wwPDB EM Validation Summary Report (i)

Apr 24, 2023 – 02:30 PM JST

PDB ID	:	7VPX
EMDB ID	:	EMD-32074
Title	:	The cryo-EM structure of the human pre-A complex
Authors	:	Zhang, X.; Zhan, X.; Shi, Y.
Deposited on	:	2021-10-18
Resolution	:	3.00 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:

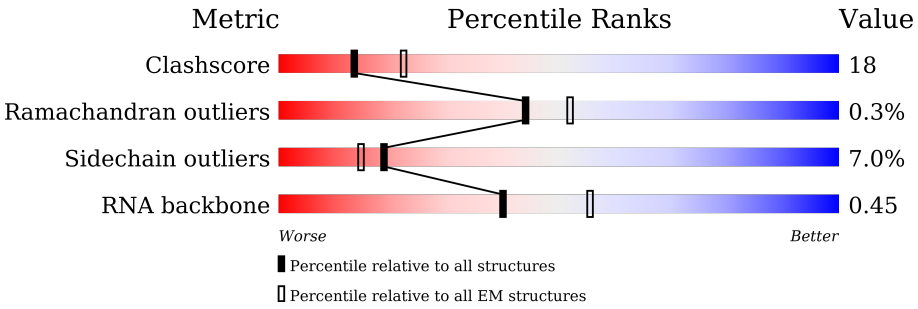
EMDB validation analysis	:	0.0.1.dev50
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.2

1 Overall quality at a glance

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

The reported resolution of this entry is 3.00 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=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	B	464	23% 28% 68%
2	D	639	14% 19% 81%
3	J	253	5% 17% 79%
4	H	188	52% 43% 21% 5% 31%
5	1	1304	40% 20% 38%
6	2	895	7% 15% 6% 79%
7	3	1217	6% 72% 24%

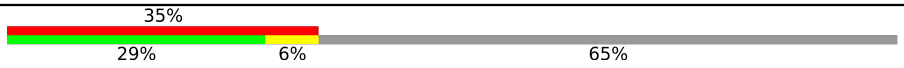
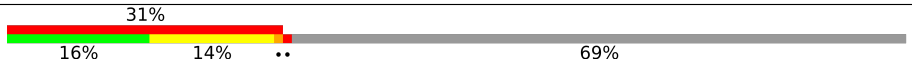
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	4	424	30% 38% 62%
9	5	86	57% 20% 23%
10	A	793	16% 16% 84%
11	C	501	64% 77% 7% 15%
12	E	1031	41% 41% 56%
13	F	255	64% 64% 36%
14	G	225	74% 73% 26%
15	a	118	76% 76% 24%
15	i	118	80% 70% 8% 20%
16	b	86	86% 86% 14%
16	m	86	86% 76% 9% 14%
17	c	92	86% 86% 14%
17	l	92	84% 67% 15% 16%
18	d	76	97% 97%
18	n	76	96% 86% 11%
19	e	126	66% 66% 34%
19	j	126	64% 56% 6% 36%
20	f	240	30% 30% 70%
20	k	240	36% 32% 64%
21	g	119	69% 69% 31%
21	h	119	68% 61% 5% 32%
22	I	259	94%
23	K	10	100% 30% 60% 10%
24	L	164	100% 40% 35% 23%
25	O	437	46% 31% 13% 54%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
26	M	282	
27	N	159	
28	6	110	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
30	SJT	6	204	X	-	-	-

2 Entry composition (i)

There are 30 unique types of molecules in this entry. The entry contains 43240 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 Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	B	147	799	455	173	168	3	0	0

- Molecule 2 is a protein called Splicing factor 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	D	122	604	360	122	122	0	0

- Molecule 3 is a protein called DnaJ homolog subfamily C member 8.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	J	54	369	224	73	72	0	0

- Molecule 4 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	H	130	2754	1233	478	915	128	0	0

- Molecule 5 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	1	815	6487	4163	1121	1164	39	0	0

- Molecule 6 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	2	187	1325	835	245	243	2	0	0

- Molecule 7 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	3	1180	Total	C	N	O	S	0	0
			9247	5872	1571	1759	45		

- Molecule 8 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	4	160	Total	C	N	O	0	0
			664	344	160	160		

- Molecule 9 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	5	66	Total	C	N	O	S	0	0
			539	342	93	99	5		

- Molecule 10 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	A	123	Total	C	N	O	0	0
			504	258	123	123		

- Molecule 11 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	C	425	Total	C	N	O	S	0	0
			2127	1191	470	462	4		

- Molecule 12 is a protein called Probable ATP-dependent RNA helicase DDX46.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	E	455	Total	C	N	O	S	0	0
			2153	1157	487	507	2		

- Molecule 13 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	F	162	Total	C	N	O	0	0
			666	342	162	162		

- Molecule 14 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms	AltConf	Trace
14	G	166	Total C N O 685 353 166 166	0	0

- Molecule 15 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms	AltConf	Trace
15	a	90	Total C N O 372 192 90 90	0	0
15	i	94	Total C N O S 764 478 139 141 6	0	0

- Molecule 16 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms	AltConf	Trace
16	b	74	Total C N O 308 160 74 74	0	0
16	m	74	Total C N O S 576 373 95 103 5	0	0

- Molecule 17 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms	AltConf	Trace
17	c	79	Total C N O 319 161 79 79	0	0
17	l	77	Total C N O S 638 405 113 115 5	0	0

- Molecule 18 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms	AltConf	Trace
18	d	74	Total C N O 305 157 74 74	0	0
18	n	73	Total C N O S 568 358 102 102 6	0	0

- Molecule 19 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms	AltConf	Trace
19	e	83	Total C N O 341 175 83 83	0	0
19	j	81	Total C N O S 637 400 112 119 6	0	0

- Molecule 20 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms	AltConf	Trace
20	f	71	Total C N O 293 151 71 71	0	0
20	k	86	Total C N O S 692 435 126 124 7	0	0

- Molecule 21 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms	AltConf	Trace
21	g	82	Total C N O 337 173 82 82	0	0
21	h	81	Total C N O S 641 408 112 118 3	0	0

- Molecule 22 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms	AltConf	Trace
22	I	15	Total C N O P 314 141 51 107 15	0	0

- Molecule 23 is a RNA chain called 5SS.

Mol	Chain	Residues	Atoms	AltConf	Trace
23	K	10	Total C N O P 216 97 42 68 9	0	0

- Molecule 24 is a RNA chain called U1 snRNA.

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

- Molecule 25 is a protein called U1 small nuclear ribonucleoprotein 70 kDa.

Mol	Chain	Residues	Atoms	AltConf	Trace
25	O	201	Total C N O S 1643 1030 317 291 5	0	0

- Molecule 26 is a protein called U1 small nuclear ribonucleoprotein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	M	98	Total	C	N	O	S	0	0
			787	506	134	143	4		

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

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

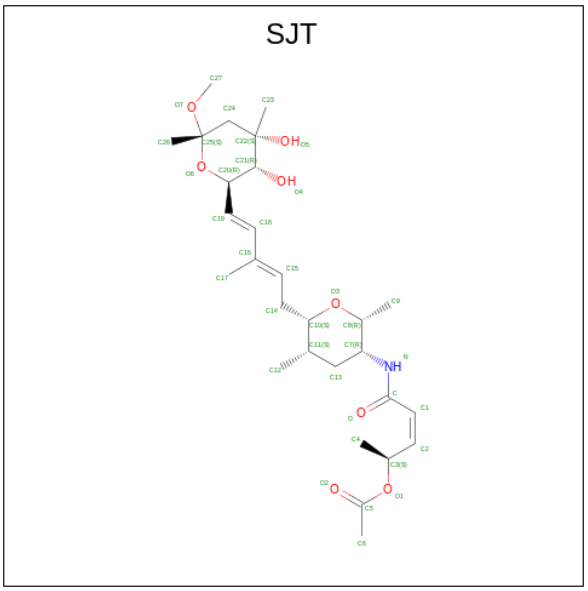
- Molecule 28 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	6	81	Total	C	N	O	S	0	0
			613	376	109	115	13		

- Molecule 29 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
29	B	1	Total	Zn	0
			1	1	
29	C	1	Total	Zn	0
			1	1	
29	N	1	Total	Zn	0
			1	1	
29	6	3	Total	Zn	0
			3	3	

- Molecule 30 is spliceostatin A (form II) (three-letter code: SJT) (formula: $C_{28}H_{45}NO_8$) (labeled as "Ligand of Interest" by depositor).



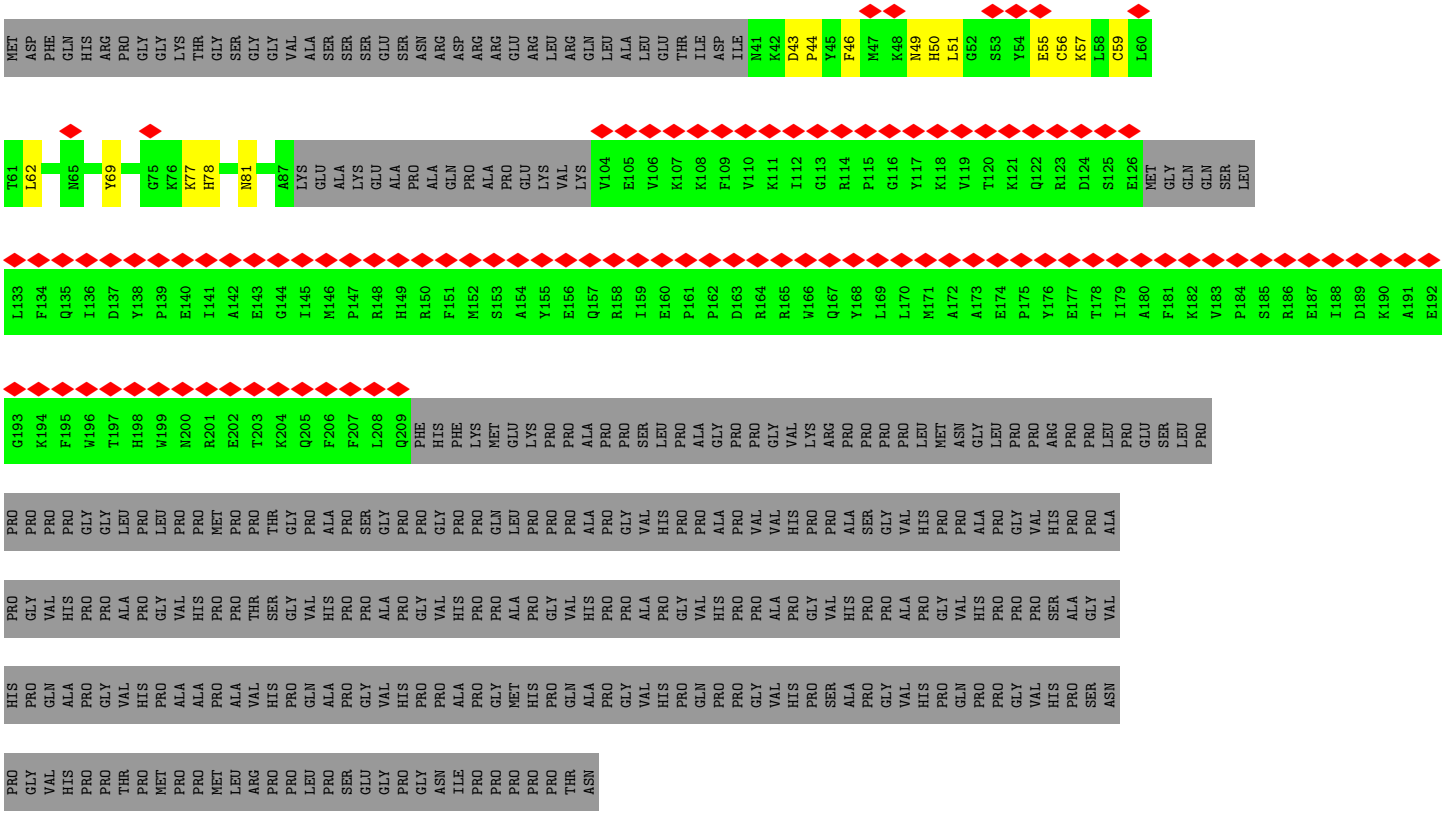
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
30	6	1	37	28	1	8	0

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.

