wwPDB EM Validation Summary Report (i)

Nov 19, 2022 – 09:01 am GMT

PDB ID	:	5LQW
EMDB ID	:	EMD-4099
Title	:	yeast activated spliceosome
Authors	:	Rauhut, R.; Luehrmann, R.
Deposited on	:	2016-08-17
Resolution	:	5.80 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 (i)) were used in the production of this report:

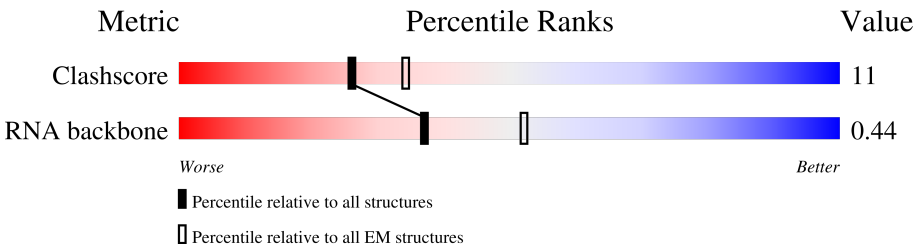
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 5.80 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	158937	4297
RNA backbone	4643	859

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

Mol	Chain	Length	Quality of chain
1	A	2413	8% 87% • 12%
2	B	1008	5% 82% • 16%
3	C	2163	10% 83% 16%
4	D	364	6% 51% 49%
5	E	157	10% 83% • 12%
6	F	339	6% 60% • 36%
7	H	577	43% 75% • 24%
8	J	148	5% 69% • 30%
9	K	451	• 73% • 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	L	266	12% 88%
11	M	379	10% 41% 55%
12	N	204	53% 77% 22%
13	O	876	20% 68% 28%
14	P	859	5% 23% 76%
15	Q	971	6% 80% 19%
16	R	687	32% 68%
17	W	590	21% 78%
18	X	1361	5% 79% 20%
19	Y	107	5% 79% 17%
20	Z	85	80% 20%
21	b	196	5% 41% 59%
22	d	101	13% 81% 19%
23	e	94	16% 80% 20%
24	f	86	20% 84% 16%
25	g	77	17% 90% 10%
26	h	146	5% 56% 44%
27	j	110	18% 85% 15%
28	2	1175	93%
29	5	179	7% 44% 26% 9% 21%
30	6	112	6% 39% 43% 7% 9%
31	9	572	5% 91%

2 Entry composition (i)

There are 31 unique types of molecules in this entry. The entry contains 18351 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 protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
1	A	2130	2130	2130	0	2130

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	153	ASN	MET	conflict	UNP P33334

- Molecule 2 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
2	B	843	843	843	0	843

- Molecule 3 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
3	C	1811	1811	1811	0	1811

- Molecule 4 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
4	D	187	187	187	0	187

- Molecule 5 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
5	E	138	138	138	0	138

- Molecule 6 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
6	F	218	218	218	0	218

- Molecule 7 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
7	H	437	437	437	0	437

- Molecule 8 is a protein called U2 snRNP component IST3.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
8	J	104	104	104	0	104

- Molecule 9 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
9	K	338	338	338	0	338

- Molecule 10 is a protein called Pre-mRNA-splicing factor CWC26.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
10	L	32	32	32	0	32

- Molecule 11 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
11	M	169	169	169	0	169

- Molecule 12 is a protein called Pre-mRNA leakage protein 1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
12	N	159	159	159	0	159

- Molecule 13 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
13	O	628	628	628	0	628

- Molecule 14 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
14	P	202	202	202	0	202

- Molecule 15 is a protein called U2 snRNP component HSH155.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
15	Q	791	791	791	0	791

- Molecule 16 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
16	R	219	219	219	0	219

- Molecule 17 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
17	W	128	128	128	0	128

- Molecule 18 is a protein called Pre-mRNA-splicing factor RSE1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
18	X	1095	1095	1095	0	1095

- Molecule 19 is a protein called Pre-mRNA-splicing factor RDS3.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
19	Y	89	89	89	0	89

- Molecule 20 is a protein called RDS3 complex subunit 10.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
20	Z	68	68	68	0	68

- Molecule 21 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
21	b	80	80	80	0	80

- Molecule 22 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
22	d	82	82	82	0	82

- Molecule 23 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
23	e	75	75	75	0	75

- Molecule 24 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
24	f	72	72	72	0	72

- Molecule 25 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
25	g	69	69	69	0	69

- Molecule 26 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
26	h	82	82	82	0	82

- Molecule 27 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
27	j	94	94	94	0	94

- Molecule 28 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
28	2	81	1707	764	285	577	81	0	0

- Molecule 29 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
29	5	141	2999	1342	530	986	141	0	0

- Molecule 30 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
30	6	102	2170	972	386	710	102	0	0

- Molecule 31 is a RNA chain called actin pre-mRNA.

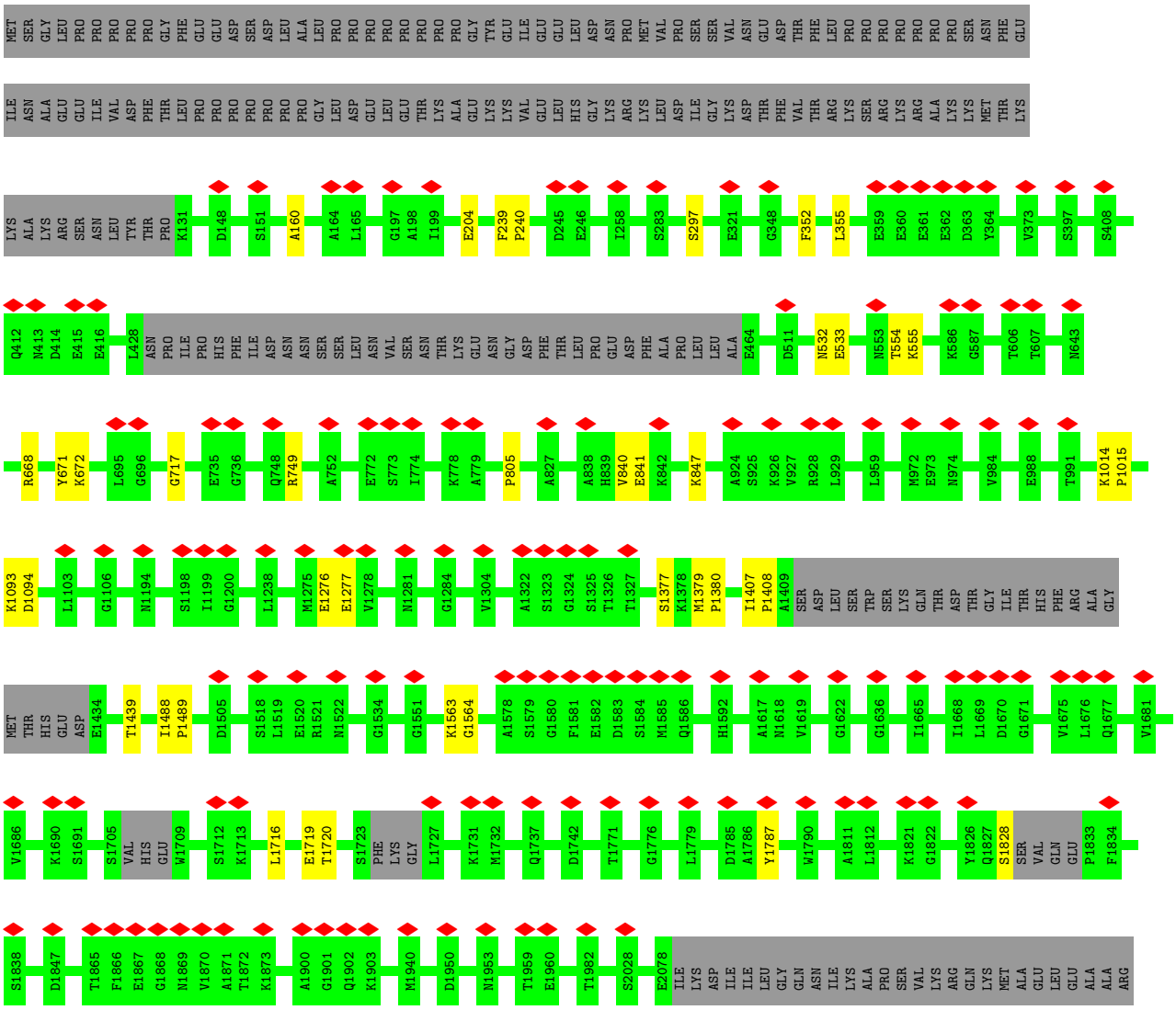
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
31	9	54	1135	509	187	385	54	0	0

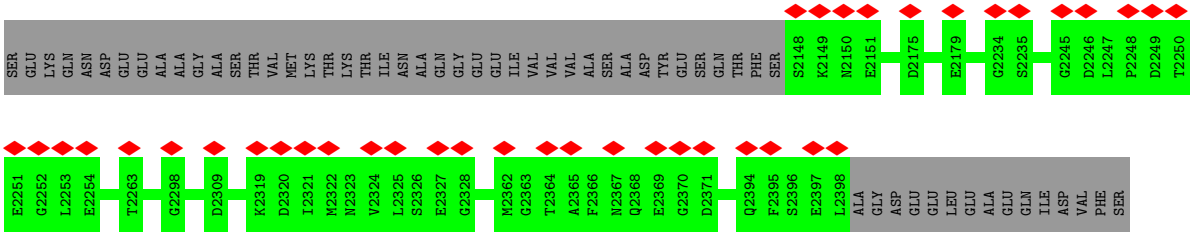
3 Residue-property plots (i)

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: Pre-mRNA-splicing factor 8

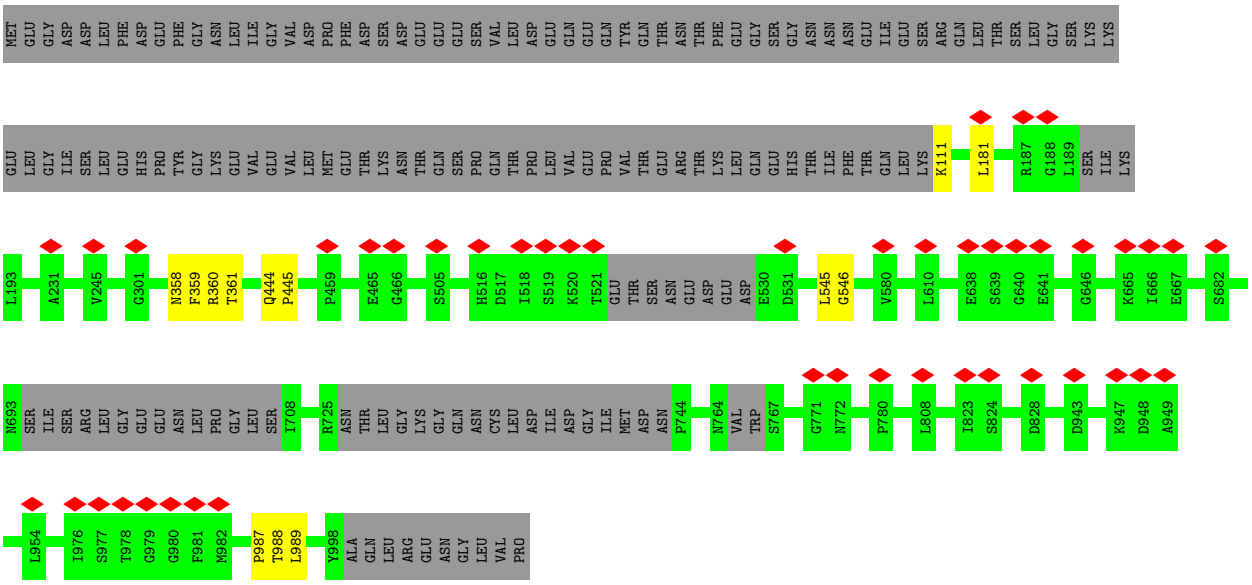
Chain A:





