	o	73	SEP	OG-CB-CA	2.70	110.77	108.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 15 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	$\#\|Z\| > 2$
43	KGN	A	2500	-	36,36,36	0.78	0	54,60,60	0.51	0
44	GTP	C	1101	41	26,34,34	1.13	1 (3%)	32,54,54	1.43	5 (15%)

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
43	KGN	A	2500	-	-	4/30/54/54	0/1/1/1
44	GTP	C	1101	41	-	3/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
44	C	1101	GTP	C5-C6	-3.98	1.39	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
44	C	1101	GTP	PA-O3A-PB	-3.55	120.66	132.83
44	C	1101	GTP	C5-C6-N1	3.15	119.51	113.95
44	C	1101	GTP	C8-N7-C5	3.02	108.74	102.99
44	C	1101	GTP	C2-N1-C6	-2.85	119.86	125.10
44	C	1101	GTP	PB-O3B-PG	-2.78	123.29	132.83

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

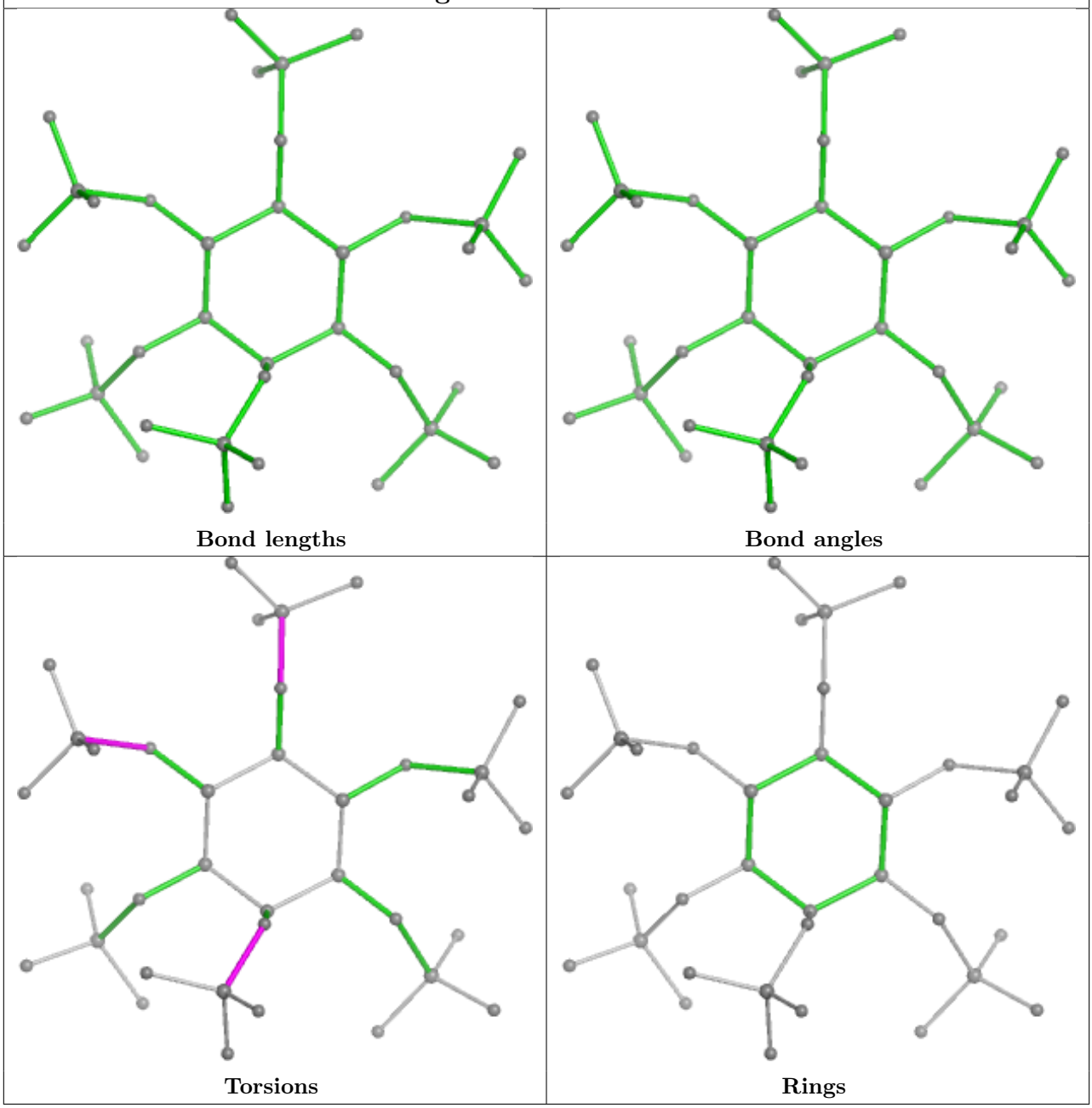
Mol	Chain	Res	Type	Atoms
43	A	2500	KGN	C3-O13-P3-O23
43	A	2500	KGN	C6-O16-P6-O46
44	C	1101	GTP	C5'-O5'-PA-O3A
44	C	1101	GTP	PA-O3A-PB-O1B
43	A	2500	KGN	C5-O15-P5-O45