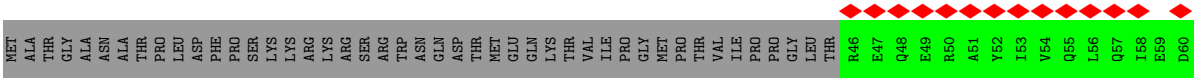
• Molecule 1: Splicing factor 3A subunit 2

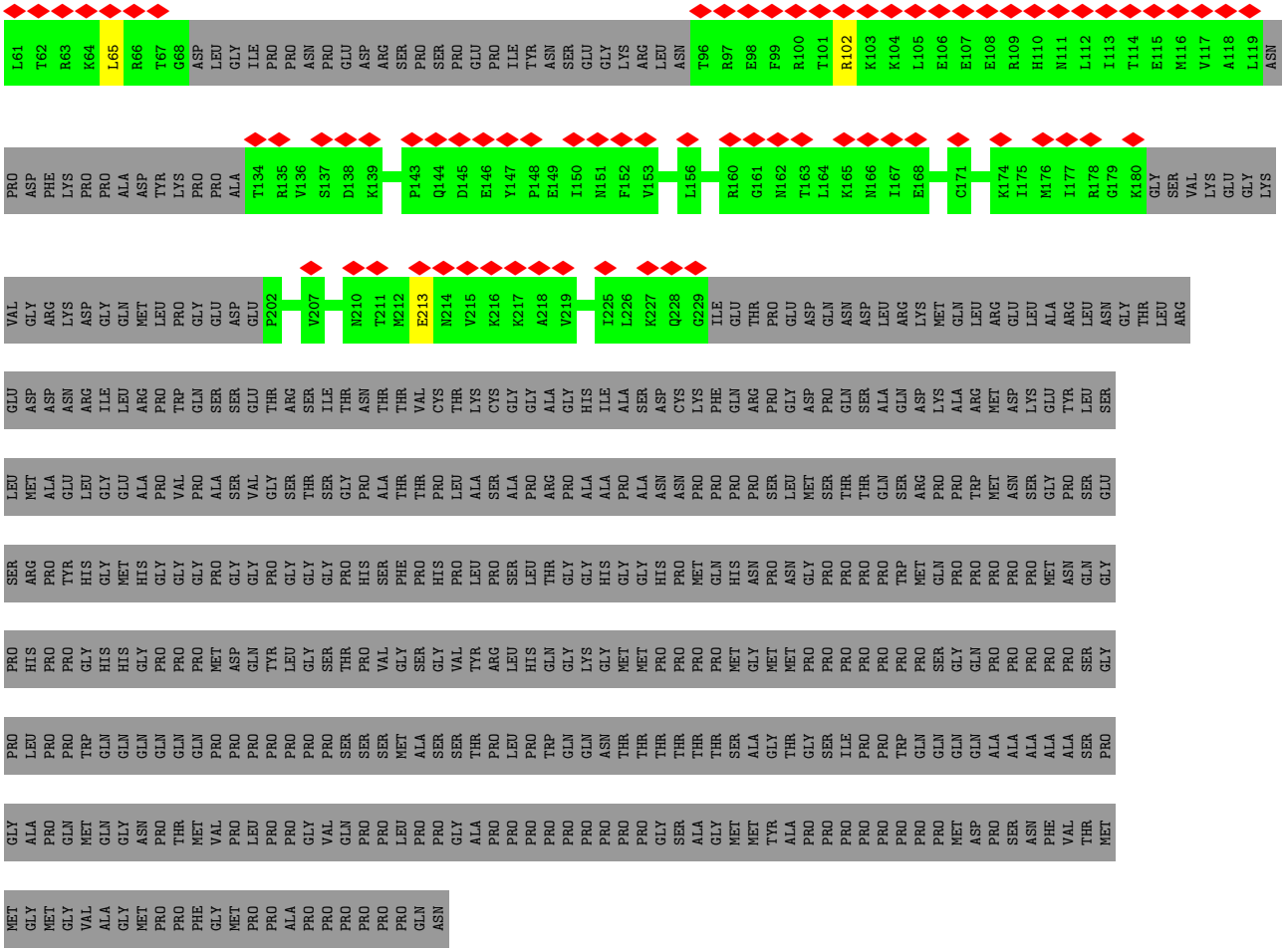
Chain B:



• Molecule 2: Splicing factor 1

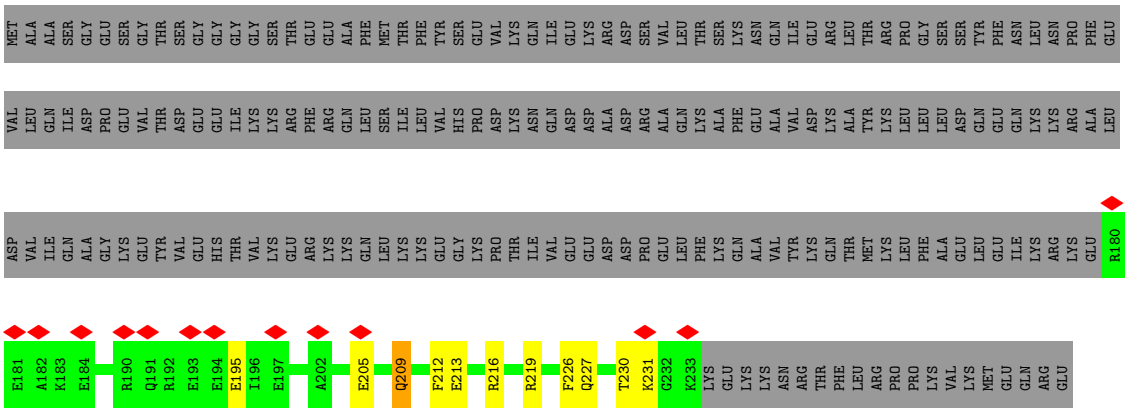
Chain D:





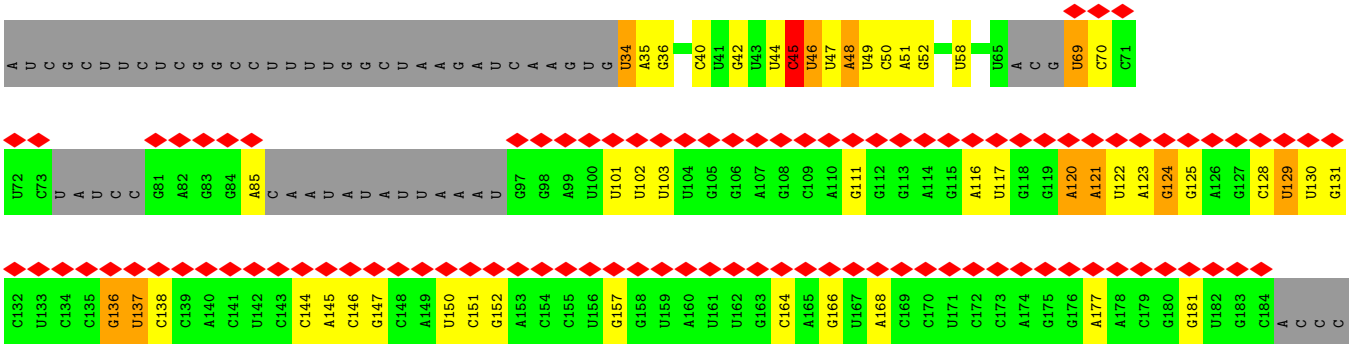
• Molecule 3: DnaJ homolog subfamily C member 8

Chain J:



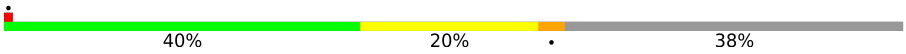
• Molecule 4: U2 snRNA

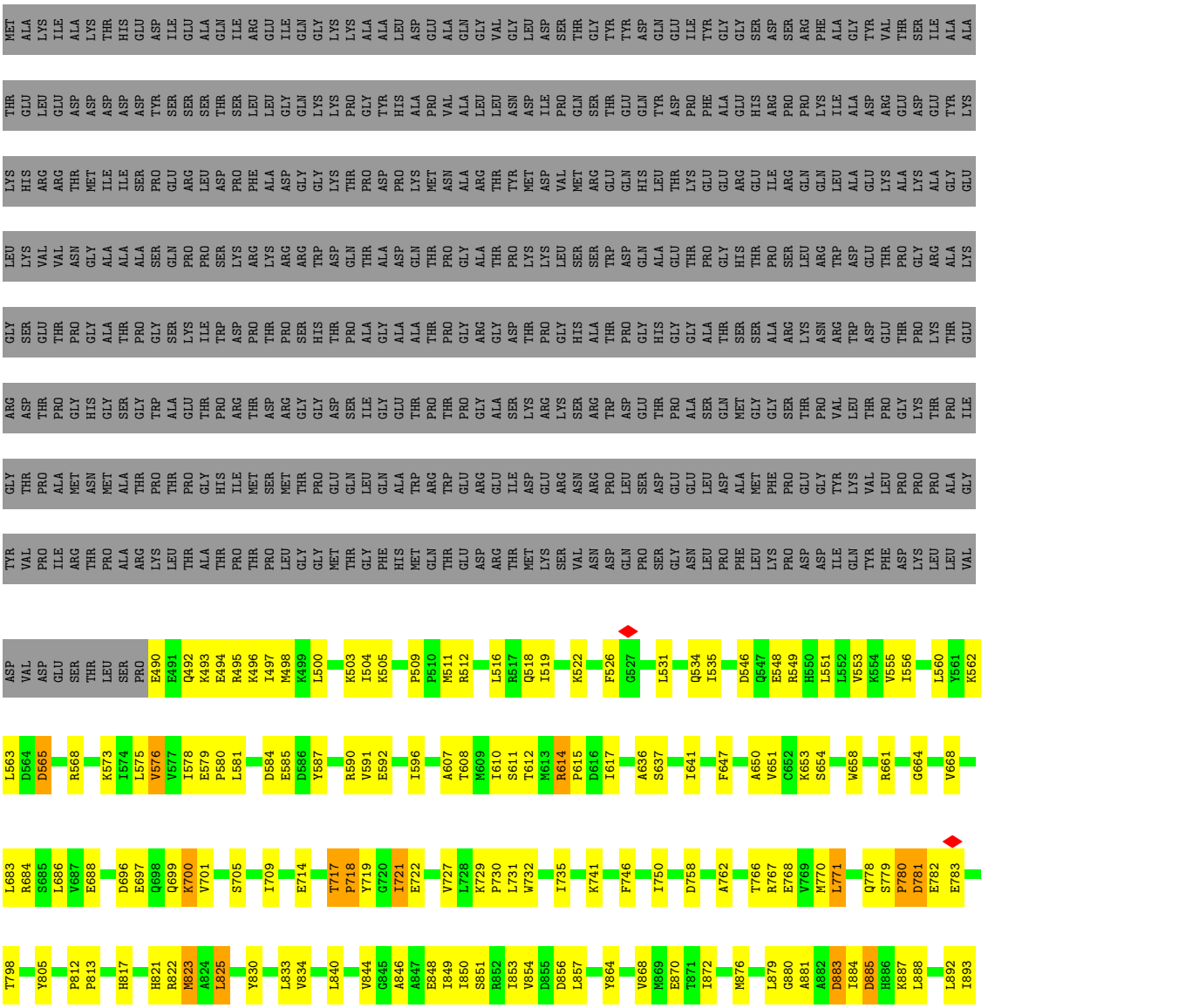
Chain H:

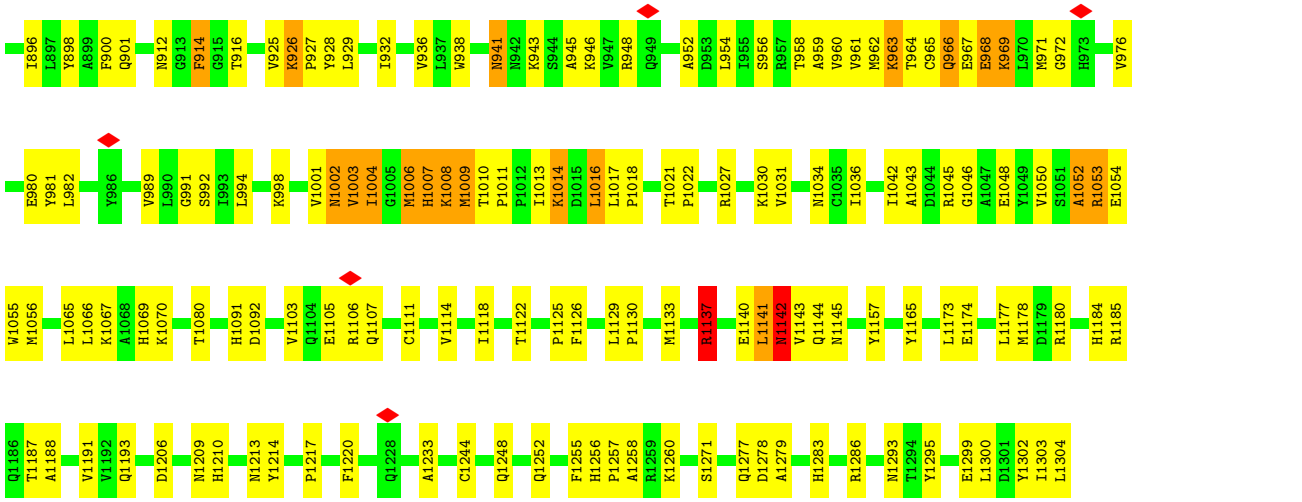


• Molecule 5: Splicing factor 3B subunit 1

Chain 1:

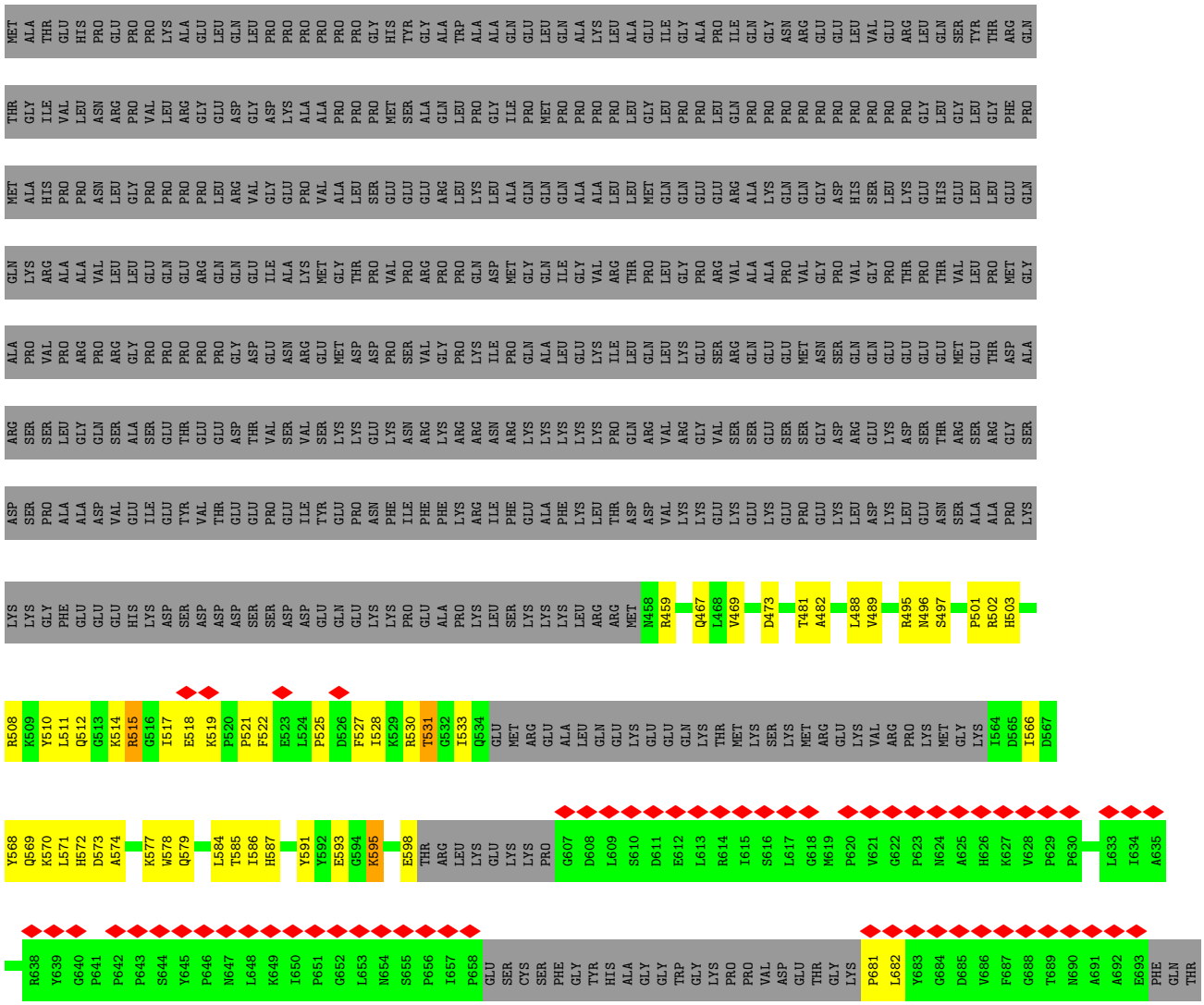


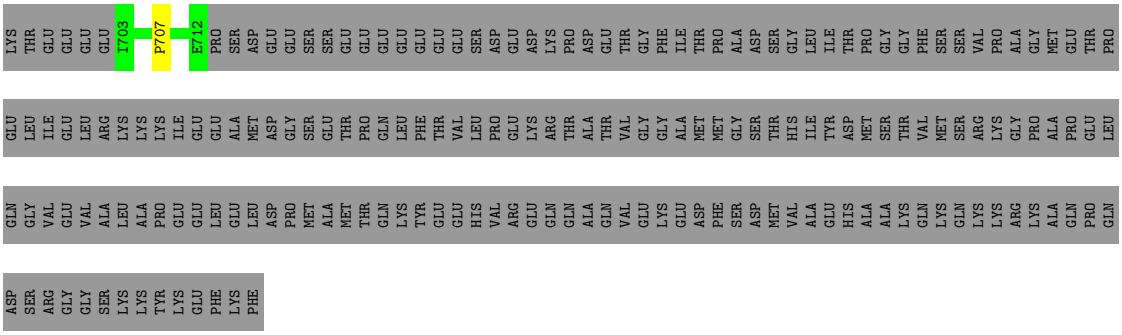




• Molecule 6: Splicing factor 3B subunit 2

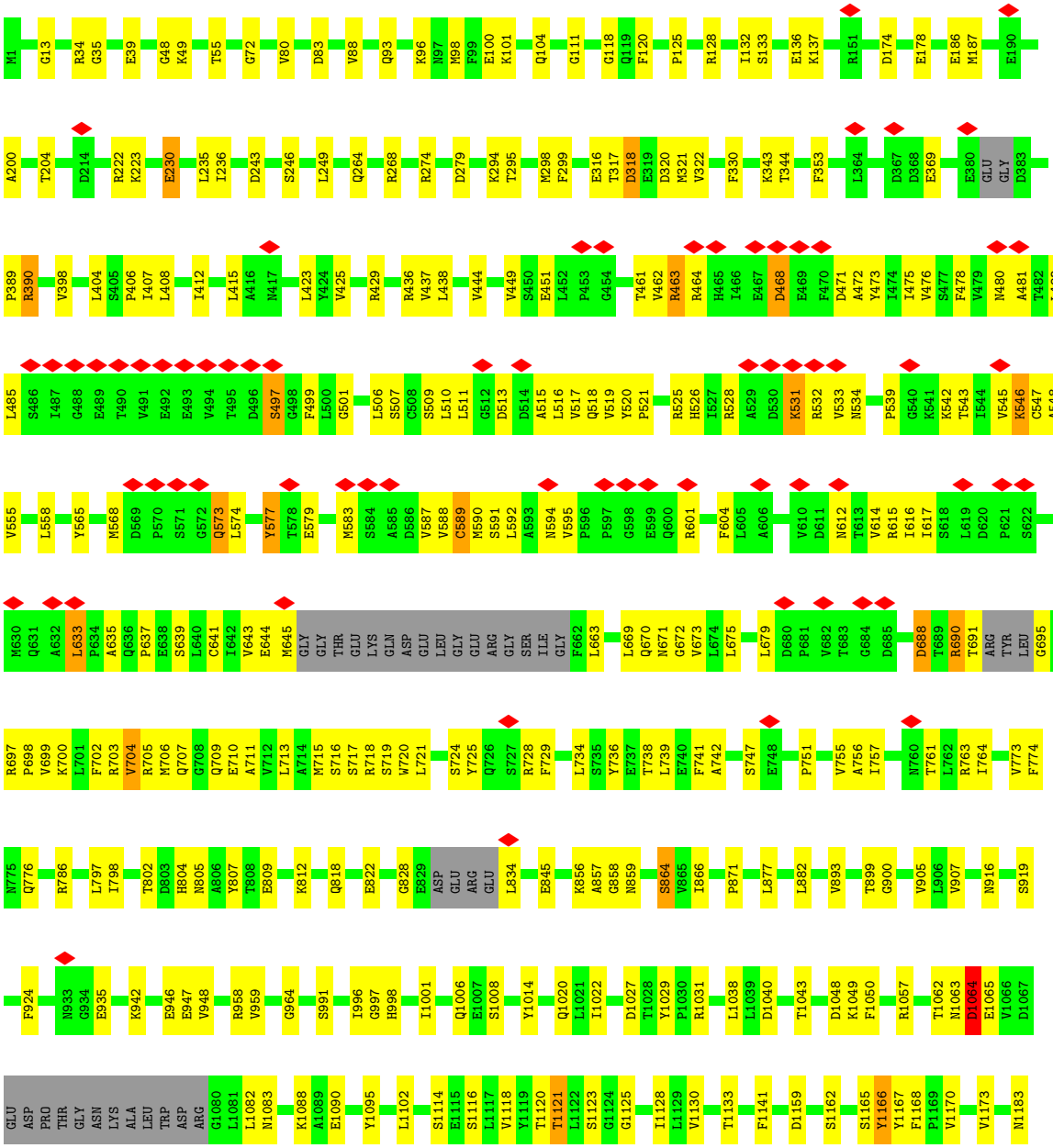
Chain 2:





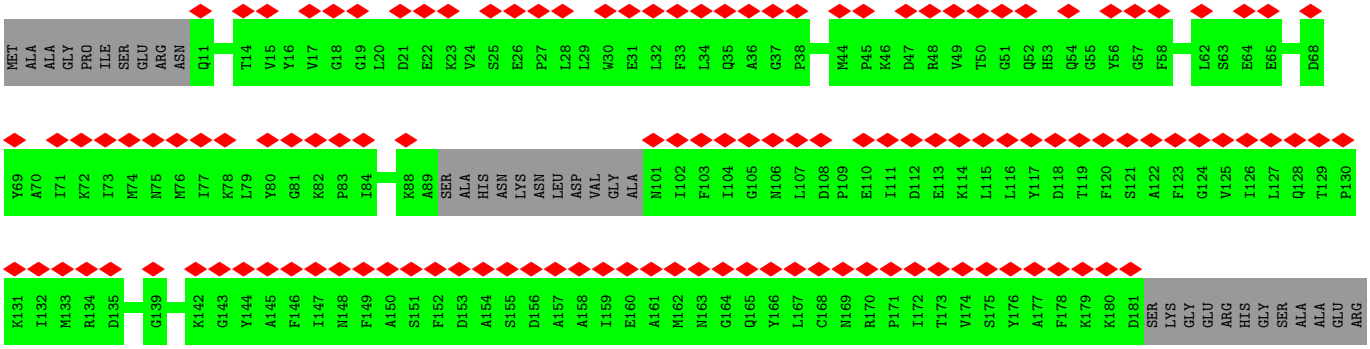
• Molecule 7: Splicing factor 3B subunit 3

Chain 3:



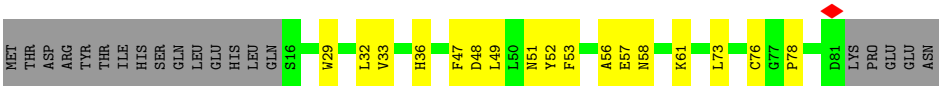
• Molecule 8: Splicing factor 3B subunit 4

Chain 4:



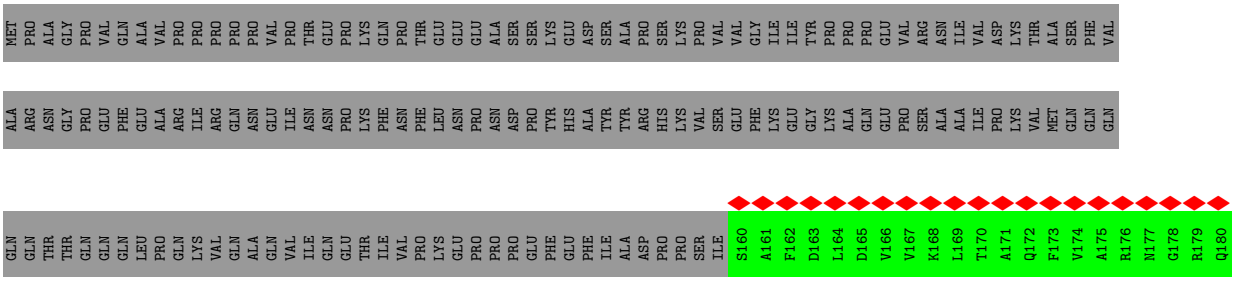
• Molecule 9: Splicing factor 3B subunit 5

Chain 5:



• Molecule 10: Splicing factor 3A subunit 1

Chain A:



F181 L182 T183 Q184 L185 M186 Q187 K188 E189 Q190 R191 N192 Y193 Q194 F195 D196 F197 L198 R199 P200 Q201 H202 S203 L204 F205 N206 Y207 F208 T209 K210 L211 V212 E213 Q214 Y215 T216 K217 I218 L219 I220 P221 P222 K223 G224 L225 F226 S227 K228 L229 K230 K231 E232 A233 E234 N235 P236 R237 E238 V239 L240

D241 Q242 V243 C244 Y245 R246 V247 E248 W249 A250 K251 F252 Q253 E254 R255 E256 R257 K258 K259 E260 E261 E262 E263 K264 E265 K266 E267 R268 V269 A270 Y271 A272 Q273 I274 D275 W276 H277 D278 F279 V280 V281 V282 GLU THR VAL ASP PHE GLN PRO ASN GLU GLN GLY ASN PHE PRO PRO PRO THR THR

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

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

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

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

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

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

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

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

ILE HIS LEU ALA LEU LYS GLU ARG GLY GLY ARG LYS LYS

• Molecule 11: Splicing factor 3A subunit 3

Chain C:

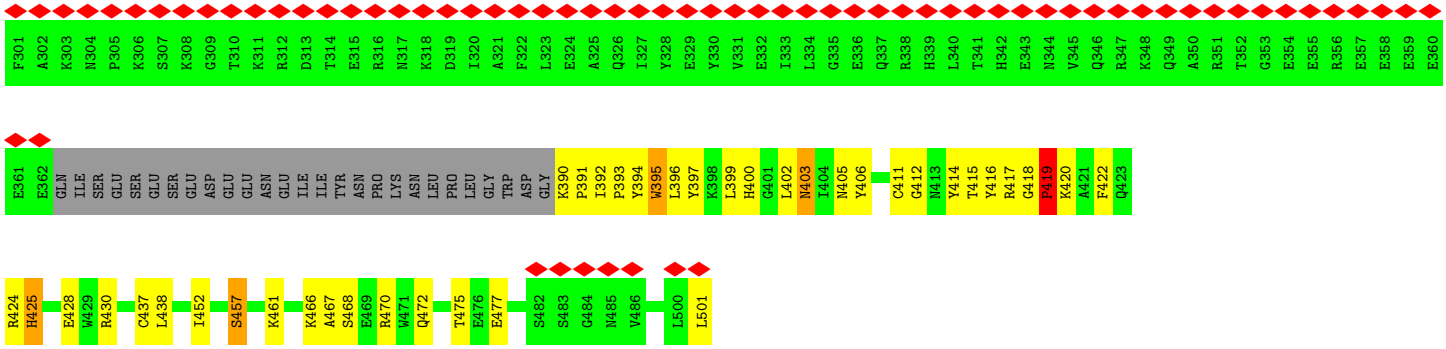
M1 E2 T3 I4 L5 E6 Q7 Q8 R9 R10 Y11 H12 E13 E14 K15 E16 R17 L18 M19 D20 V21 M22 A23 K24 E25 M26 L27 T28 K29 K30 S31 T32 L33 R34 D35 Q36 I37 N38 S39 D40 H41 R42 T43 R44 A45 M46 Q47 D48 R49 Y50 M51 E52 V53 S54 G55 N56 L57 R58 D59 L60

Y61 D62 D63 K64 D65 G66 L67 R68 K69 E70 E71 L72 N73 A74 I75 S76 G77 P78 N79 E80 F81 A82 E83 F84 Y85 N86 R87 L88 K89 Q90 I91 K92 E93 F94 H95 R96 K97 H98 P99 N100 E101 I102 C103 V104 P105 M106 S107 V108 E109 F110 E111 E112 L113 L114 K115 A116 R117 E118 N119 P120

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

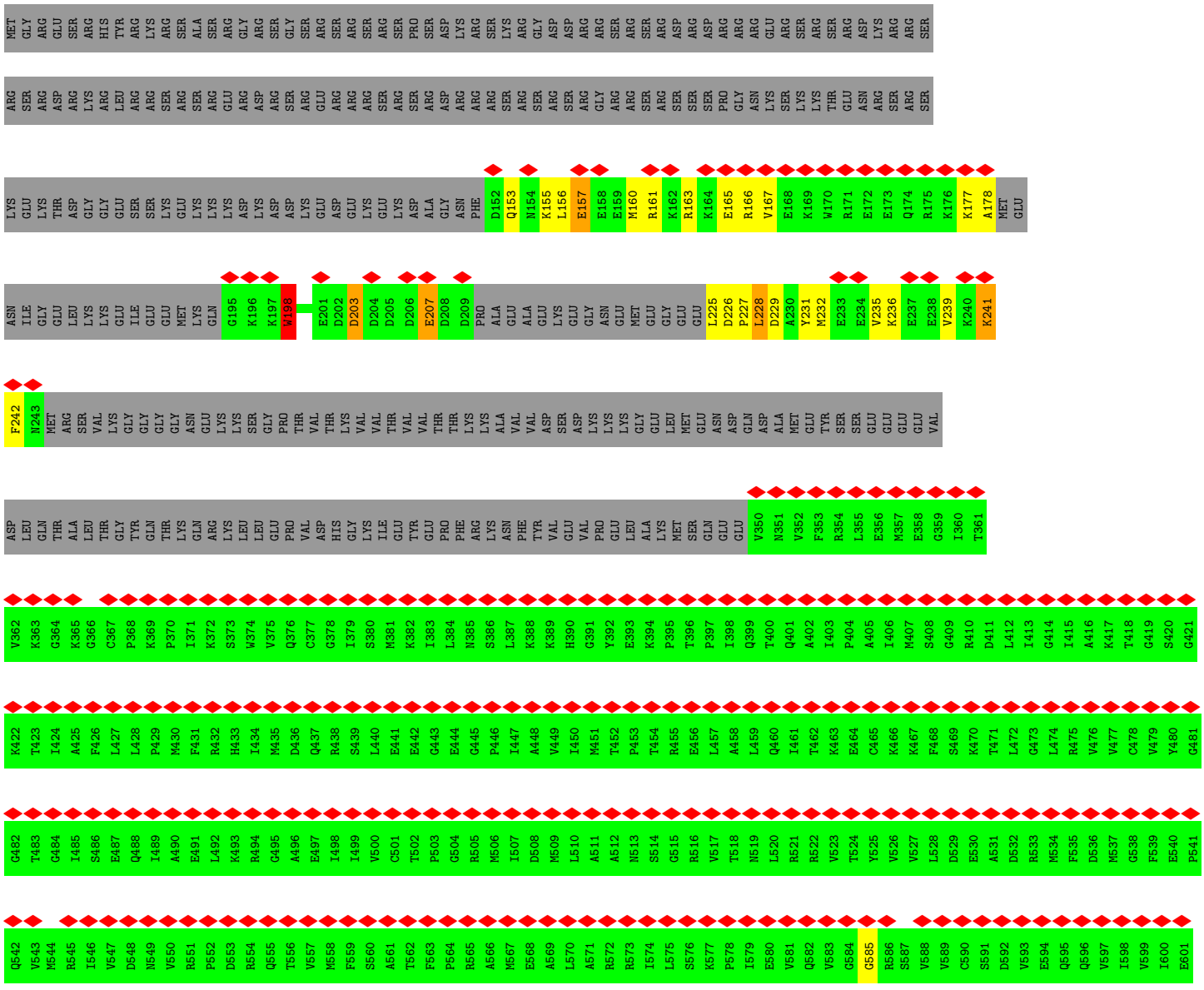
E181 Y182 K183 R184 Y185 L186 E187 M188 L189 L190 E191 Y192 L193 Q194 D195 Y196 T197 D198 R199 V200 K201 P202 L203 Q204 D205 Q206 N207 E208 L209 F210 G211 K212 I213 Q214 A215 E216 F217 E218 K219 K220 W221 E222 N223 G224 T225 F226 P227 G228 W229 PRO LYS GLU THR SER SER ALA LEU THR HIS ALA