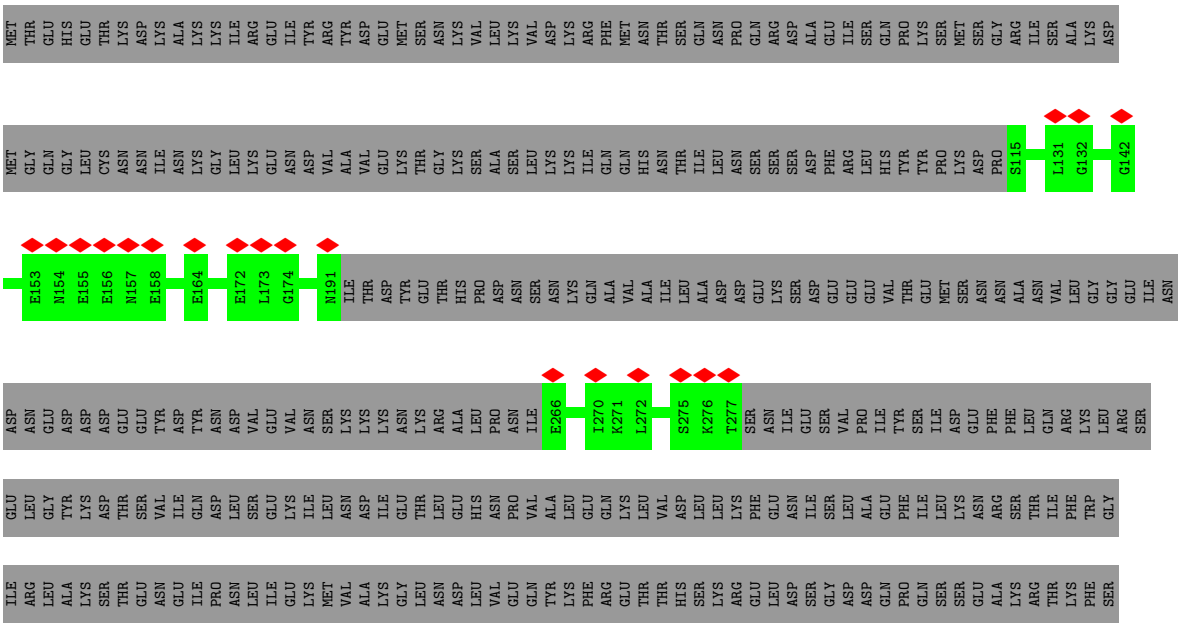
- Molecule 2: Pre-mRNA-splicing factor SNU114

Chain B:



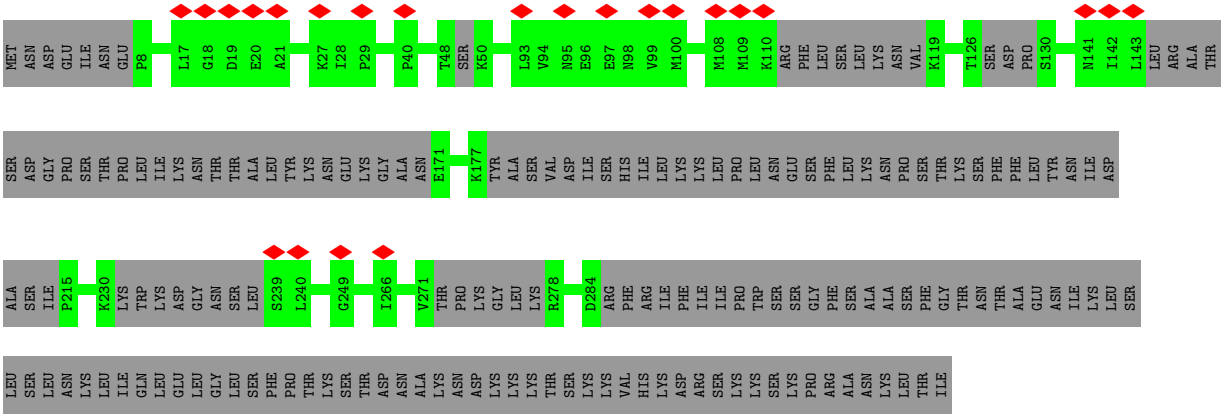
- Molecule 3: Pre-mRNA-splicing helicase BRR2

Chain C:



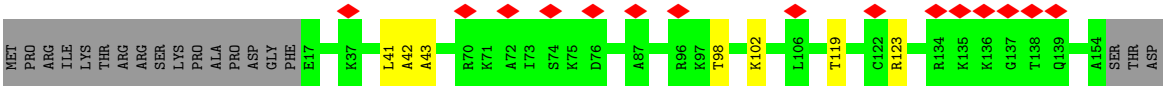
- Molecule 4: Pre-mRNA-splicing factor SLT11

Chain D:



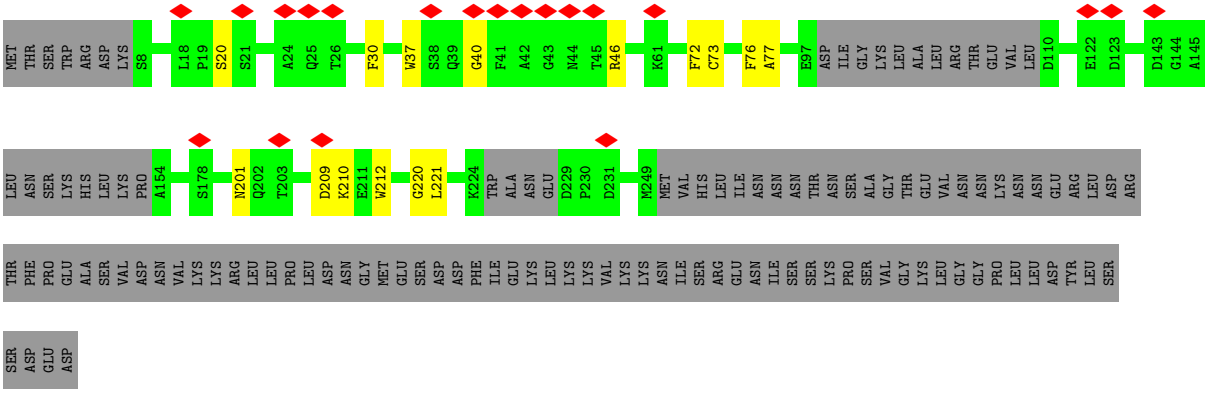
- Molecule 5: Pre-mRNA-splicing factor BUD31

Chain E:



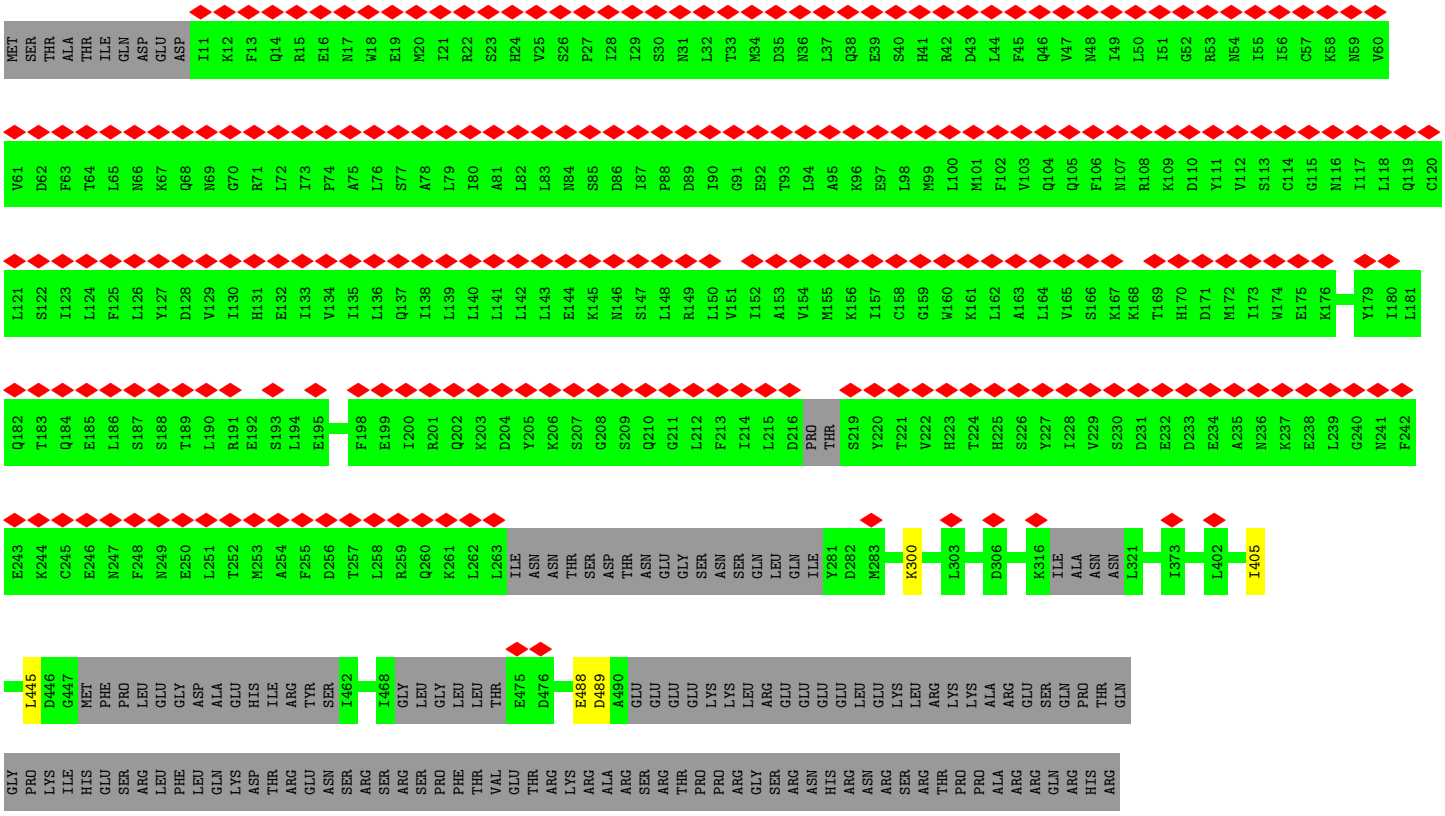
- Molecule 6: Pre-mRNA-splicing factor CWC2

Chain F:



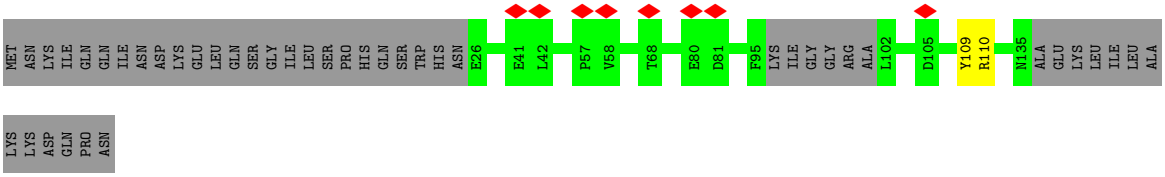
- Molecule 7: Pre-mRNA-splicing factor CWC22

Chain H:



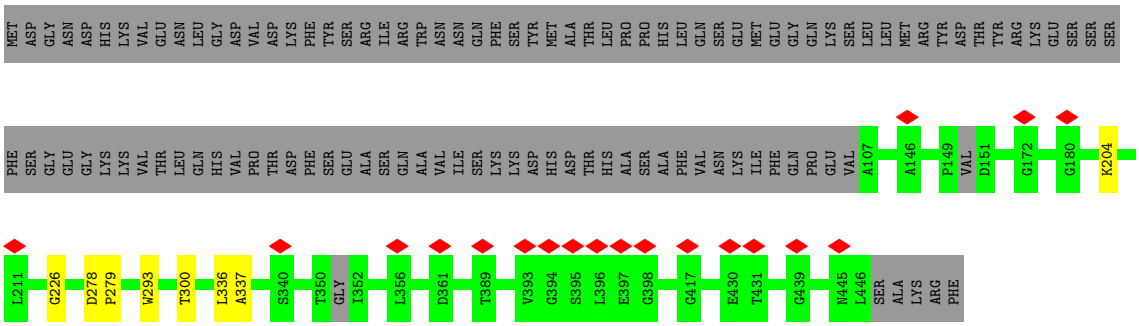
- Molecule 8: U2 snRNP component IST3

Chain J:



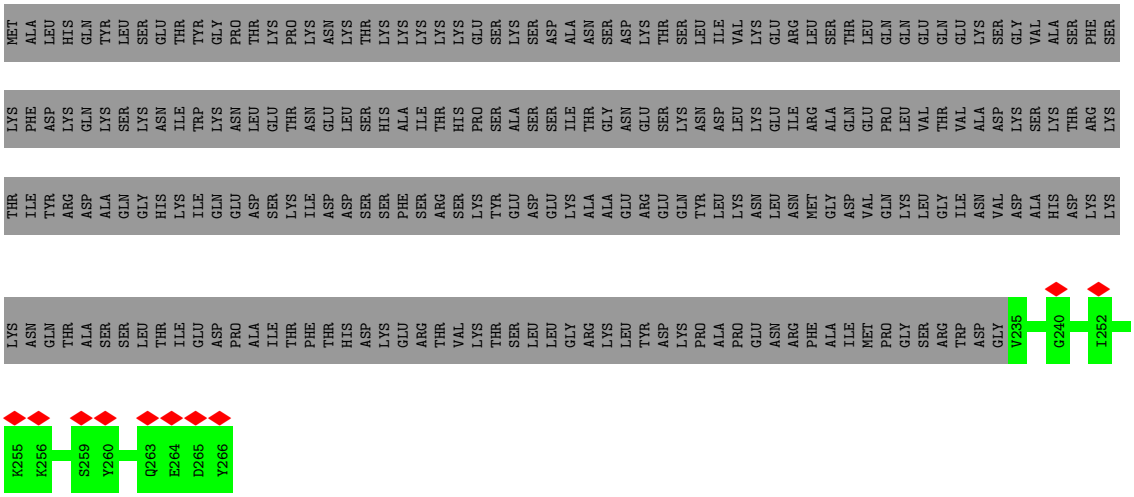
- Molecule 9: Pre-mRNA-splicing factor PRP46

Chain K:



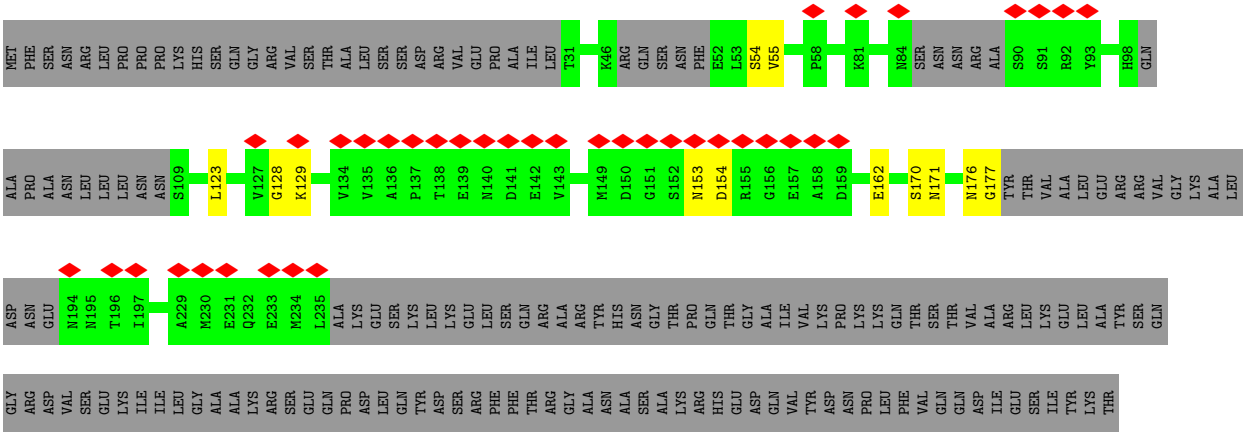
- Molecule 10: Pre-mRNA-splicing factor CWC26

Chain L:



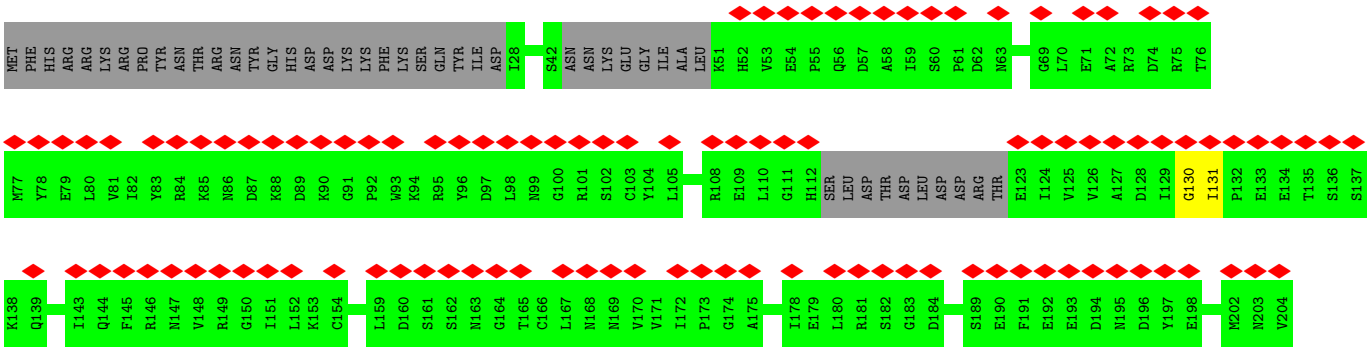
- Molecule 11: Pre-mRNA-processing protein 45

Chain M:



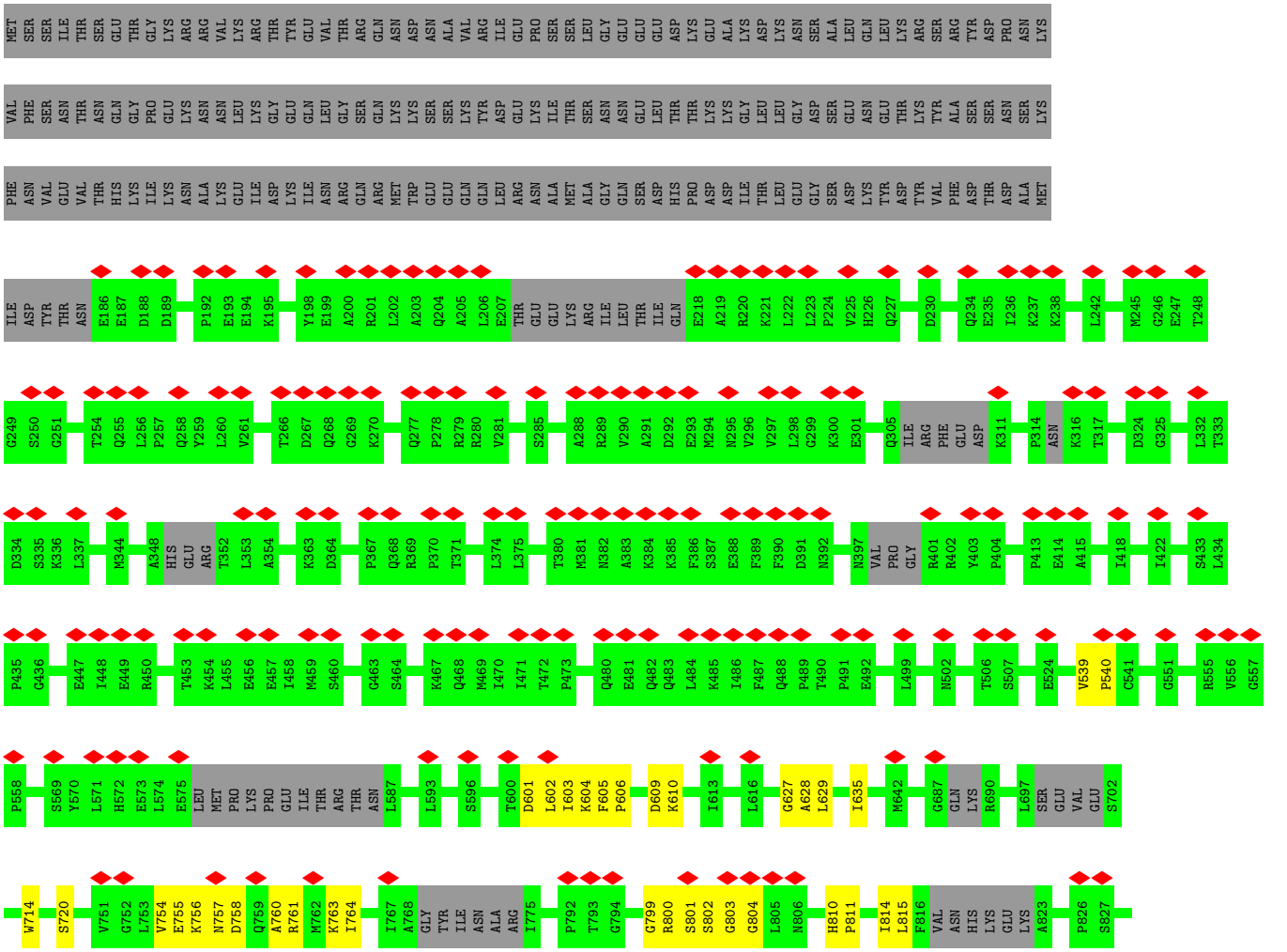
• Molecule 12: Pre-mRNA leakage protein 1

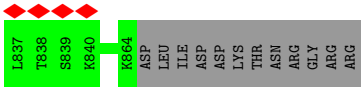
Chain N:



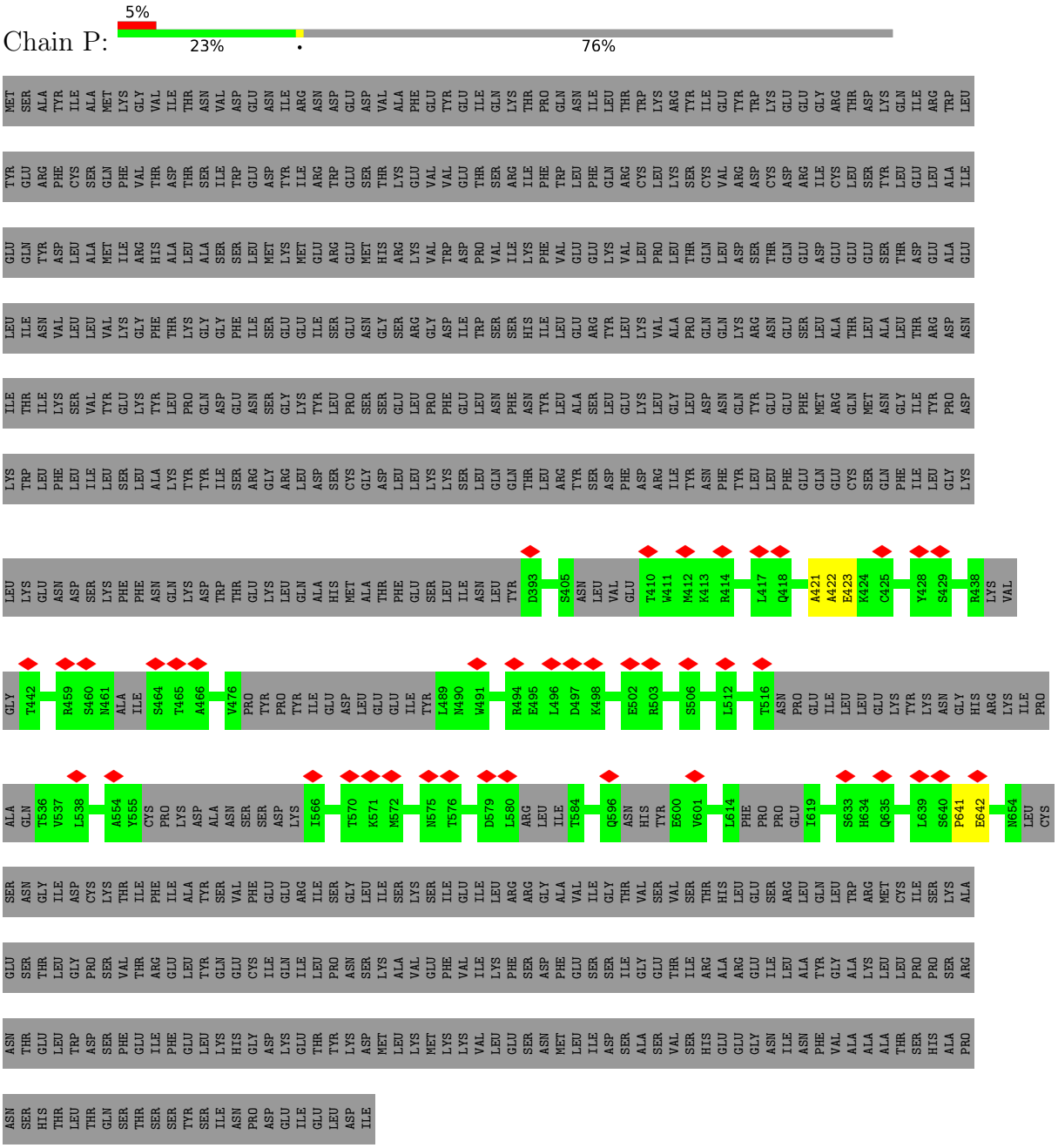
• Molecule 13: Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2