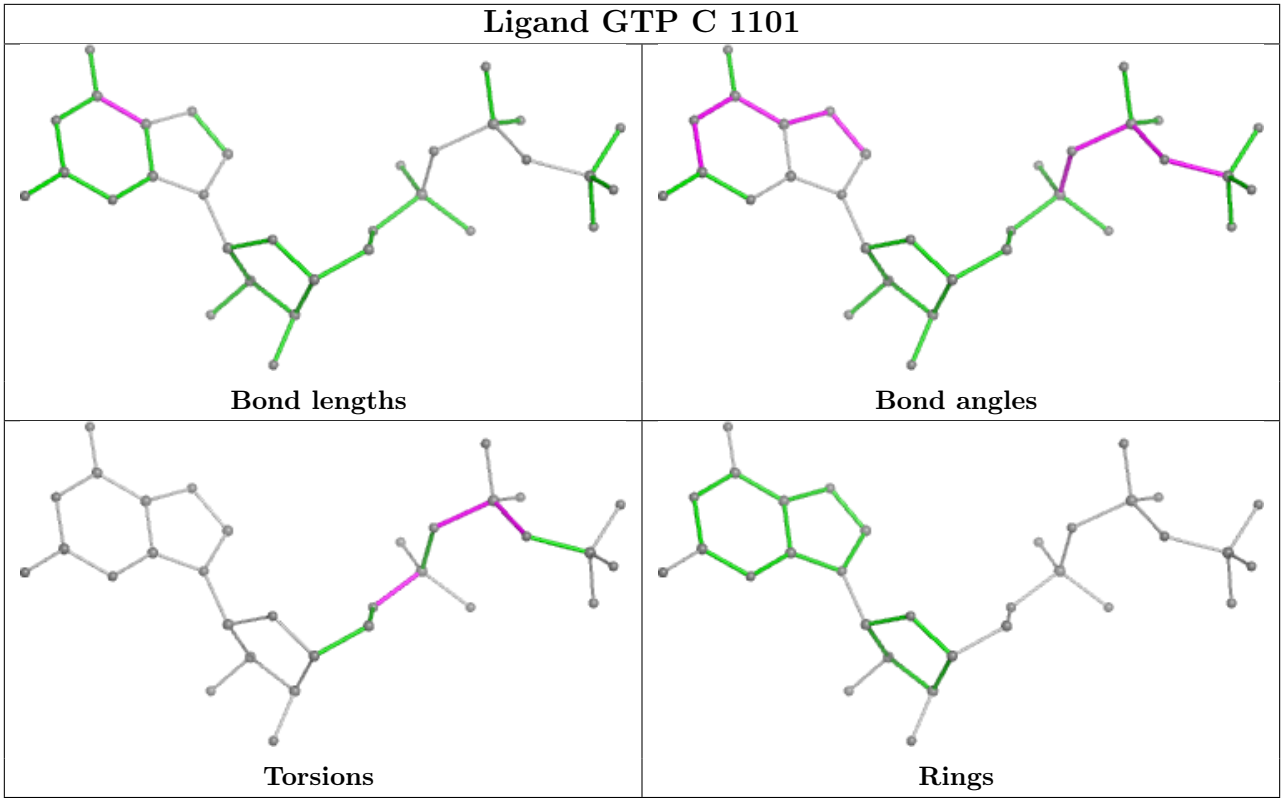
There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand KGN A 2500





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
25	X	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	X	18:UNK	C	35:UNK	N	201.98
1	X	51:UNK	C	200:UNK	N	106.75

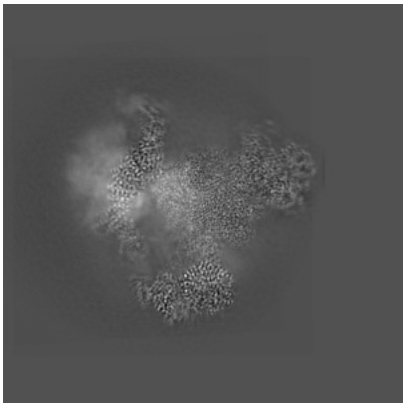
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-12106. These allow visual inspection of the internal detail of the map and identification of artifacts.