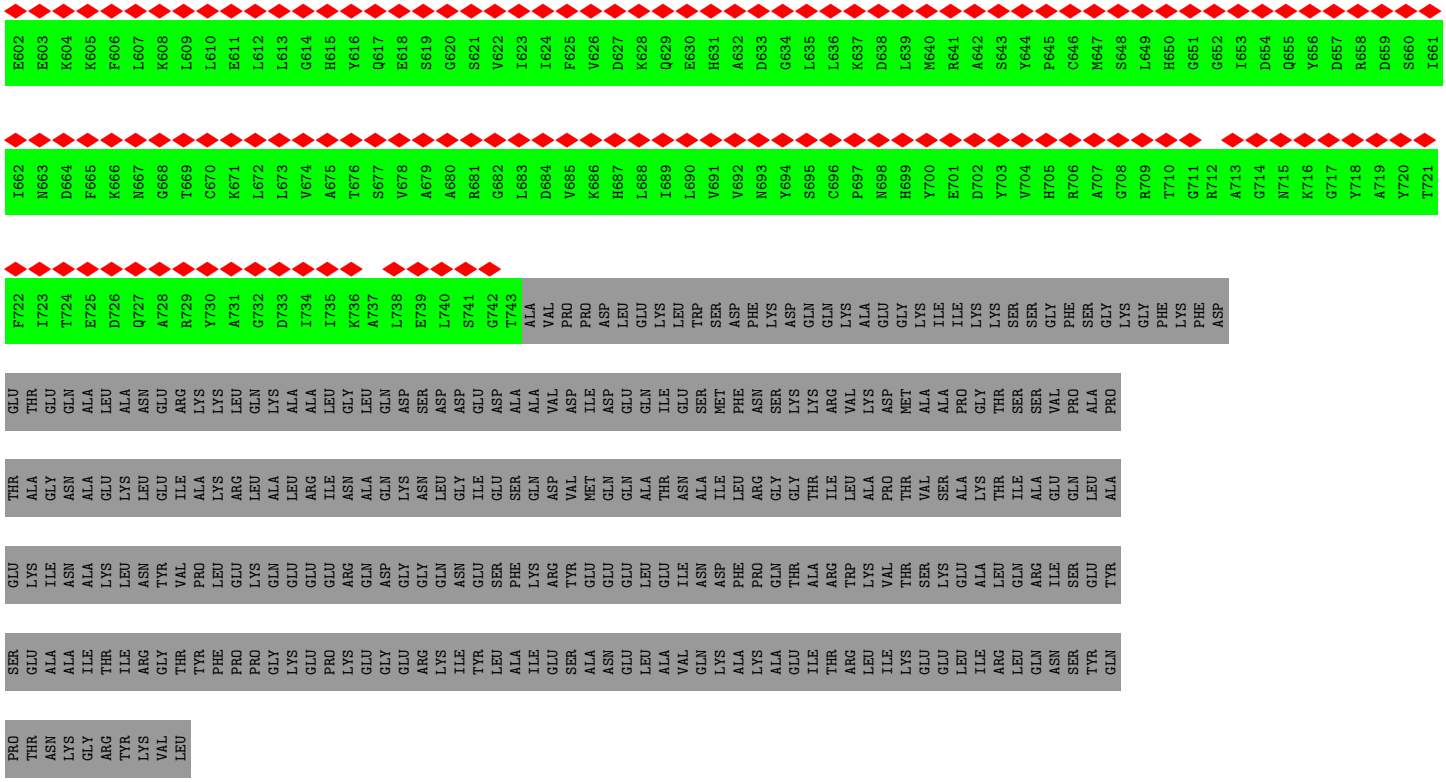
GLY ALA HIS LEU ASP LEU SER ALA PHE SER SER TRP GLU GLU LEU ALA SER LEU GLY LEU ASP ARG LEU LYS SER ALA LEU LEU ALA LEU GLY LEU LYS CYS GLY GLY THR LEU E279 E280 R281 A282 Q283 R284 L285 F286 S287 T288 K289 G290 K291 S292 L293 E294 S295 L296 D297 T298 S299 L300



• Molecule 12: Probable ATP-dependent RNA helicase DDX46

Chain E:



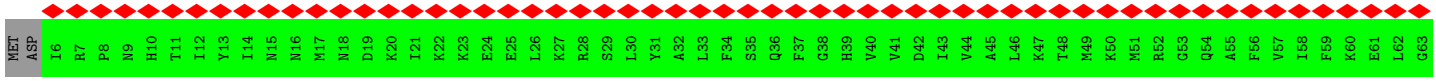


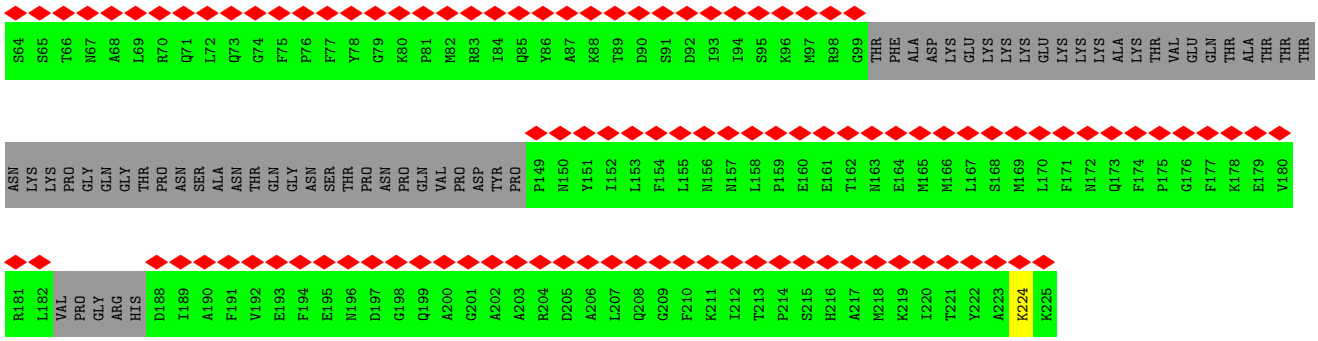
• Molecule 13: U2 small nuclear ribonucleoprotein A'

Chain F:

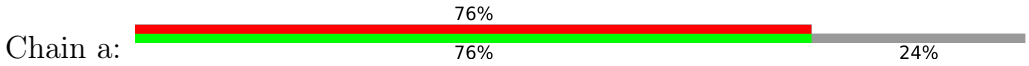
• Molecule 14: U2 small nuclear ribonucleoprotein B"

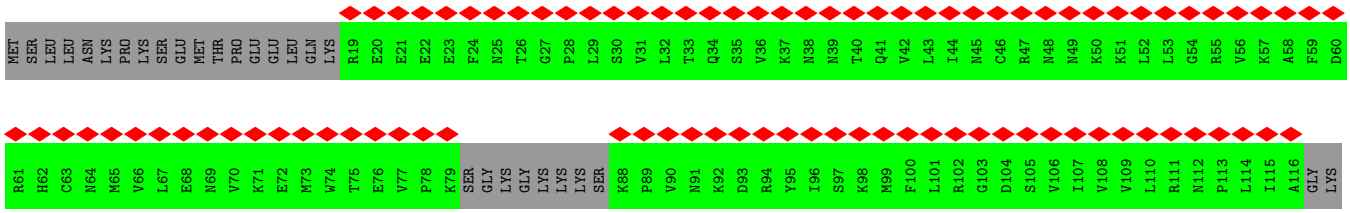
Chain G:



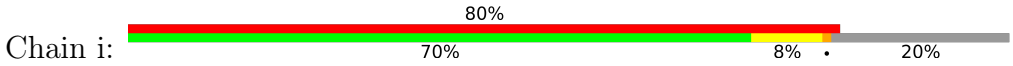


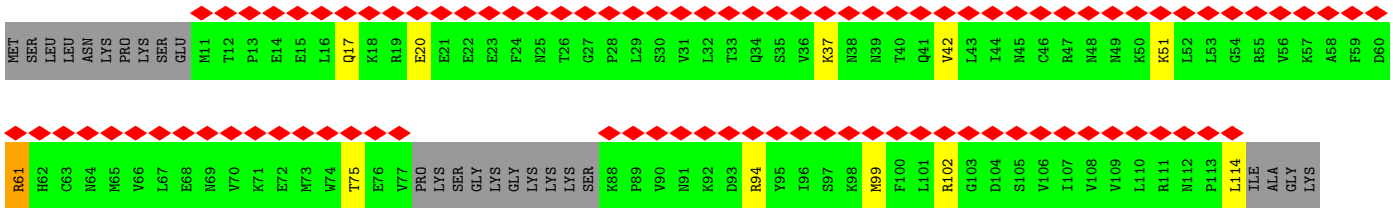
• Molecule 15: Small nuclear ribonucleoprotein Sm D2



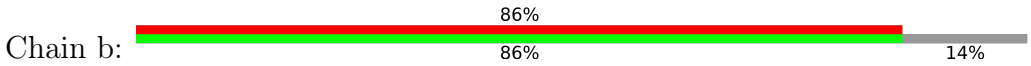


• Molecule 15: Small nuclear ribonucleoprotein Sm D2

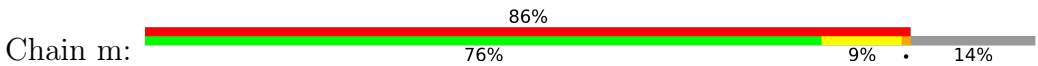


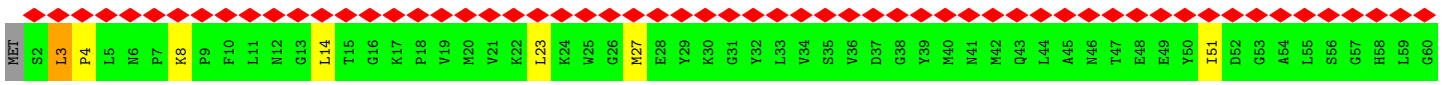


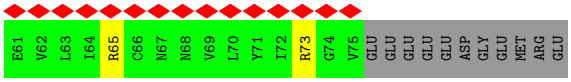
• Molecule 16: Small nuclear ribonucleoprotein F



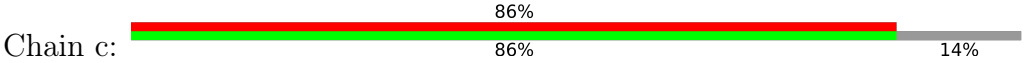
• Molecule 16: Small nuclear ribonucleoprotein F