Chain O:

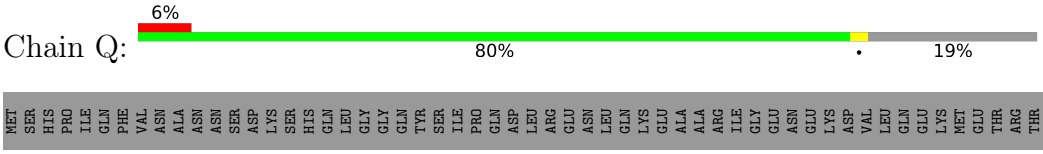


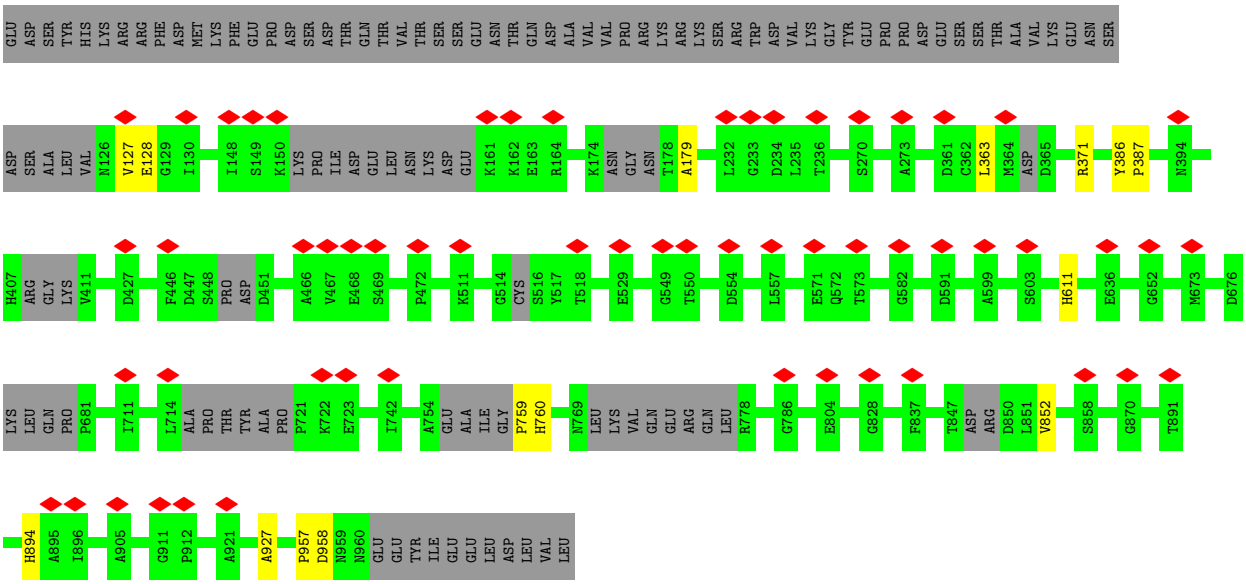


• Molecule 14: Pre-mRNA-splicing factor SYF1



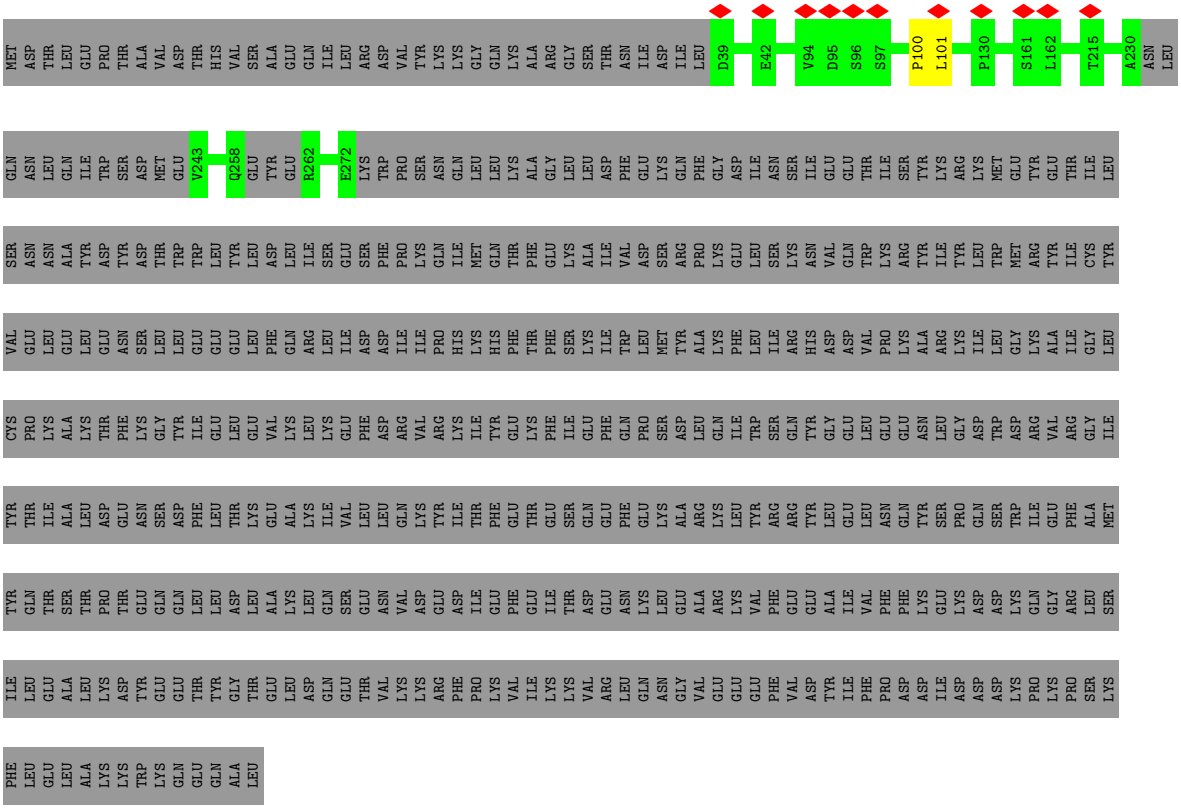
• Molecule 15: U2 snRNP component HSH155





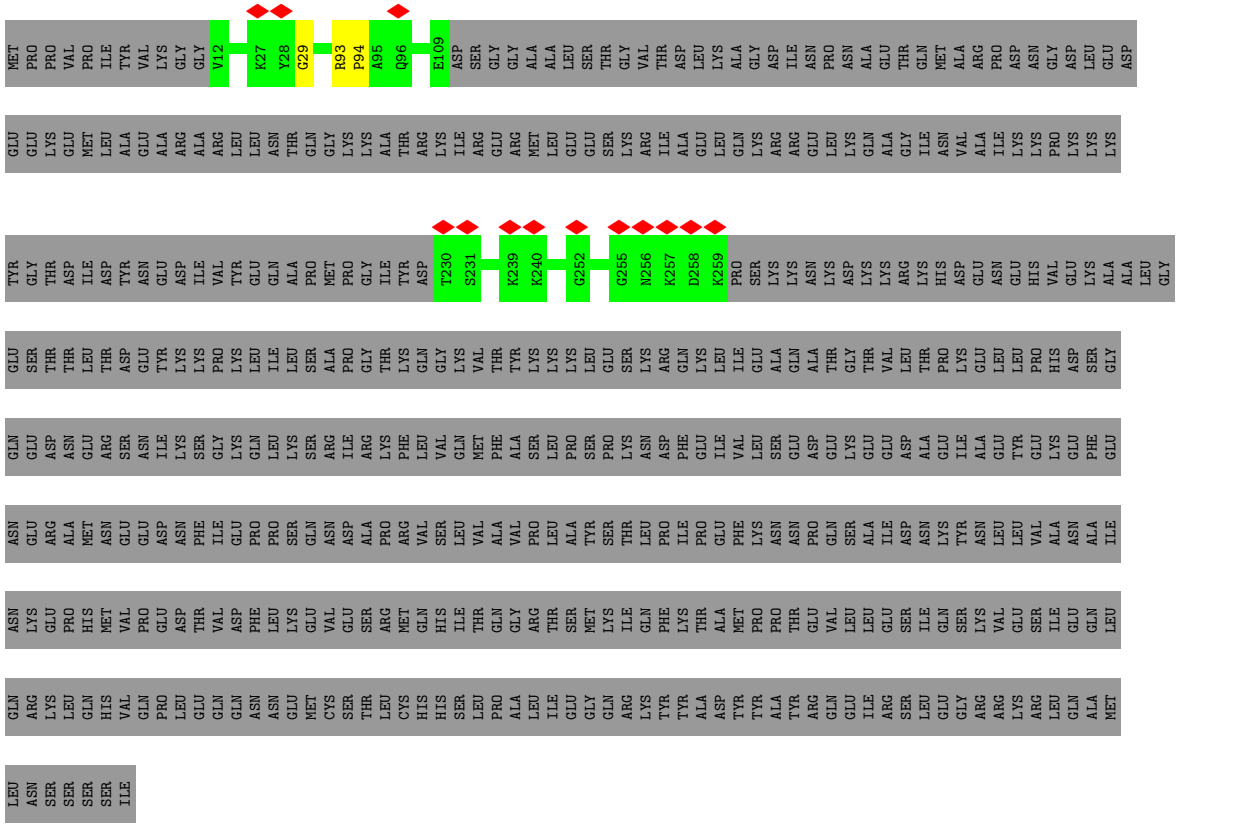
• Molecule 16: Pre-mRNA-splicing factor CLF1

Chain R: 32% 68%



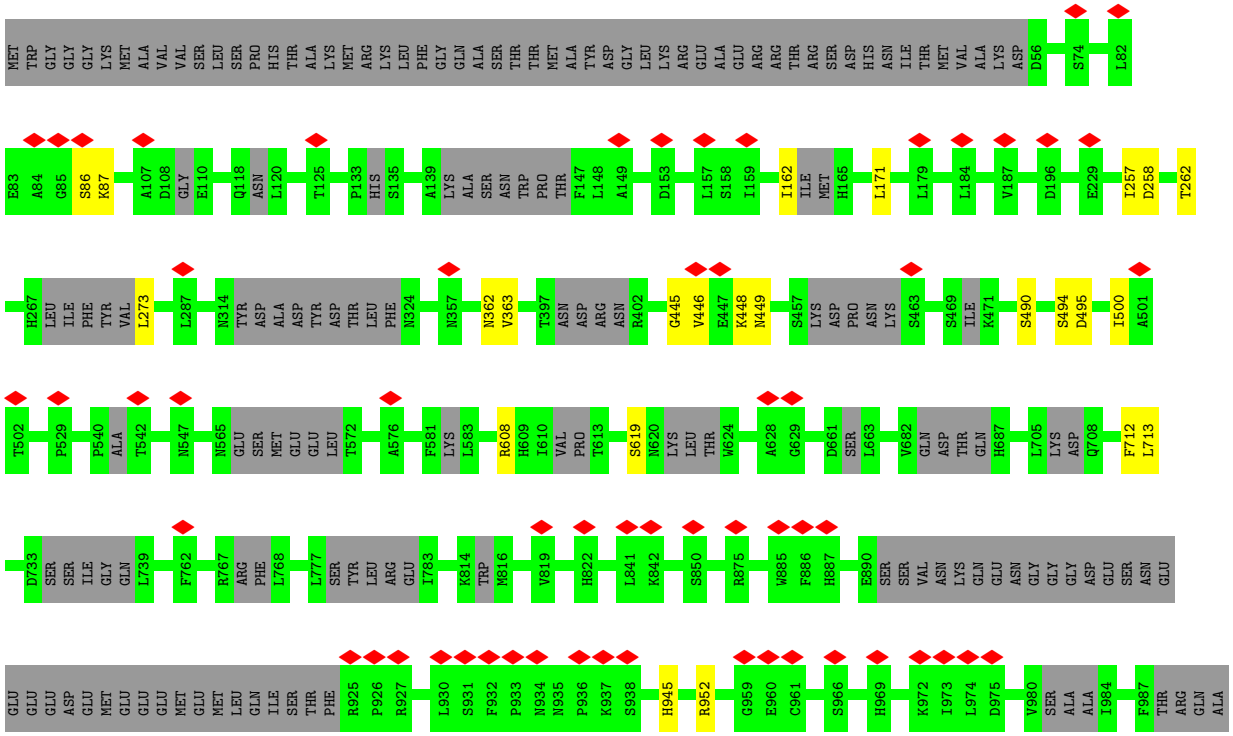
• Molecule 17: Pre-mRNA-splicing factor CEF1