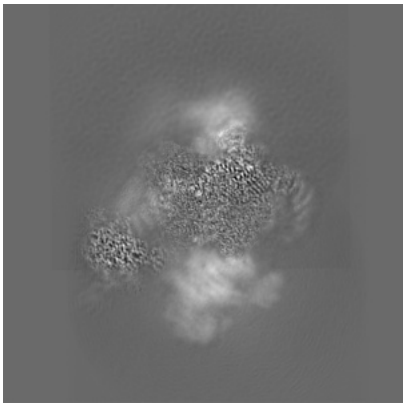
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

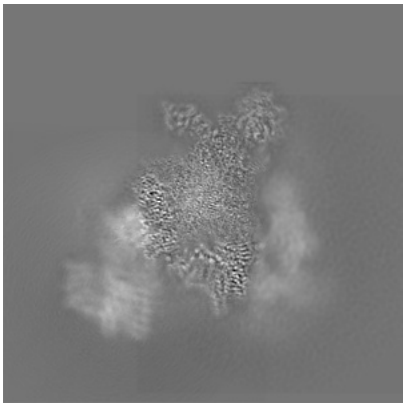
6.1.1 Primary map



X



Y

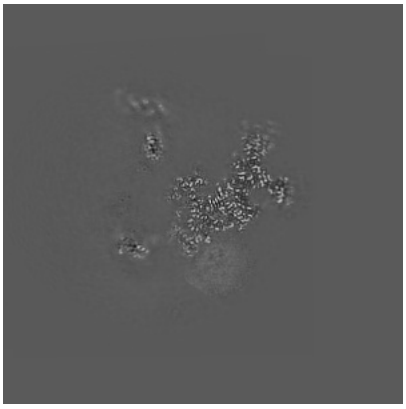


Z

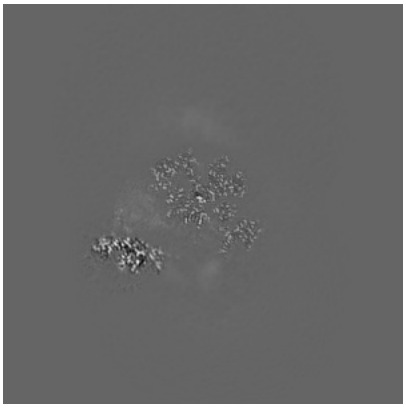
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

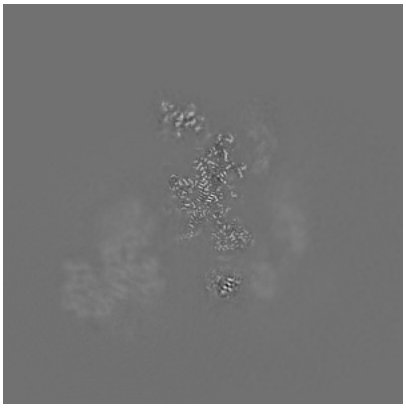
6.2.1 Primary map



X Index: 200



Y Index: 200



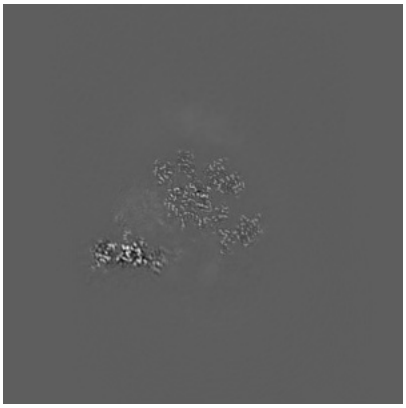
Z Index: 200

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

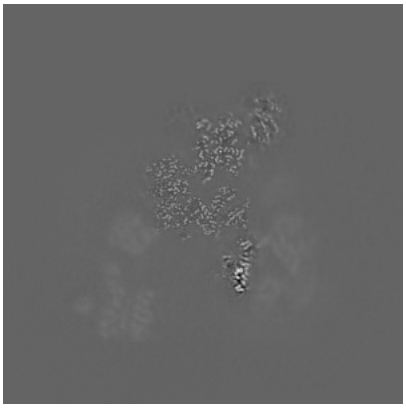
6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 235



Y Index: 205

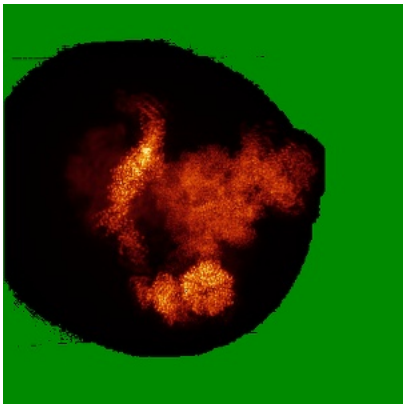


Z Index: 224

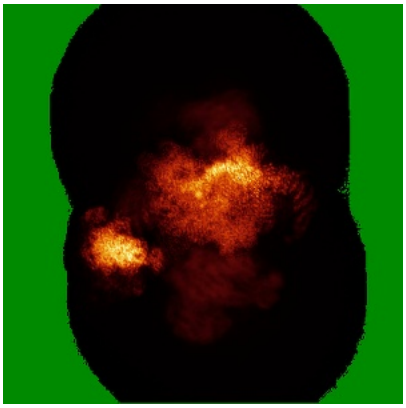
The images above show the largest variance slices of the map in three orthogonal directions.