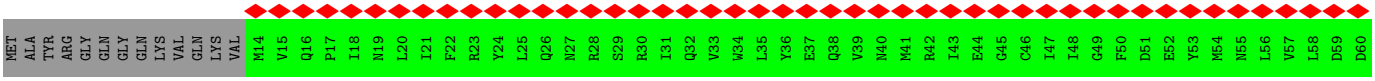


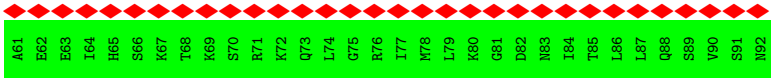




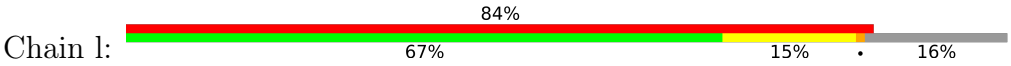
• Molecule 17: Small nuclear ribonucleoprotein E

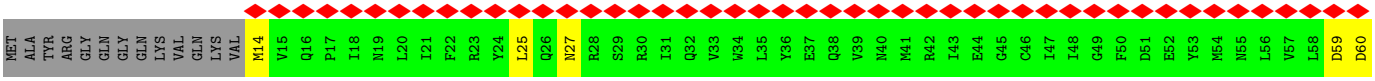


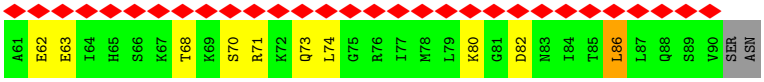




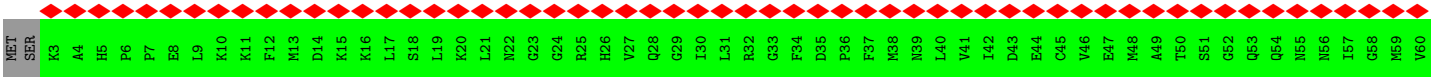
• Molecule 17: Small nuclear ribonucleoprotein E

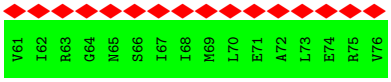




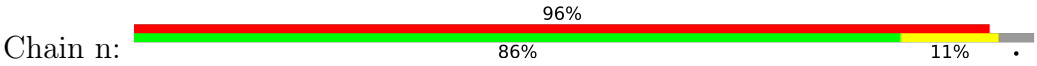


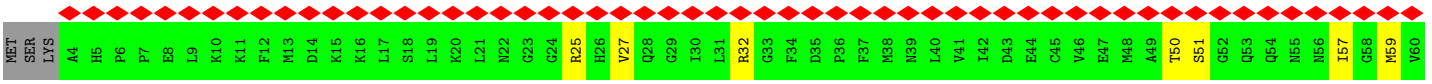
• Molecule 18: Small nuclear ribonucleoprotein G

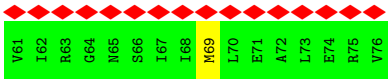




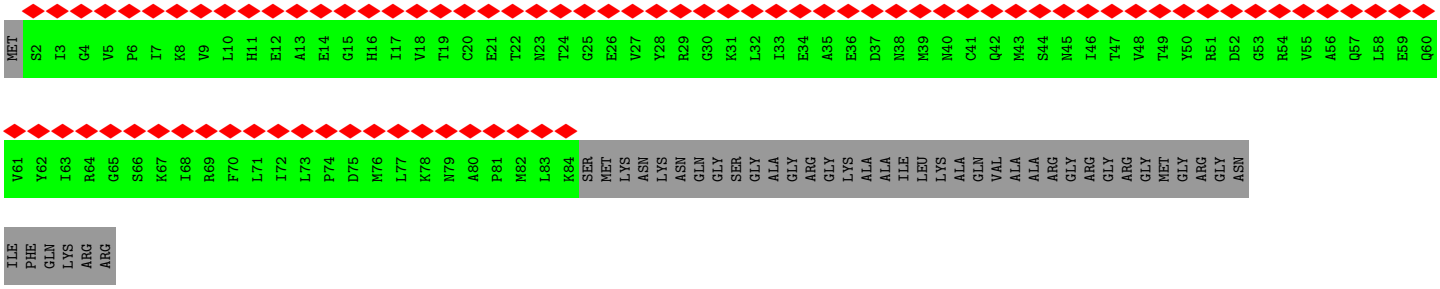
• Molecule 18: Small nuclear ribonucleoprotein G





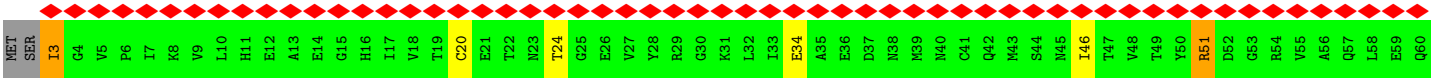


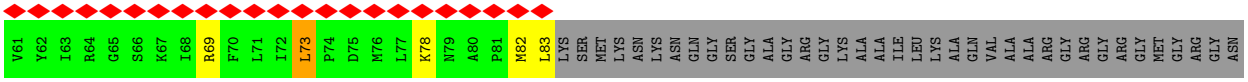
• Molecule 19: Small nuclear ribonucleoprotein Sm D3



• Molecule 19: Small nuclear ribonucleoprotein Sm D3

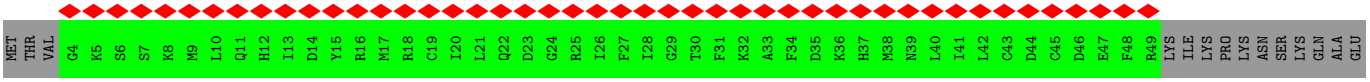
Chain j:

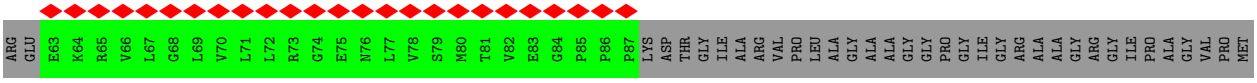




• Molecule 20: Small nuclear ribonucleoprotein-associated proteins B and B'

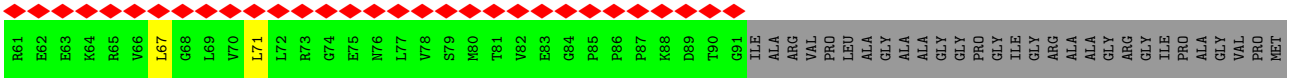
Chain f:

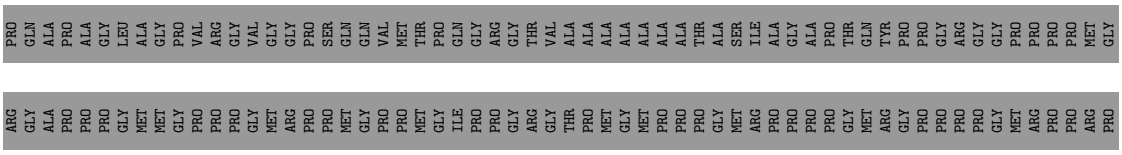




• Molecule 20: Small nuclear ribonucleoprotein-associated proteins B and B'

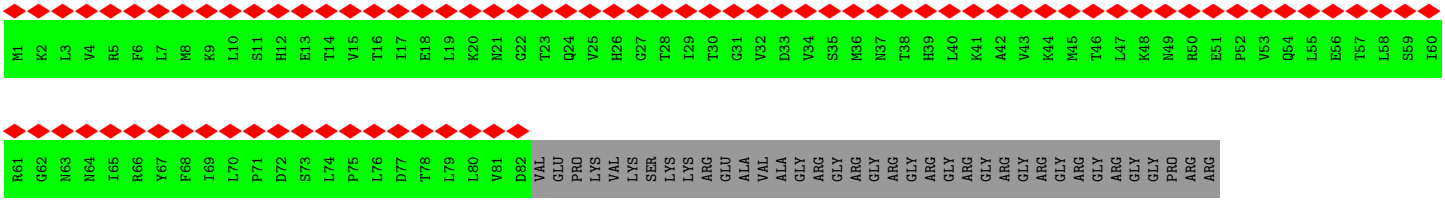
Chain k:





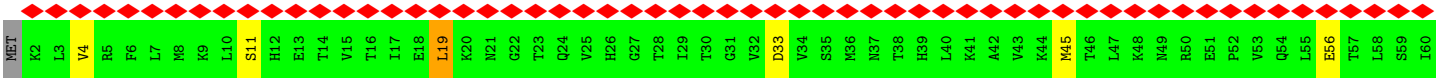
• Molecule 21: Small nuclear ribonucleoprotein Sm D1

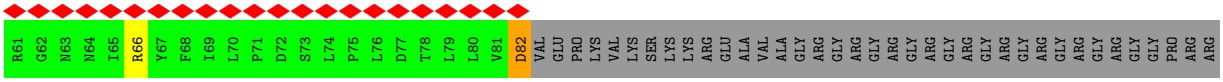
Chain g:



• Molecule 21: Small nuclear ribonucleoprotein Sm D1

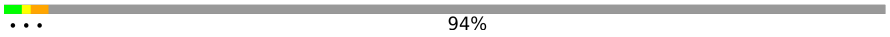
Chain h:

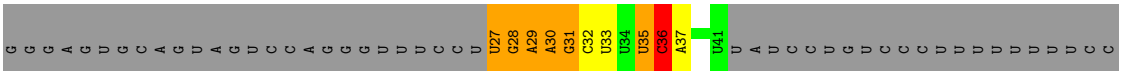


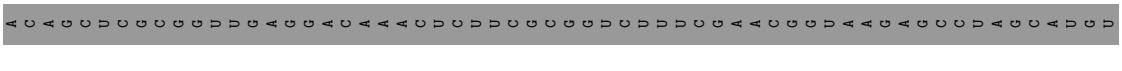


• Molecule 22: pre-mRNA

Chain I:

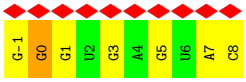






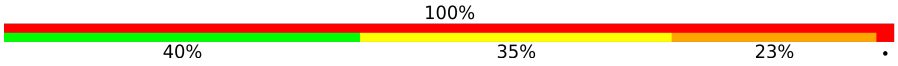
• Molecule 23: 5SS

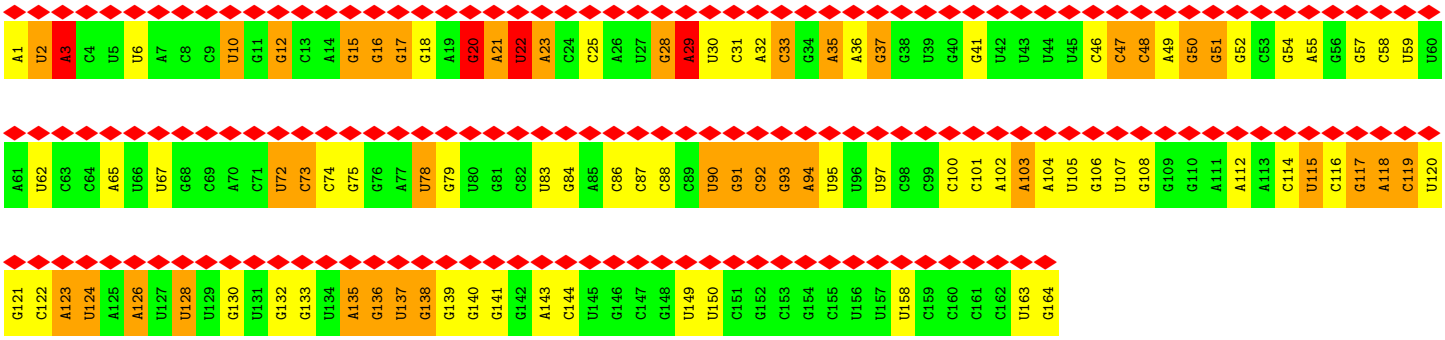
Chain K:



• Molecule 24: U1 snRNA

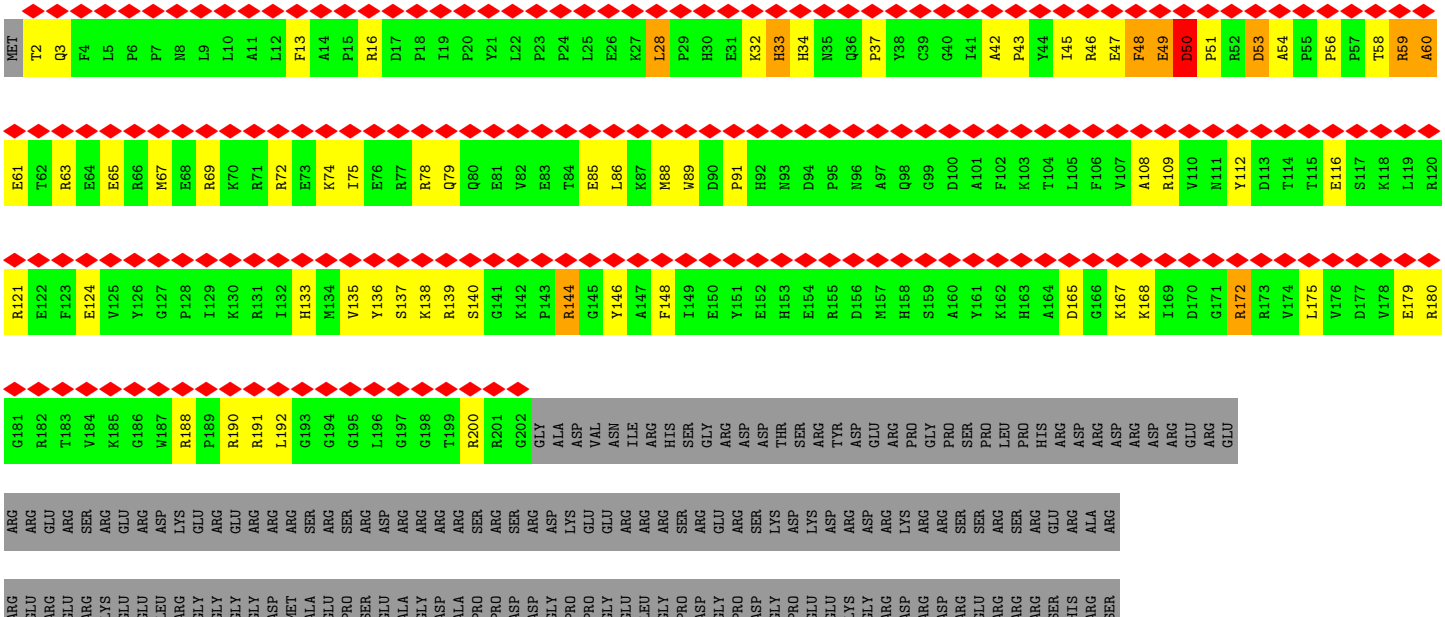
Chain L:





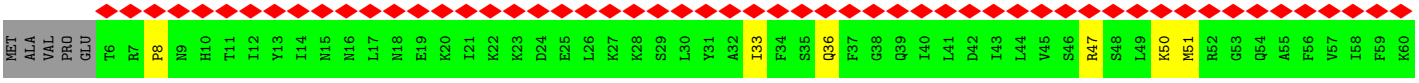
• Molecule 25: U1 small nuclear ribonucleoprotein 70 kDa

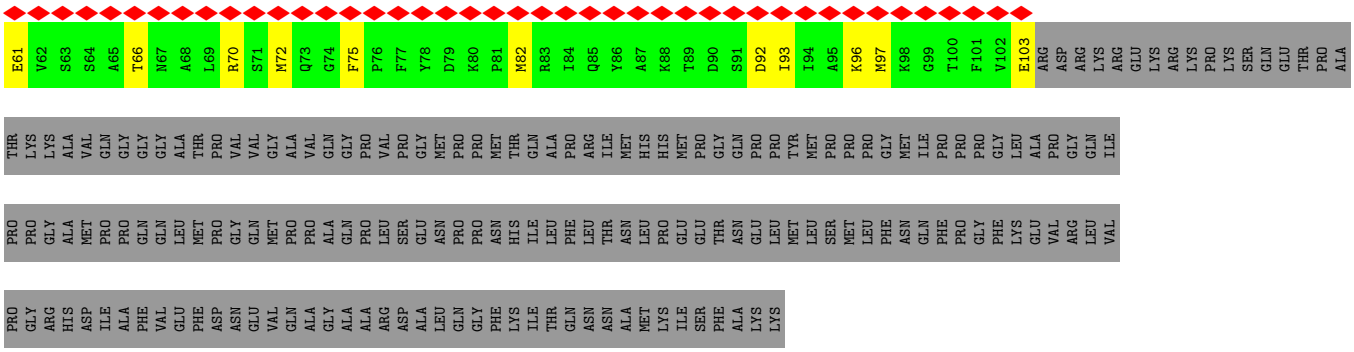
Chain O:



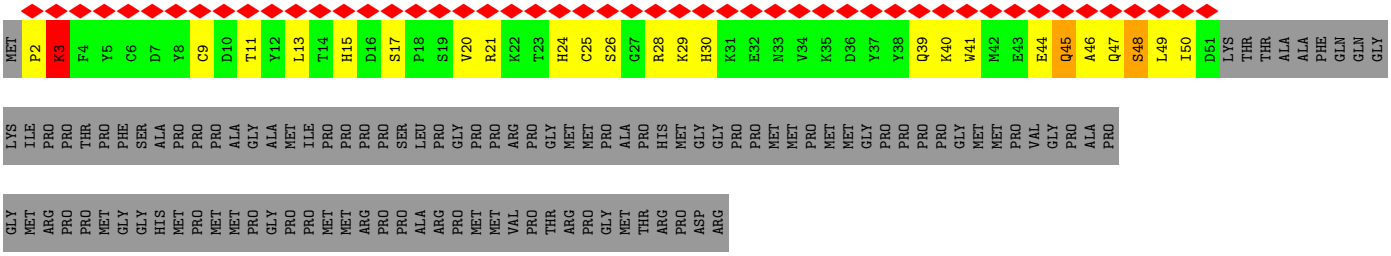
• Molecule 26: U1 small nuclear ribonucleoprotein A

Chain M:

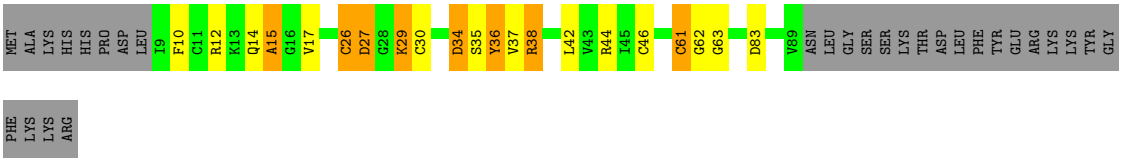




• Molecule 27: U1 small nuclear ribonucleoprotein C



• Molecule 28: PHD finger-like domain-containing protein 5A



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	419522	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	3.684	Depositor
Minimum map value	-2.084	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.043	Depositor
Recommended contour level	0.25	Depositor
Map size (Å)	521.76, 521.76, 521.76	wwPDB
Map dimensions	480, 480, 480	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.087, 1.087, 1.087	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: SJT, ZN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	B	0.31	0/812	0.52	0/1060
2	D	0.33	0/600	0.61	0/830
3	J	0.70	0/374	0.66	0/507
4	H	0.34	0/3073	1.09	26/4779 (0.5%)
5	1	0.38	2/6609 (0.0%)	0.57	9/8946 (0.1%)
6	2	0.44	0/1364	0.52	0/1831
7	3	0.27	0/9435	0.52	3/12802 (0.0%)
8	4	0.24	0/670	0.45	0/850
9	5	0.49	1/555 (0.2%)	0.52	0/750
10	A	0.24	0/507	0.41	0/640
11	C	0.43	2/2158 (0.1%)	0.51	1/2808 (0.0%)
12	E	0.42	1/2171 (0.0%)	0.55	1/2779 (0.0%)
13	F	0.24	0/671	0.49	0/849
14	G	0.24	0/689	0.46	0/869
15	a	0.24	0/374	0.50	0/472
15	i	0.69	0/773	1.15	3/1039 (0.3%)
16	b	0.25	0/311	0.51	0/395
16	m	0.65	1/588 (0.2%)	1.18	4/795 (0.5%)
17	c	0.24	0/319	0.49	0/399
17	l	0.62	0/646	1.27	5/867 (0.6%)
18	d	0.25	0/307	0.51	0/388
18	n	0.58	0/575	1.17	4/768 (0.5%)
19	e	0.25	0/343	0.51	0/433
19	j	0.53	0/645	1.19	6/870 (0.7%)
20	f	0.24	0/294	0.49	0/370
20	k	0.64	0/702	1.16	3/936 (0.3%)
21	g	0.24	0/339	0.51	0/428
21	h	0.68	0/649	1.24	7/878 (0.8%)
22	I	0.52	1/349 (0.3%)	1.19	5/540 (0.9%)
23	K	0.67	0/242	1.09	0/377
24	L	0.62	7/3891 (0.2%)	0.90	12/6061 (0.2%)
25	O	1.14	2/1688 (0.1%)	1.19	16/2279 (0.7%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
26	M	1.04	0/801	1.02	2/1074 (0.2%)
27	N	0.56	0/437	1.16	4/587 (0.7%)
28	6	0.68	2/621 (0.3%)	1.10	5/833 (0.6%)
All	All	0.49	19/44582 (0.0%)	0.78	116/61089 (0.2%)

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	1	0	1
15	i	0	1
25	O	0	1
All	All	0	3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	C	419	PRO	N-CA	13.18	1.69	1.47
5	1	718	PRO	N-CA	11.30	1.66	1.47
28	6	38	ARG	C-N	9.43	1.52	1.34
9	5	78	PRO	C-N	8.70	1.50	1.34
24	L	35	A	O3'-P	-7.27	1.52	1.61

The worst 5 of 116 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	3	1064	ASP	N-CA-CB	17.17	141.51	110.60
27	N	28	ARG	NE-CZ-NH1	12.38	126.49	120.30
24	L	3	A	N9-C1'-C2'	-10.85	99.89	114.00
19	j	73	LEU	CB-CG-CD2	10.23	128.39	111.00
16	m	73	ARG	NE-CZ-NH1	-10.09	115.26	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	1	1002	ASN	Mainchain
25	O	56	PRO	Peptide
15	i	17	GLN	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	799	0	514	31	0
2	D	604	0	261	2	0
3	J	369	0	275	22	0
4	H	2754	0	1398	21	0
5	1	6487	0	6699	498	0
6	2	1325	0	1184	131	0
7	3	9247	0	9179	197	0
8	4	664	0	230	0	0
9	5	539	0	505	38	0
10	A	504	0	150	0	0
11	C	2127	0	1138	76	0
12	E	2153	0	1054	143	0
13	F	666	0	209	0	0
14	G	685	0	226	7	0
15	a	372	0	118	0	0
15	i	764	0	783	0	0
16	b	308	0	115	0	0
16	m	576	0	589	0	0
17	c	319	0	92	0	0
17	l	638	0	657	0	0
18	d	305	0	105	0	0
18	n	568	0	590	0	0
19	e	341	0	112	0	0
19	j	637	0	652	0	0
20	f	293	0	99	0	0
20	k	692	0	717	0	0
21	g	337	0	110	0	0
21	h	641	0	680	0	0
22	I	314	0	160	33	0
23	K	216	0	110	6	0
24	L	3485	0	1759	257	0
25	O	1643	0	1605	145	0
26	M	787	0	797	7	0
27	N	425	0	392	20	0
28	6	613	0	596	20	0
29	6	3	0	0	0	0
29	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	C	1	0	0	0	0
29	N	1	0	0	0	0
30	6	37	0	0	16	0
All	All	43240	0	33860	1223	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 18.