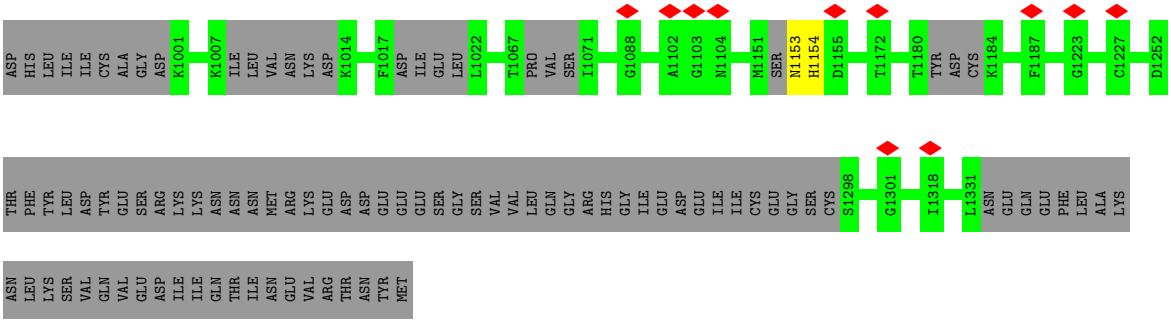
Chain W: 21% 78%



- Molecule 18: Pre-mRNA-splicing factor RSE1

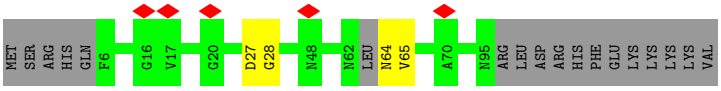
Chain X:





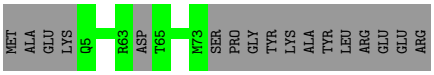
- Molecule 19: Pre-mRNA-splicing factor RDS3

Chain Y:



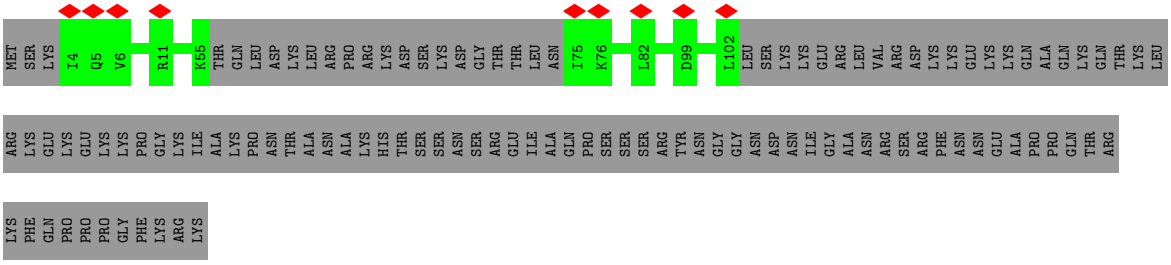
- Molecule 20: RDS3 complex subunit 10

Chain Z:



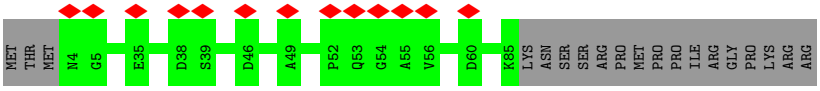
- Molecule 21: Small nuclear ribonucleoprotein-associated protein B

Chain b:



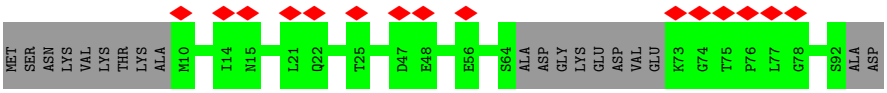
- Molecule 22: Small nuclear ribonucleoprotein Sm D3

Chain d:



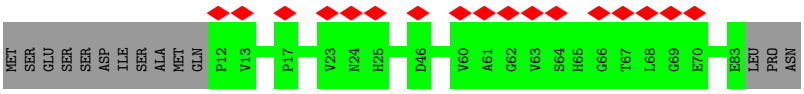
- Molecule 23: Small nuclear ribonucleoprotein E

Chain e:



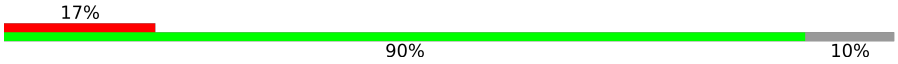
• Molecule 24: Small nuclear ribonucleoprotein F

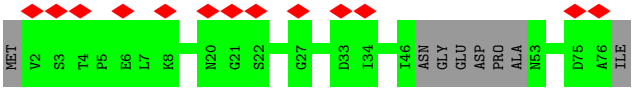
Chain f:



• Molecule 25: Small nuclear ribonucleoprotein G

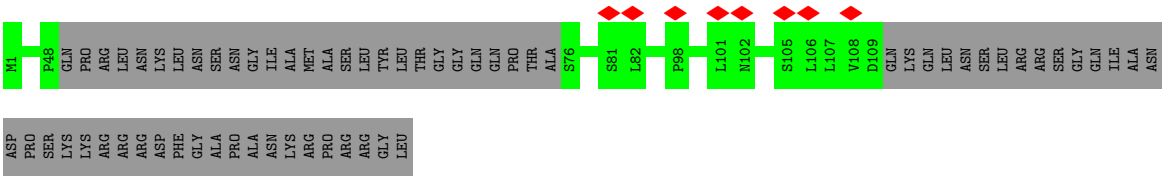
Chain g:





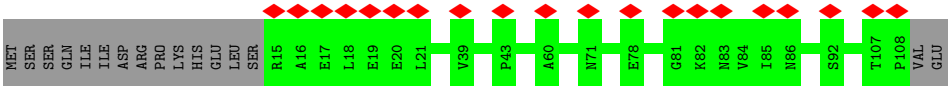
• Molecule 26: Small nuclear ribonucleoprotein Sm D1

Chain h:



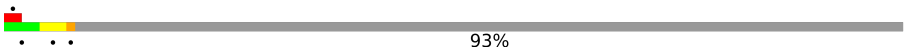
• Molecule 27: Small nuclear ribonucleoprotein Sm D2

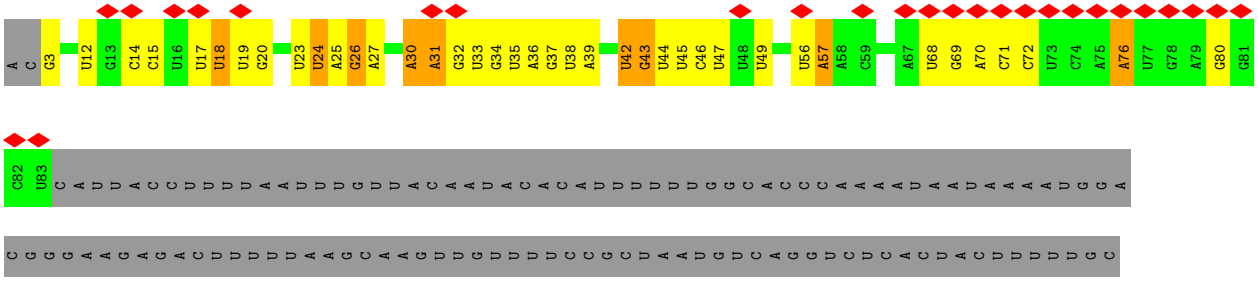
Chain j:

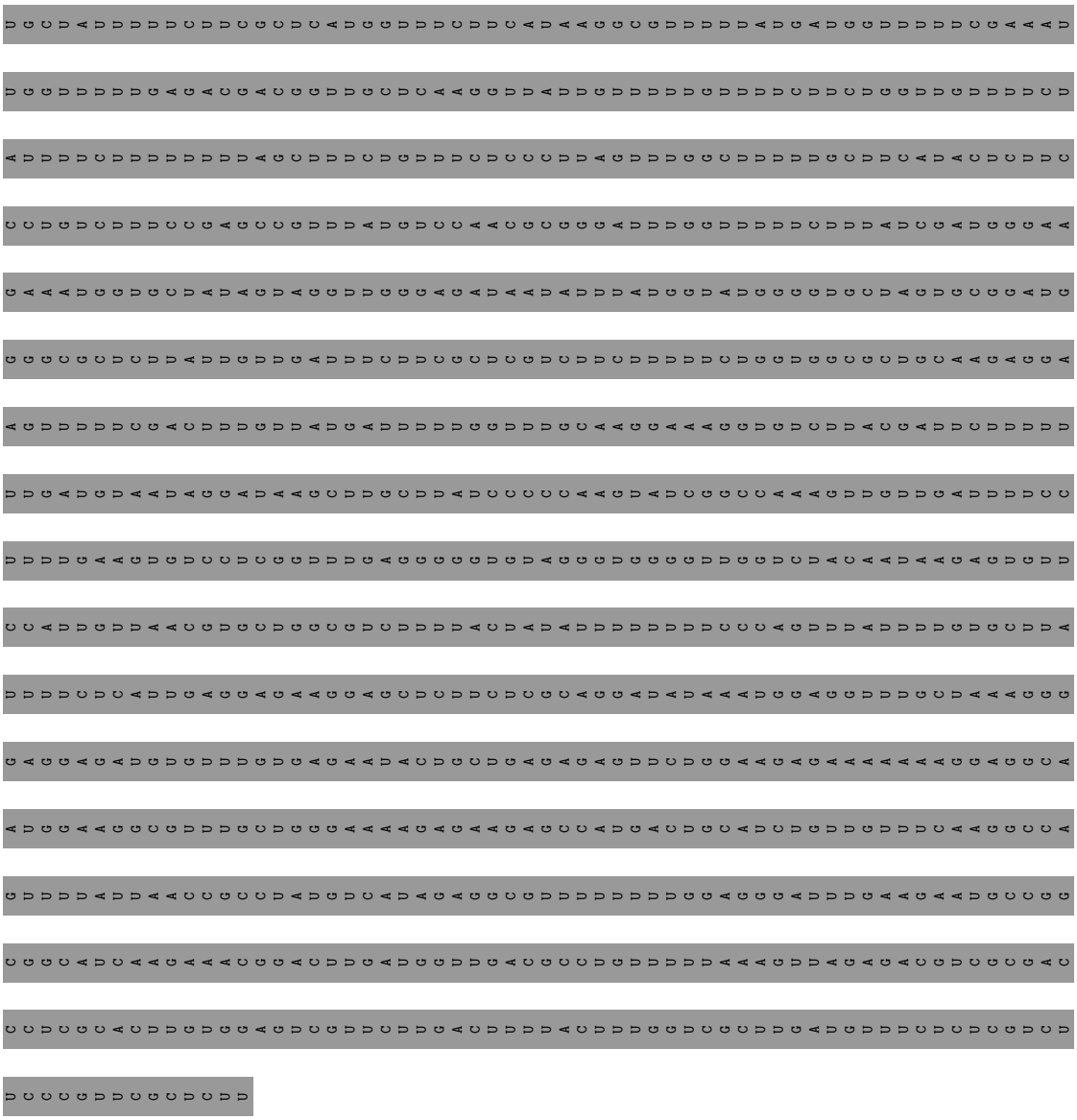


• Molecule 28: U2 snRNA

Chain 2:

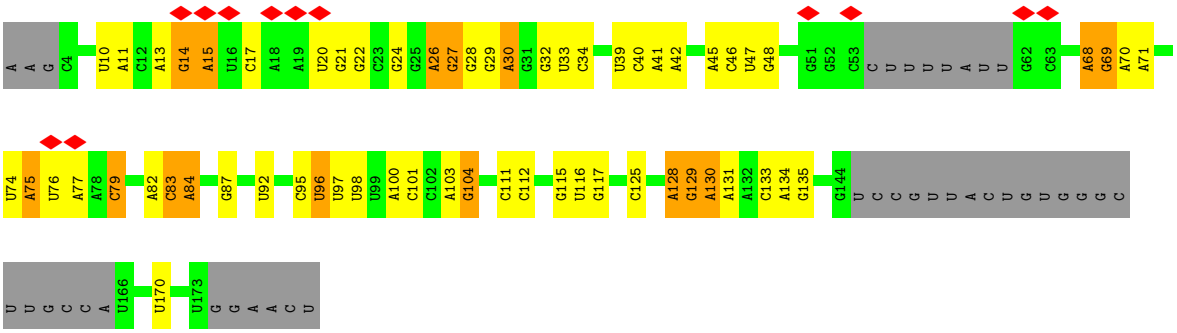






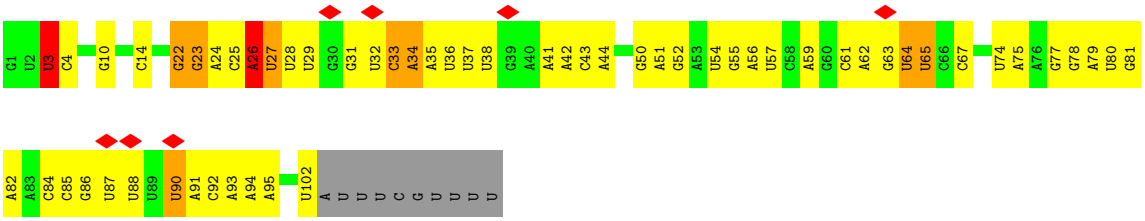
- Molecule 29: U5 snRNA

Chain 5:



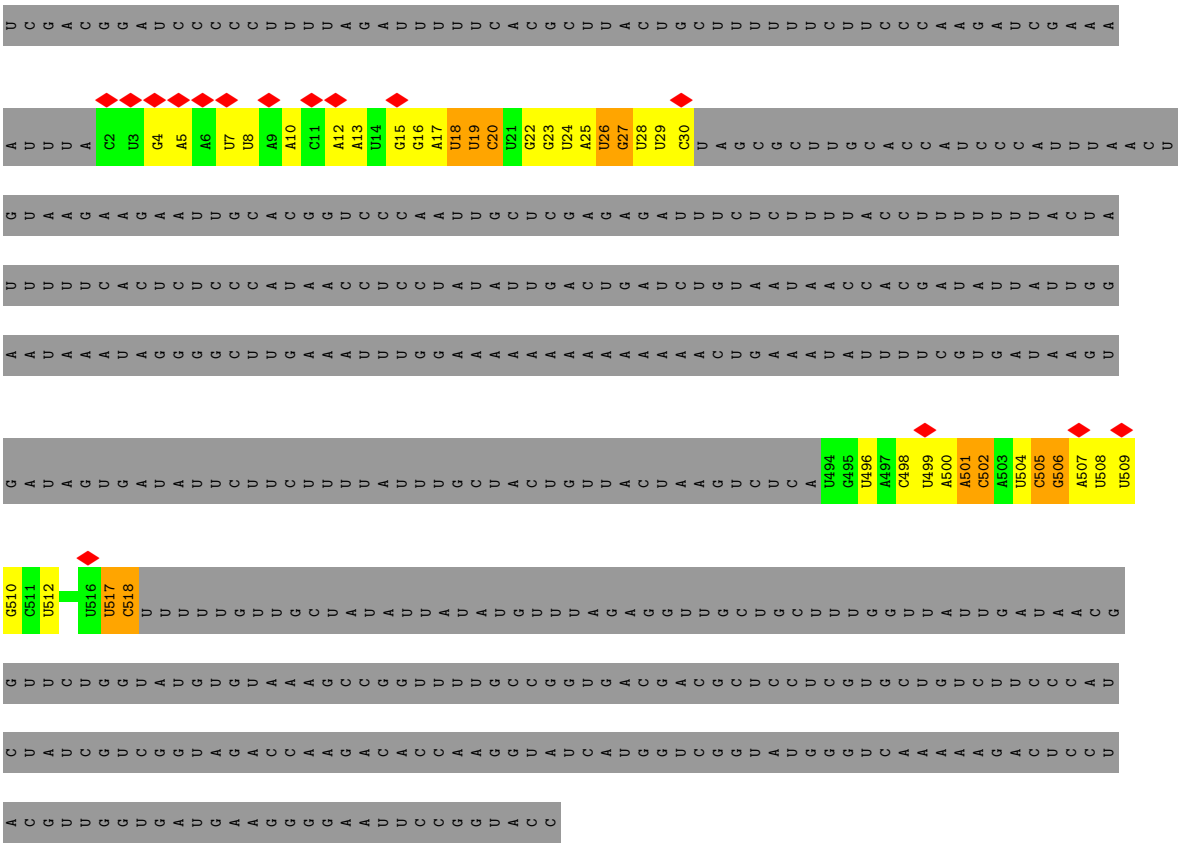
- Molecule 30: U6 snRNA

Chain 6:



- Molecule 31: actin pre-mRNA

Chain 9:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	122000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	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	74000	Depositor
Image detector	FEI FALCON II (4k x 4k)	Depositor
Maximum map value	0.343	Depositor
Minimum map value	-0.079	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.019	Depositor
Recommended contour level	0.11	Depositor
Map size (Å)	512.0, 512.0, 512.0	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.0, 2.0, 2.0	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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	$\#\|Z\|>5$
28	2	0.13	0/1902	0.79	2/2956 (0.1%)
29	5	0.12	0/3350	0.68	0/5209
30	6	0.16	0/2427	0.76	4/3778 (0.1%)
31	9	0.17	0/1264	0.65	0/1961
All	All	0.14	0/8943	0.72	6/13904 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
28	2	76	A	C2-N3-C4	15.59	118.39	110.60
28	2	76	A	N1-C2-N3	10.07	134.34	129.30
30	6	3	U	C2-N1-C1'	7.71	126.95	117.70
30	6	3	U	N1-C2-O2	7.28	127.90	122.80
30	6	3	U	N3-C2-O2	-6.56	117.61	122.20

There are no chirality outliers.

There are no planarity outliers.

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	A	2130	0	0	48	0
2	B	843	0	0	8	0
3	C	1811	0	0	5	0
4	D	187	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	138	0	0	17	0
6	F	218	0	0	22	0
7	H	437	0	0	3	0
8	J	104	0	0	1	0
9	K	338	0	0	4	0
10	L	32	0	0	0	0
11	M	169	0	0	10	0
12	N	159	0	0	1	0
13	O	628	0	0	19	0
14	P	202	0	0	3	0
15	Q	791	0	0	13	0
16	R	219	0	0	4	0
17	W	128	0	0	3	0
18	X	1095	0	0	13	0
19	Y	89	0	0	7	0
20	Z	68	0	0	0	0
21	b	80	0	0	0	0
22	d	82	0	0	0	0
23	e	75	0	0	0	0
24	f	72	0	0	0	0
25	g	69	0	0	0	0
26	h	82	0	0	0	0
27	j	94	0	0	0	0
28	2	1707	0	860	31	0
29	5	2999	0	1515	50	0
30	6	2170	0	1094	53	0
31	9	1135	0	575	32	0
All	All	18351	0	4044	247	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 11.

The worst 5 of 247 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:A:1377:SER:CA	31:9:18:U:C5	1.87	1.56
5:E:98:THR:CA	30:6:4:C:C2	1.88	1.52
15:Q:927:ALA:CA	28:2:57:A:H1'	1.05	1.51
6:F:72:PHE:CA	30:6:34:A:C2	1.94	1.50
6:F:72:PHE:CA	30:6:34:A:H2	1.20	1.49

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
28	2	80/1175 (6%)	21 (26%)	2 (2%)
29	5	137/179 (76%)	31 (22%)	4 (2%)
30	6	101/112 (90%)	41 (40%)	6 (5%)
31	9	52/572 (9%)	27 (51%)	1 (1%)
All	All	370/2038 (18%)	120 (32%)	13 (3%)

5 of 120 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
28	2	12	U
28	2	14	C
28	2	15	C
28	2	18	U
28	2	20	G

5 of 13 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
30	6	26	A
30	6	42	A
31	9	505	C
30	6	64	U
30	6	86	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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
29	5	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	5	166:U	O3'	167:A	P	3.66

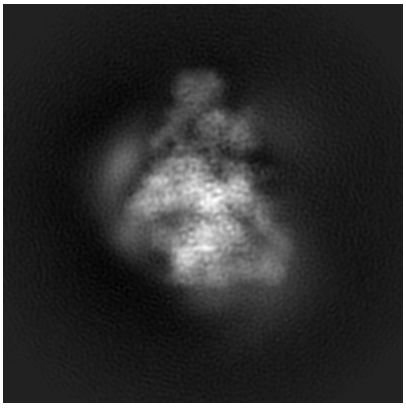
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-4099. These allow visual inspection of the internal detail of the map and identification of artifacts.

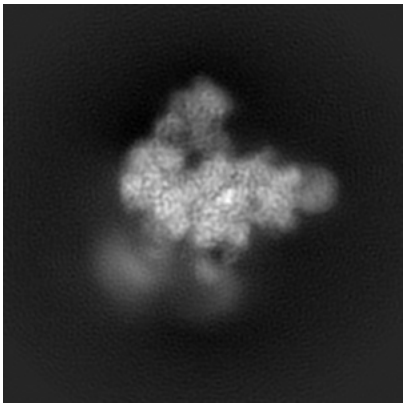
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

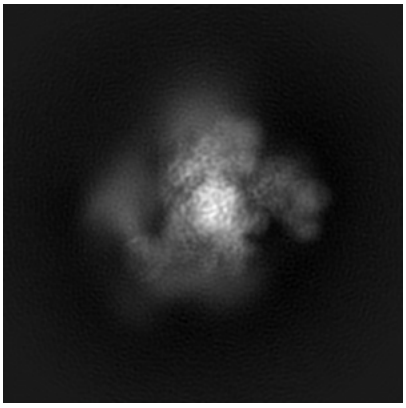
6.1.1 Primary map



X



Y

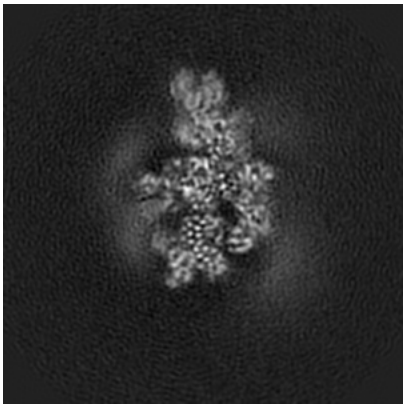


Z

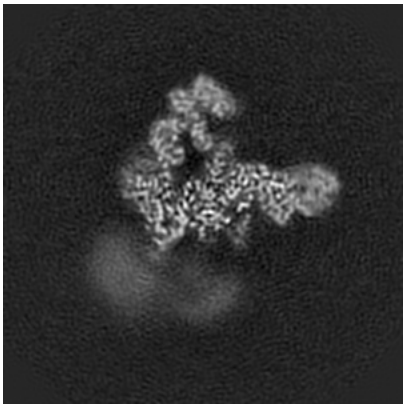
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