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

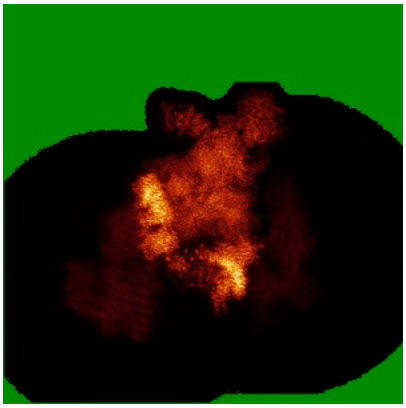
6.4.1 Primary map



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 5.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

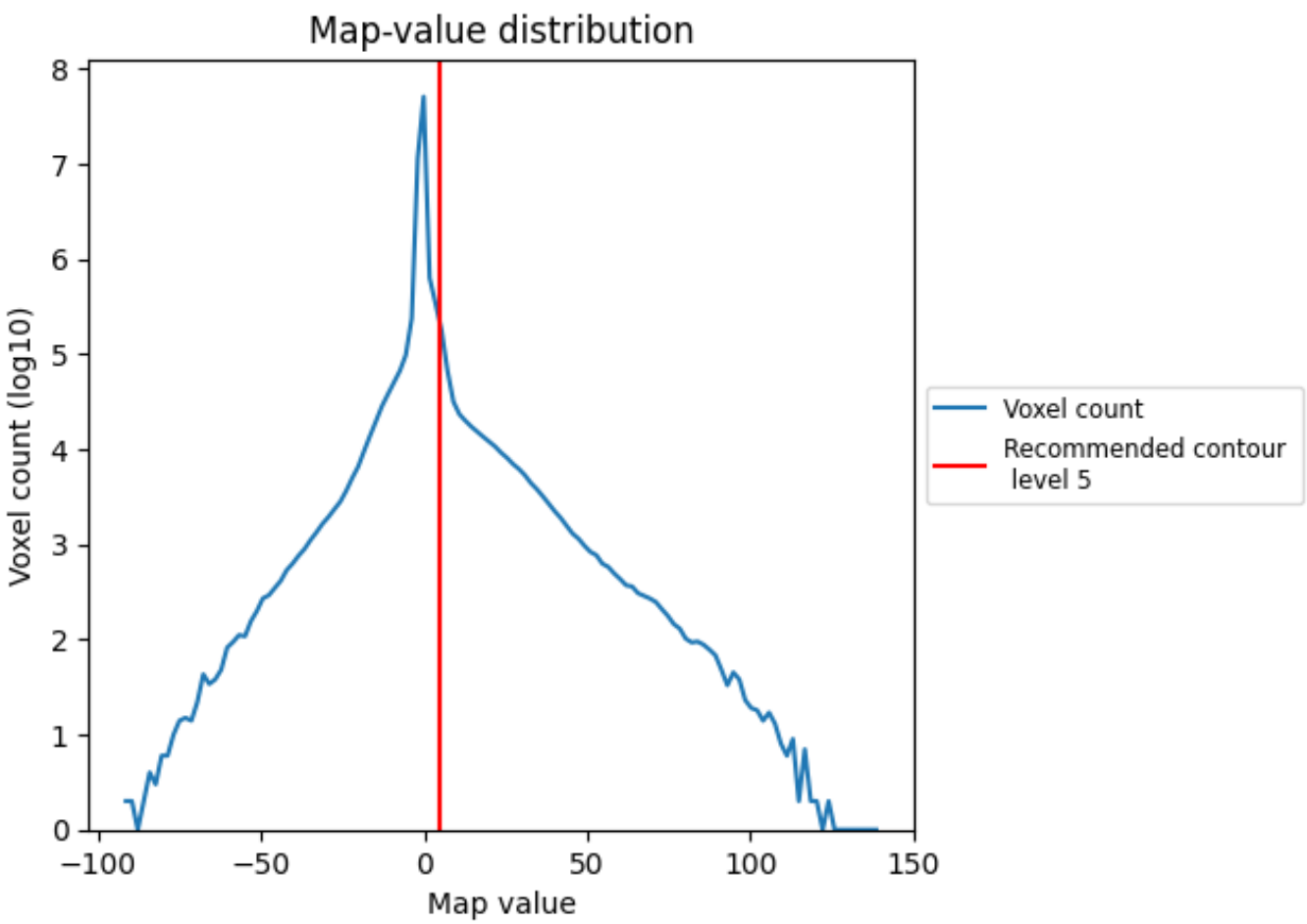
6.6 Mask visualisation (i)

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

7 Map analysis (i)

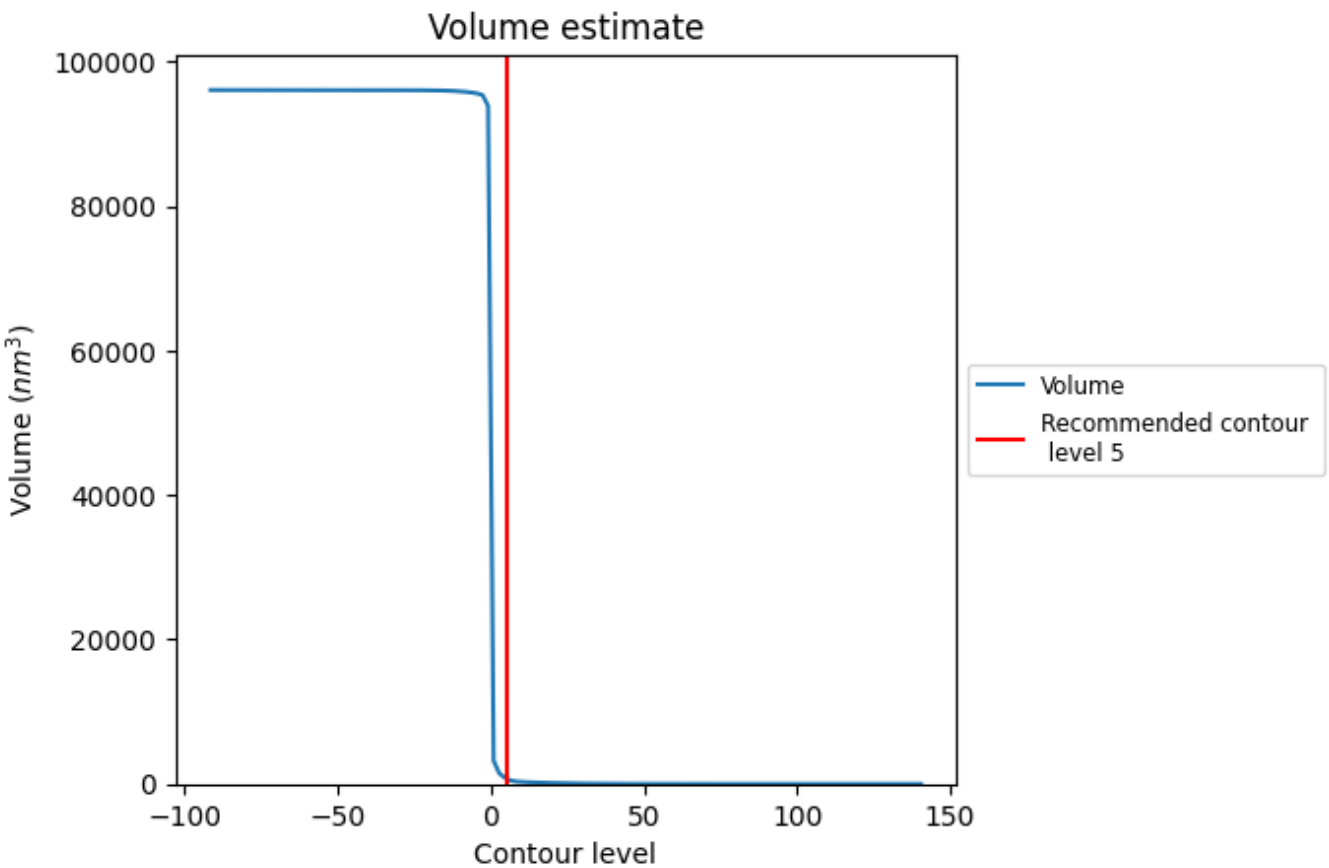
This section contains the results of statistical analysis of the map.

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.