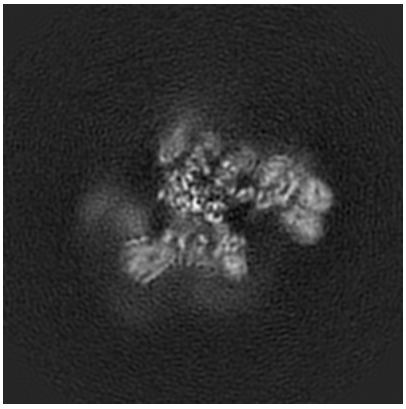
6.2.1 Primary map



X Index: 128



Y Index: 128

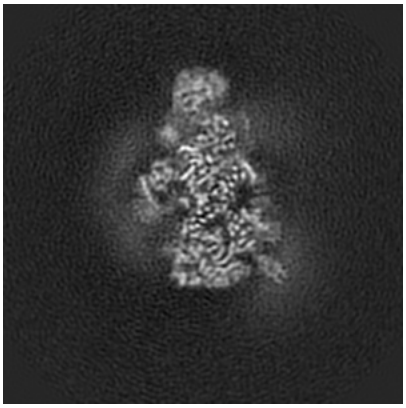


Z Index: 128

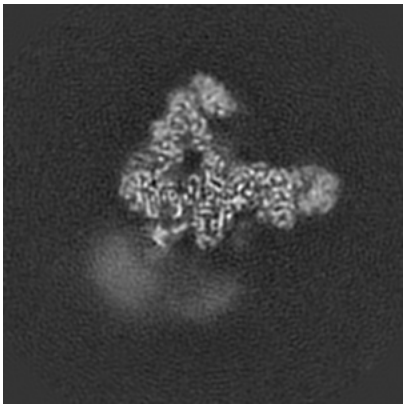
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

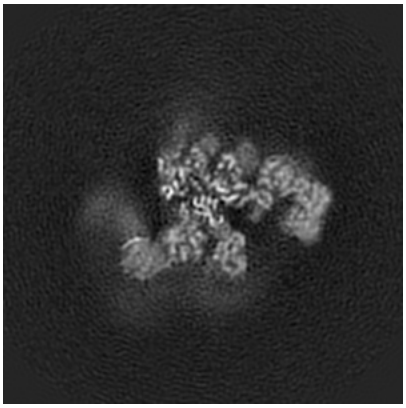
6.3.1 Primary map



X Index: 135



Y Index: 131

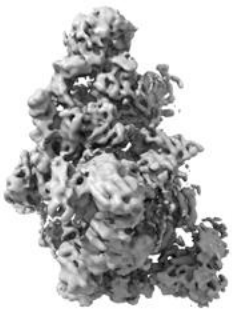


Z Index: 131

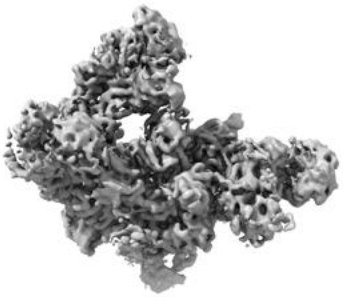
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

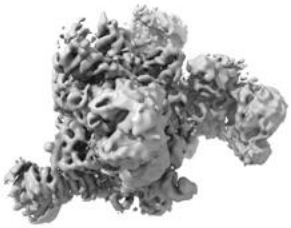
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.11. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

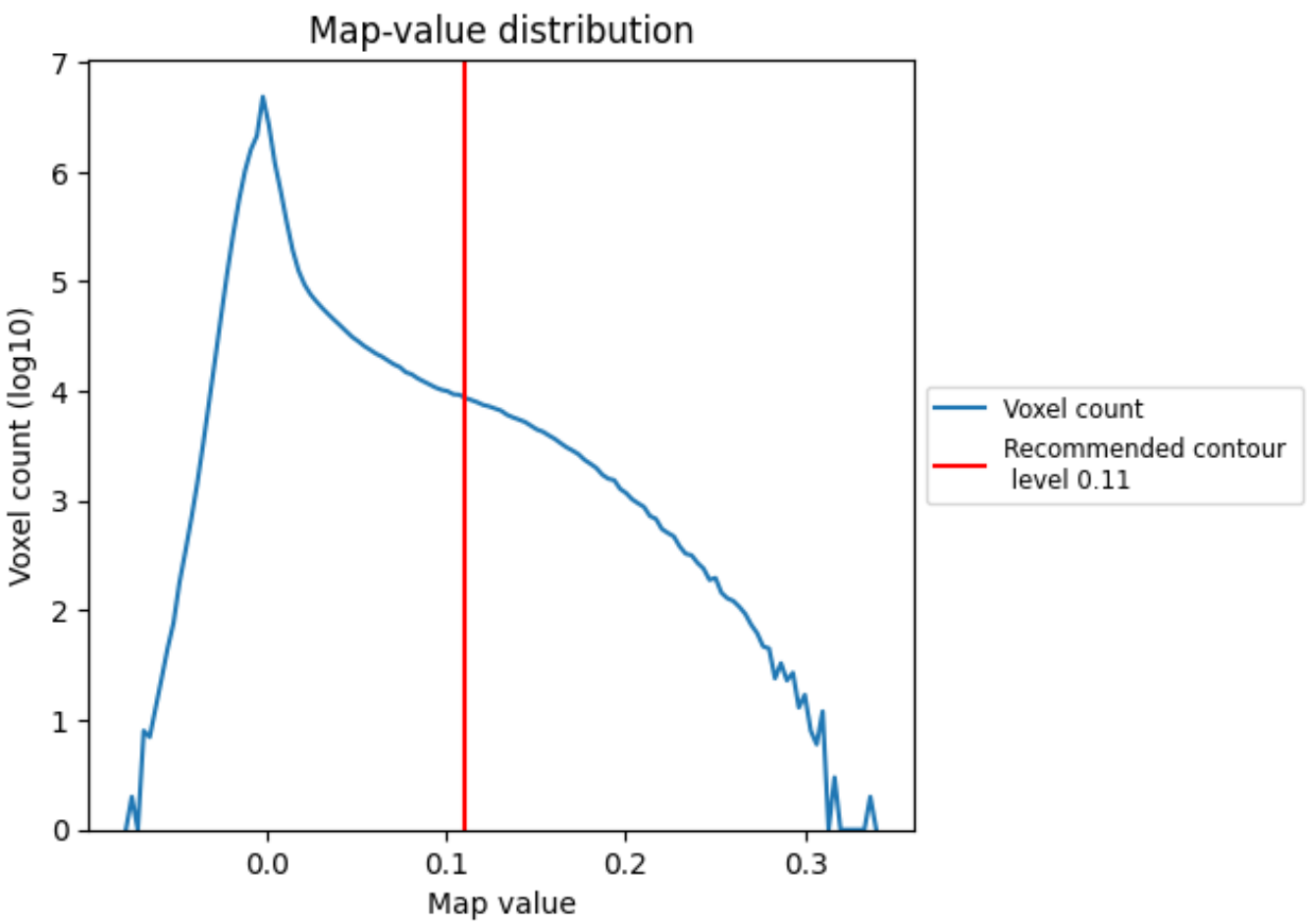
6.5 Mask visualisation (i)

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

7 Map analysis (i)

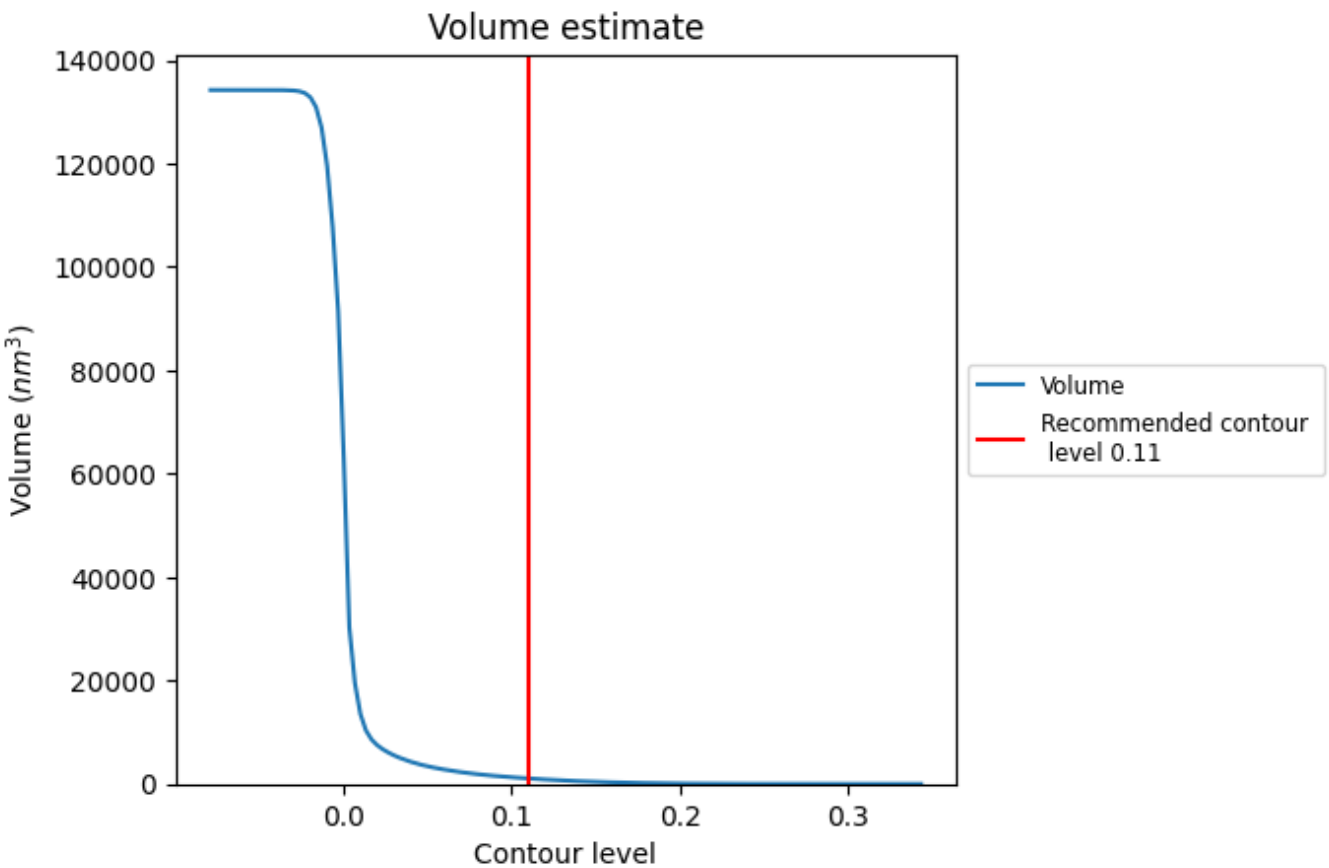
This section contains the results of statistical analysis of the map.

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.

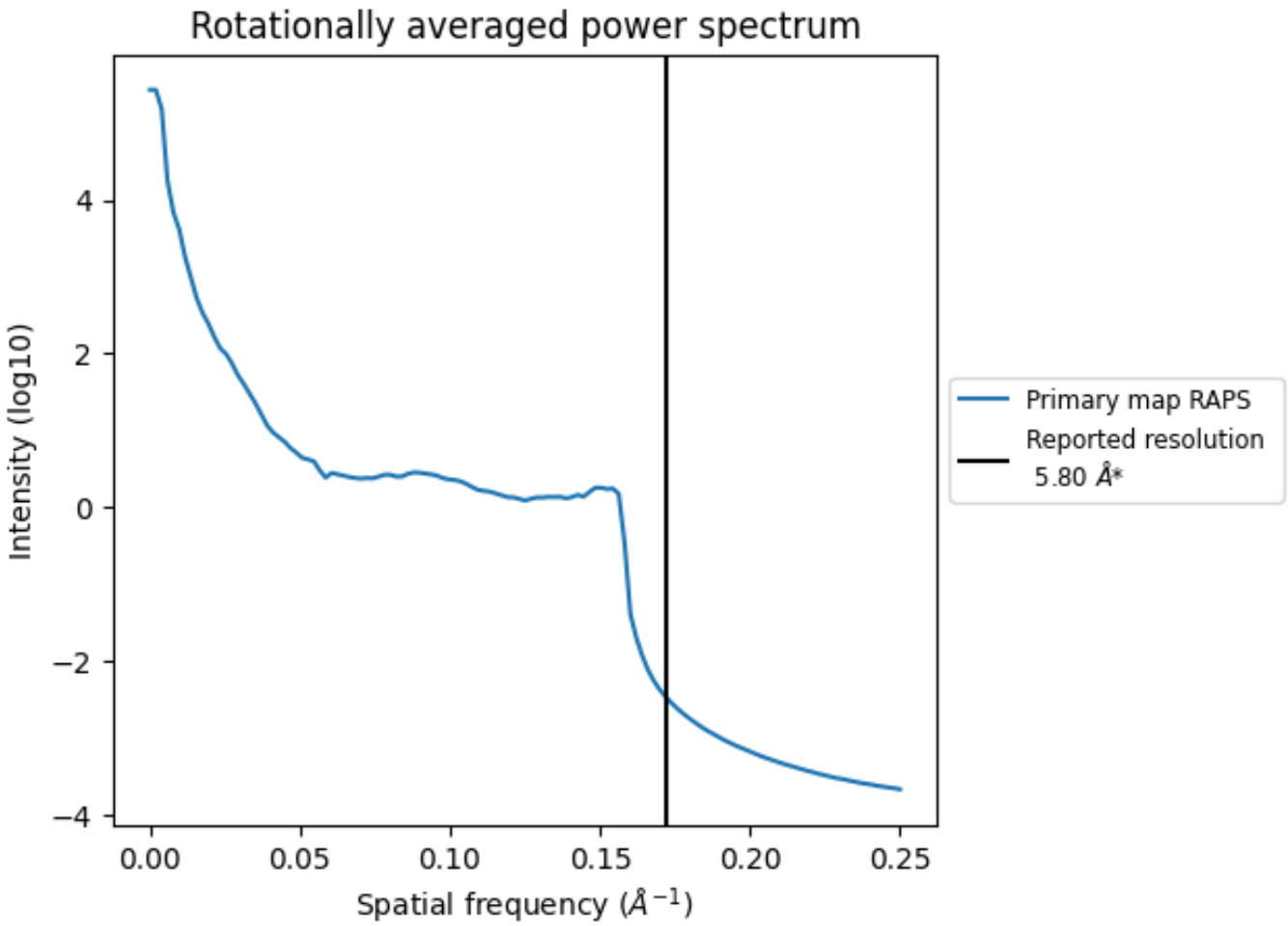
7.2 Volume estimate (i)



The volume at the recommended contour level is 1066 nm^3; this corresponds to an approximate mass of 963 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 0.172 Å^{-1}

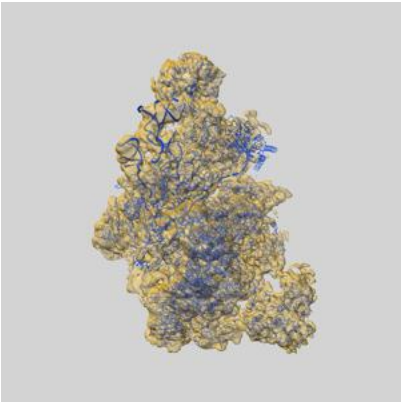
8 Fourier-Shell correlation (i)

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

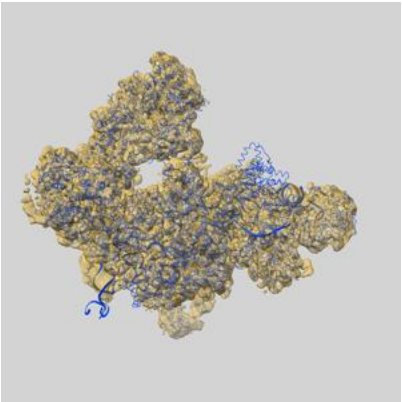
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-4099 and PDB model 5LQW. Per-residue inclusion information can be found in section 3 on page 9.

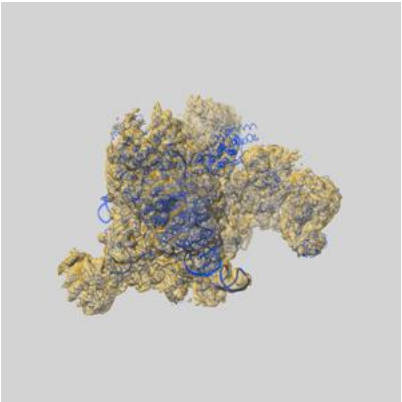
9.1 Map-model overlay (i)



X



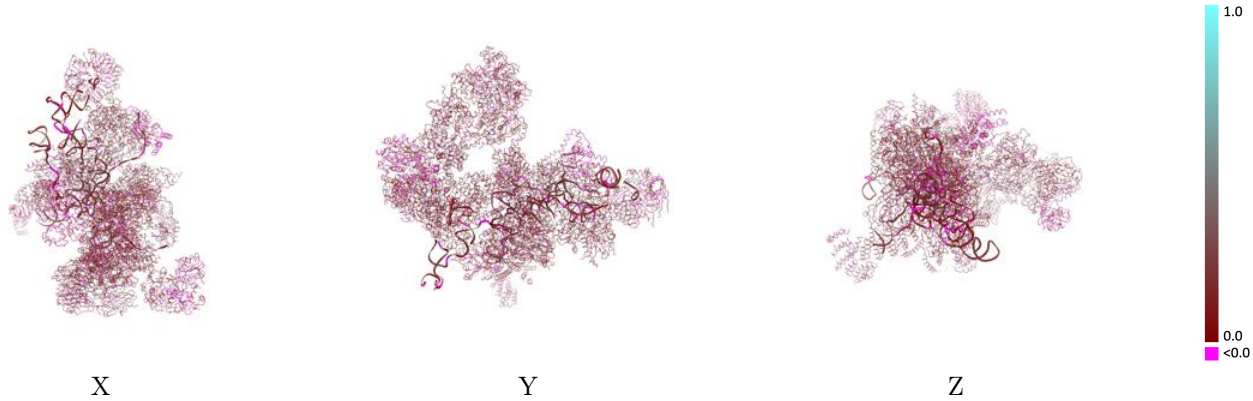
Y



Z

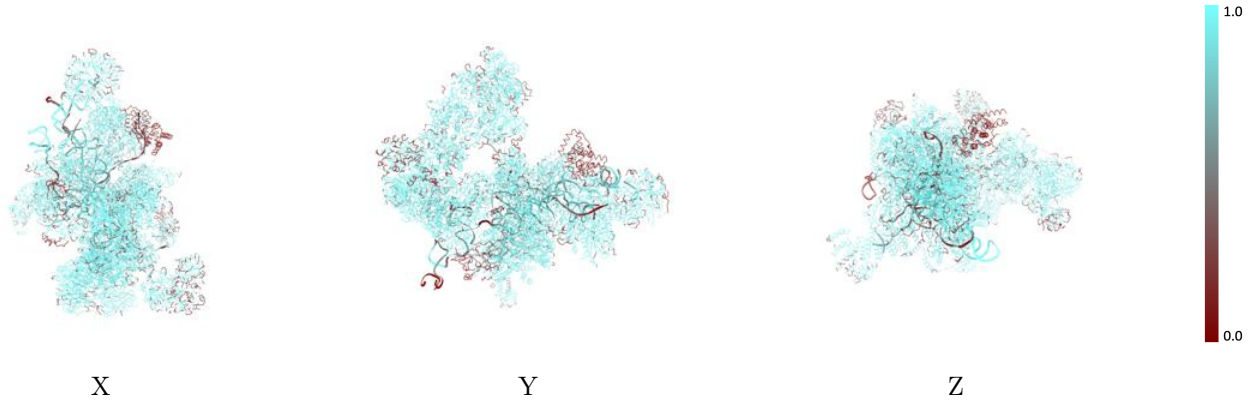
The images above show the 3D surface view of the map at the recommended contour level 0.11 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



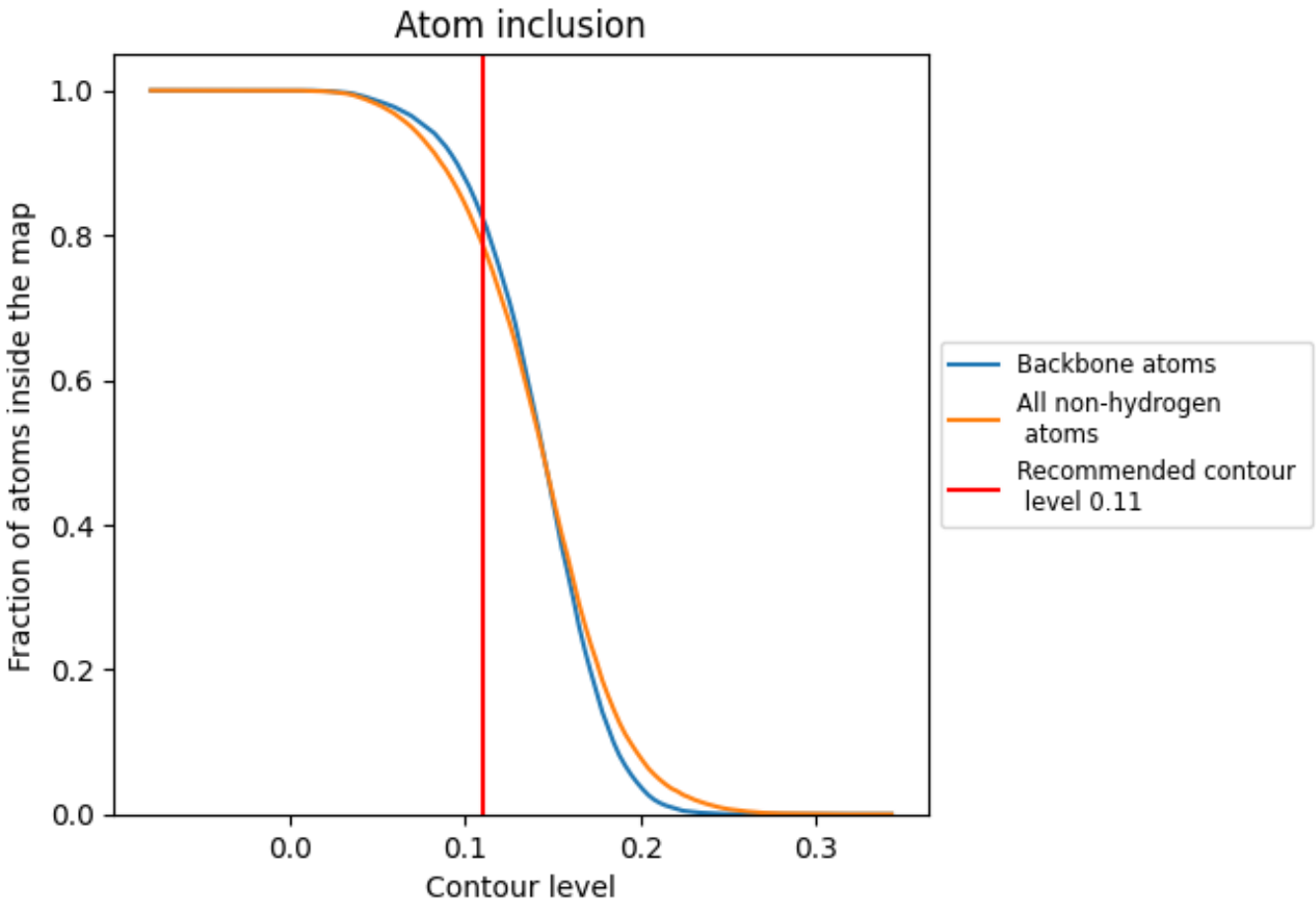
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.11).

9.4 Atom inclusion (i)



At the recommended contour level, 83% of all backbone atoms, 79% 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 (0.11) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7884	0.1440
2	0.5237	0.1370
5	0.7706	0.1210
6	0.7700	0.1500
9	0.5894	0.1550
A	0.9131	0.1640
B	0.9454	0.1700
C	0.8774	0.1370
D	0.8770	0.1460
E	0.8913	0.1170
F	0.9083	0.1660
H	0.4279	0.0900
J	0.9231	0.1610
K	0.9438	0.1790
L	0.6875	0.1720
M	0.7692	0.2040
N	0.3145	0.1200
O	0.7182	0.0910
P	0.7921	0.1200
Q	0.9267	0.1570
R	0.9498	0.1360
W	0.8984	0.1540
X	0.9361	0.1760
Y	0.9438	0.2000
Z	1.0000	0.1760
b	0.8875	0.1220
d	0.8415	0.1380
e	0.8000	0.0960
f	0.7639	0.0760
g	0.8116	0.1070
h	0.9024	0.1480
j	0.7872	0.1020