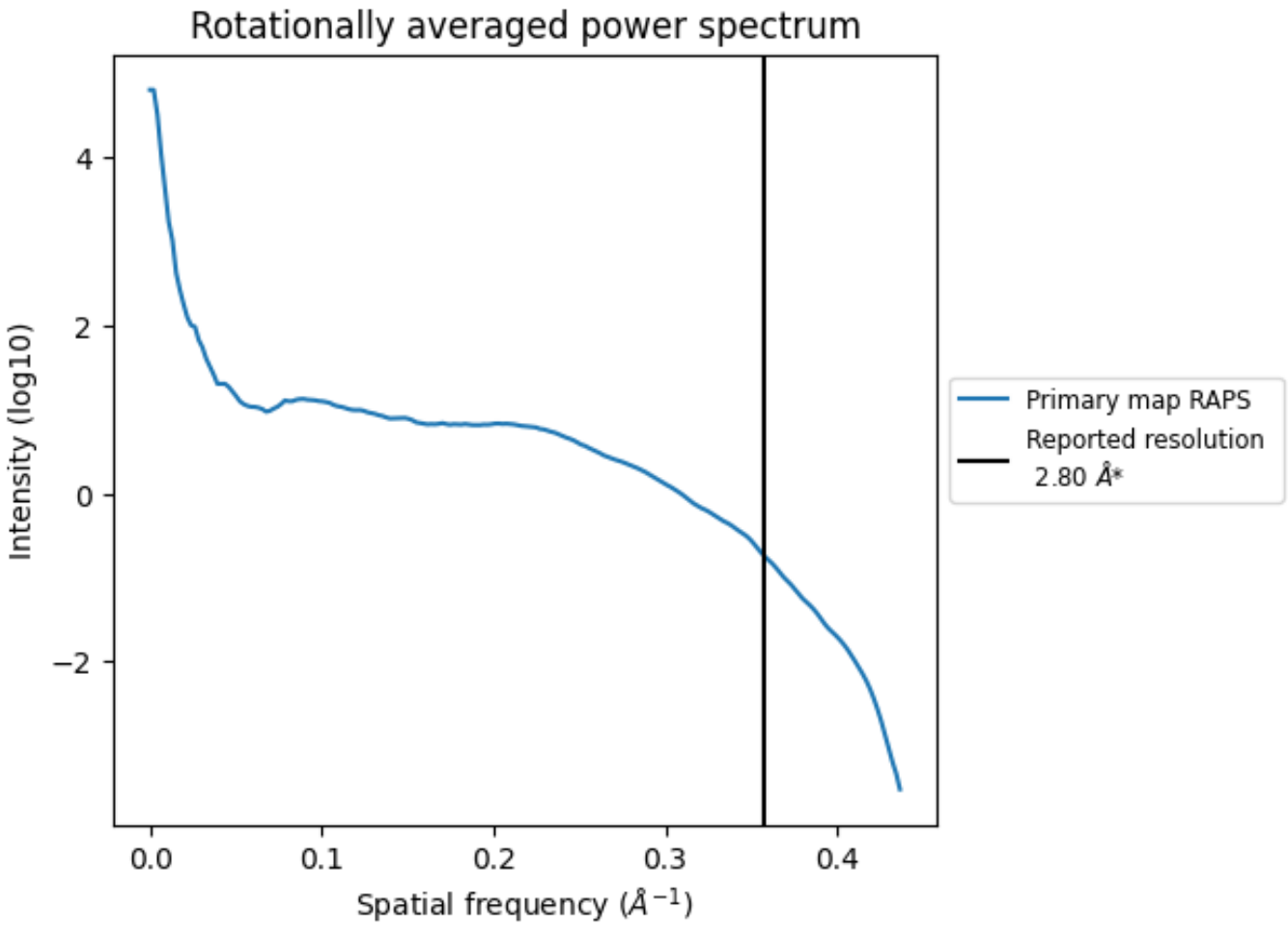
7.2 Volume estimate (i)



The volume at the recommended contour level is 773 nm^3; this corresponds to an approximate mass of 699 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 0.357 Å^{-1}

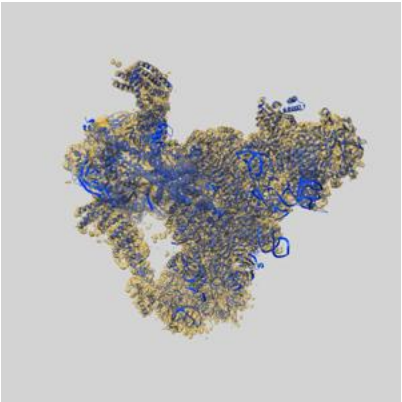
8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

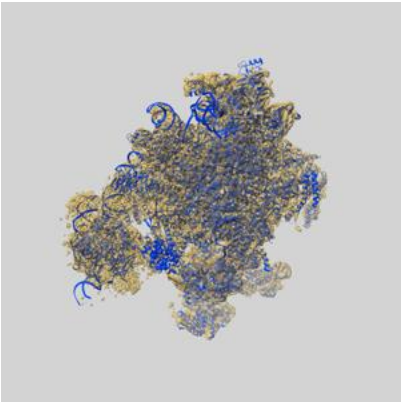
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-12106 and PDB model 7B9V. Per-residue inclusion information can be found in section 3 on page 16.

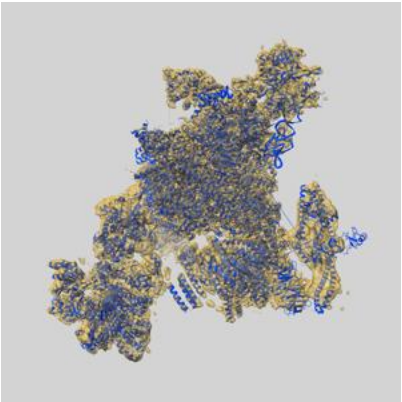
9.1 Map-model overlay (i)



X



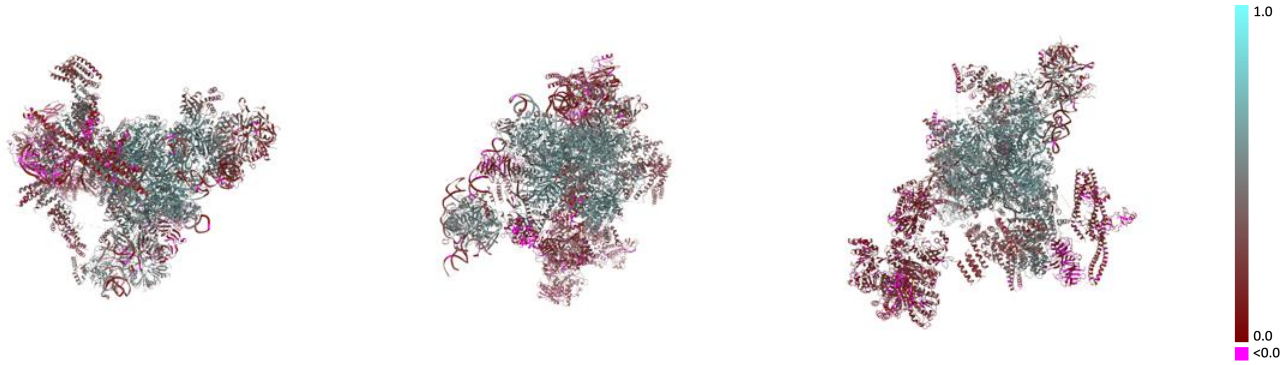
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 5.0 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.

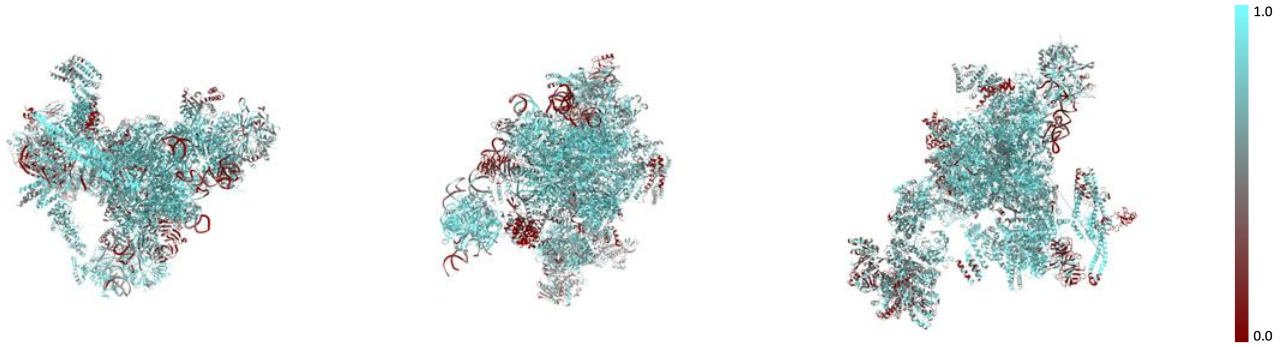
9.2 Q-score mapped to coordinate model (i)



X Y Z

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.

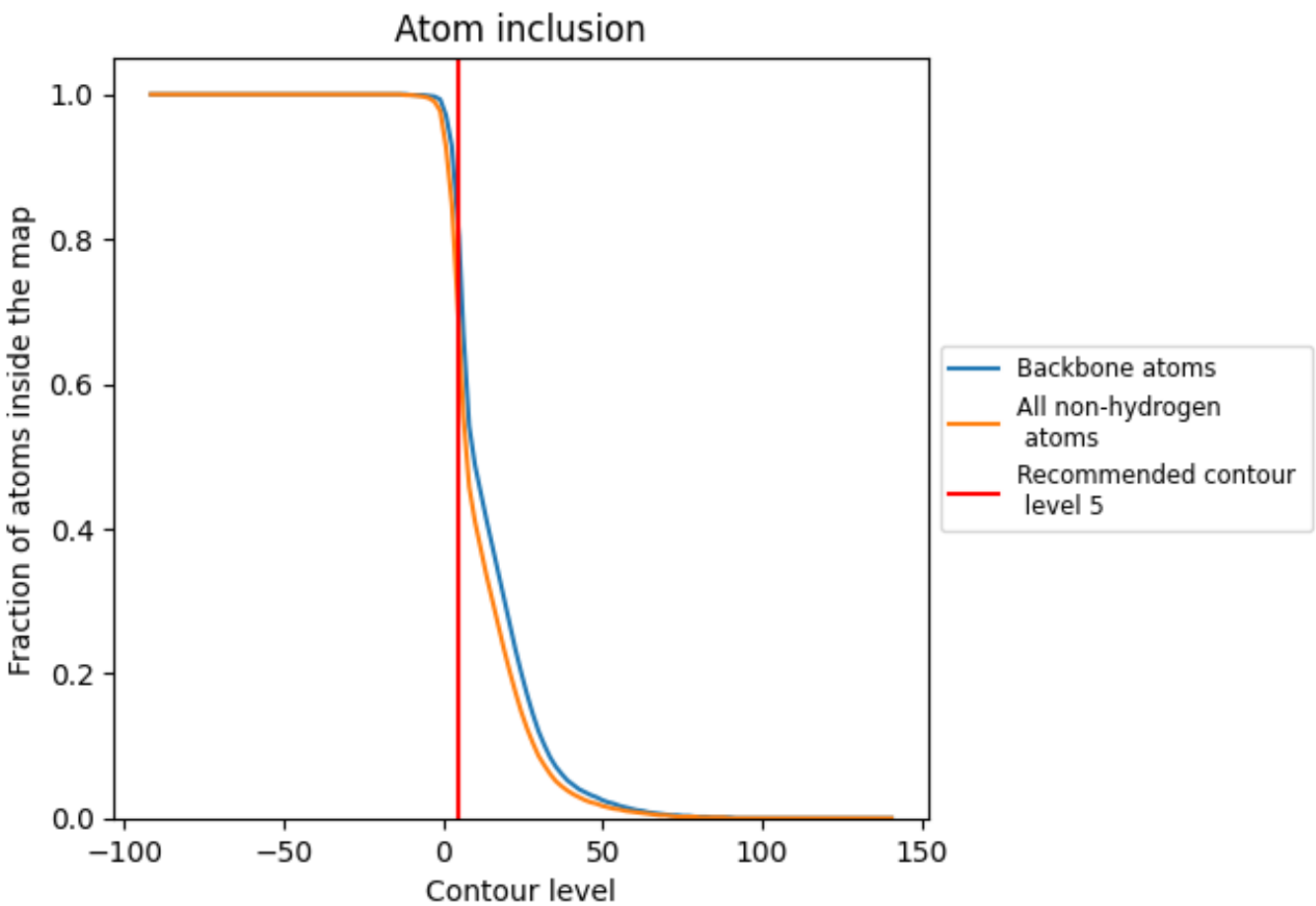
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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 (5).

9.4 Atom inclusion (i)



At the recommended contour level, 80% of all backbone atoms, 68% 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 (5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6770	0.3780
2	0.5380	0.2700
5	0.5300	0.3900
6	0.7920	0.5230
A	0.7870	0.5100
B	0.5320	0.1760
C	0.7960	0.5320
D	0.7260	0.4540
E	0.8060	0.5590
F	0.6180	0.4210
G	0.7500	0.4440
H	0.6740	0.3940
I	0.6010	0.3930
J	0.8720	0.5870
K	0.7050	0.5030
L	0.8460	0.5710
M	0.8320	0.5600
N	0.6440	0.4760
O	0.7030	0.4390
P	0.7720	0.5440
Q	0.6890	0.1520
R	0.4060	0.3950
S	0.7910	0.4500
T	0.8050	0.4100
W	0.8300	0.4660
X	0.7180	0.2720
Y	0.8490	0.4960
Z	0.7120	0.3900
a	0.0000	0.0600
b	0.6440	0.3770
c	0.1320	0.2090
d	0.7510	0.4440
e	0.6090	0.2880
f	0.5980	0.2340
g	0.6540	0.3420

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
h	0.5790	0.2710
j	0.5370	0.2290
k	0.8750	0.5160
l	0.8870	0.5170
m	0.8380	0.4800
n	0.8450	0.4850
o	0.3000	0.1940
p	0.8440	0.4680
q	0.8710	0.5020
r	0.8230	0.4780
s	0.7270	0.1840
t	0.3380	0.1080
u	0.6010	0.1290
v	0.6780	0.1380
w	0.3620	0.0630
y	0.6880	0.4420