The worst 5 of 1223 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)
6:2:574:ALA:HA	6:2:578:TRP:CE3	1.40	1.51
5:1:1114:VAL:HG21	30:6:204:SJT:C1	1.40	1.51
25:O:61:GLU:CB	25:O:65:GLU:HG2	1.41	1.50
24:L:92:C:H2'	24:L:93:G:C8	1.44	1.48
24:L:32:A:N3	25:O:135:VAL:HG22	1.22	1.47

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	
1	B	141/464 (30%)	139 (99%)	2 (1%)	0	100	100
2	D	114/639 (18%)	113 (99%)	1 (1%)	0	100	100
3	J	52/253 (21%)	51 (98%)	1 (2%)	0	100	100
5	1	813/1304 (62%)	776 (95%)	33 (4%)	4 (0%)	29	68
6	2	177/895 (20%)	162 (92%)	14 (8%)	1 (1%)	25	64
7	3	1168/1217 (96%)	1119 (96%)	47 (4%)	2 (0%)	47	82
8	4	156/424 (37%)	153 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	5	64/86 (74%)	63 (98%)	1 (2%)	0	100	100
10	A	121/793 (15%)	121 (100%)	0	0	100	100
11	C	419/501 (84%)	397 (95%)	20 (5%)	2 (0%)	29	68
12	E	447/1031 (43%)	430 (96%)	16 (4%)	1 (0%)	47	82
13	F	160/255 (63%)	160 (100%)	0	0	100	100
14	G	160/225 (71%)	158 (99%)	2 (1%)	0	100	100
15	a	86/118 (73%)	84 (98%)	2 (2%)	0	100	100
15	i	90/118 (76%)	84 (93%)	6 (7%)	0	100	100
16	b	72/86 (84%)	71 (99%)	1 (1%)	0	100	100
16	m	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
17	c	77/92 (84%)	77 (100%)	0	0	100	100
17	l	75/92 (82%)	70 (93%)	5 (7%)	0	100	100
18	d	72/76 (95%)	72 (100%)	0	0	100	100
18	n	71/76 (93%)	69 (97%)	2 (3%)	0	100	100
19	e	81/126 (64%)	79 (98%)	2 (2%)	0	100	100
19	j	79/126 (63%)	75 (95%)	4 (5%)	0	100	100
20	f	67/240 (28%)	67 (100%)	0	0	100	100
20	k	84/240 (35%)	82 (98%)	2 (2%)	0	100	100
21	g	80/119 (67%)	79 (99%)	1 (1%)	0	100	100
21	h	79/119 (66%)	77 (98%)	2 (2%)	0	100	100
25	O	199/437 (46%)	182 (92%)	13 (6%)	4 (2%)	7	34
26	M	96/282 (34%)	94 (98%)	2 (2%)	0	100	100
27	N	48/159 (30%)	47 (98%)	1 (2%)	0	100	100
28	6	79/110 (72%)	71 (90%)	5 (6%)	3 (4%)	3	18
All	All	5499/10789 (51%)	5291 (96%)	191 (4%)	17 (0%)	44	76

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	1	969	LYS
5	1	1052	ALA
5	1	1053	ARG
7	3	1064	ASP
25	O	59	ARG

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
1	B	47/382 (12%)	47 (100%)	0	100 100
3	J	24/226 (11%)	22 (92%)	2 (8%)	11 39
5	1	701/1104 (64%)	653 (93%)	48 (7%)	16 48
6	2	124/776 (16%)	118 (95%)	6 (5%)	25 62
7	3	1022/1051 (97%)	963 (94%)	59 (6%)	20 55
8	4	8/336 (2%)	8 (100%)	0	100 100
9	5	57/77 (74%)	56 (98%)	1 (2%)	59 85
10	A	4/709 (1%)	4 (100%)	0	100 100
11	C	83/446 (19%)	73 (88%)	10 (12%)	5 22
12	E	73/892 (8%)	65 (89%)	8 (11%)	6 25
13	F	6/218 (3%)	6 (100%)	0	100 100
14	G	7/195 (4%)	7 (100%)	0	100 100
15	a	4/110 (4%)	4 (100%)	0	100 100
15	i	89/110 (81%)	81 (91%)	8 (9%)	9 35
16	b	4/74 (5%)	4 (100%)	0	100 100
16	m	63/74 (85%)	56 (89%)	7 (11%)	6 25
17	c	1/84 (1%)	1 (100%)	0	100 100
17	l	72/84 (86%)	61 (85%)	11 (15%)	2 13
18	d	3/66 (4%)	3 (100%)	0	100 100
18	n	63/66 (96%)	57 (90%)	6 (10%)	8 32
19	e	3/101 (3%)	3 (100%)	0	100 100
19	j	71/101 (70%)	61 (86%)	10 (14%)	3 16
20	f	3/177 (2%)	3 (100%)	0	100 100
20	k	78/177 (44%)	70 (90%)	8 (10%)	7 28
21	g	3/101 (3%)	3 (100%)	0	100 100
21	h	76/101 (75%)	72 (95%)	4 (5%)	22 58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
25	O	169/373 (45%)	160 (95%)	9 (5%)	22 58
26	M	85/240 (35%)	82 (96%)	3 (4%)	36 71
27	N	48/135 (36%)	42 (88%)	6 (12%)	4 20
28	6	69/95 (73%)	62 (90%)	7 (10%)	7 29
All	All	3060/8681 (35%)	2847 (93%)	213 (7%)	19 47

5 of 213 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	5	48	ASP
16	m	3	LEU
19	j	34	GLU
11	C	425	HIS
12	E	225	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
7	3	594	ASN
7	3	916	ASN
7	3	612	ASN
7	3	805	ASN
7	3	983	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
22	I	15/259 (5%)	7 (46%)	4 (26%)
23	K	9/10 (90%)	2 (22%)	0
24	L	163/164 (99%)	51 (31%)	6 (3%)
4	H	126/188 (67%)	29 (23%)	2 (1%)
All	All	313/621 (50%)	89 (28%)	12 (3%)

5 of 89 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	H	35	A
4	H	36	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	45	C
4	H	46	U
4	H	47	U

5 of 12 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
24	L	90	U
24	L	92	C
24	L	128	U
24	L	123	A
22	I	28	G

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
30	SJT	6	204	-	37,38,38	1.24	6 (16%)	42,55,55	1.56	7 (16%)

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
30	SJT	6	204	-	5/5/12/18	19/26/63/63	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
30	6	204	SJT	C17-C16	-3.16	1.44	1.50
30	6	204	SJT	C1-C	-3.11	1.41	1.48
30	6	204	SJT	O1-C3	-2.66	1.44	1.47
30	6	204	SJT	O1-C5	-2.45	1.29	1.35
30	6	204	SJT	C18-C16	-2.13	1.41	1.45

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
30	6	204	SJT	C2-C1-C	-5.55	108.80	122.69
30	6	204	SJT	C26-C25-C24	-3.87	109.69	113.37
30	6	204	SJT	O7-C25-C26	-3.43	111.01	113.61
30	6	204	SJT	C13-C7-N	3.06	115.59	110.86
30	6	204	SJT	C13-C7-C8	2.50	114.03	109.34

All (5) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
30	6	204	SJT	C22
30	6	204	SJT	C25
30	6	204	SJT	C8
30	6	204	SJT	C10
30	6	204	SJT	C3

5 of 19 torsion outliers are listed below:

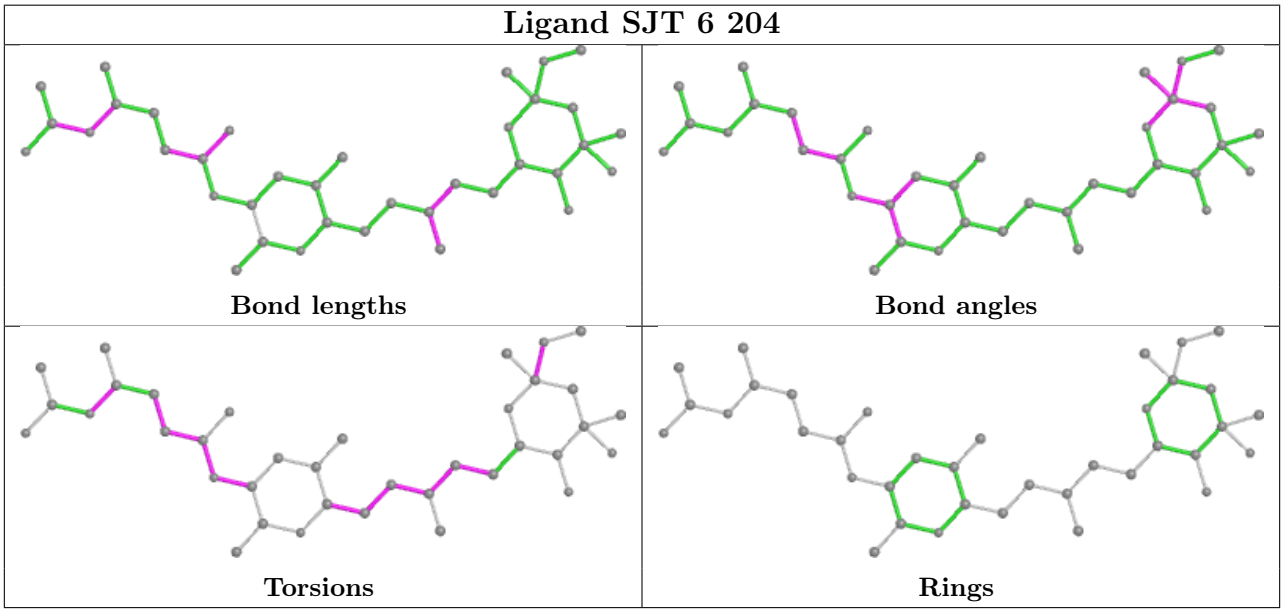
Mol	Chain	Res	Type	Atoms
30	6	204	SJT	C-C1-C2-C3
30	6	204	SJT	O-C-N-C7
30	6	204	SJT	C1-C-N-C7
30	6	204	SJT	C13-C7-N-C
30	6	204	SJT	C11-C10-C14-C15

There are no ring outliers.

1 monomer is involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
30	6	204	SJT	16	0

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



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Map visualisation

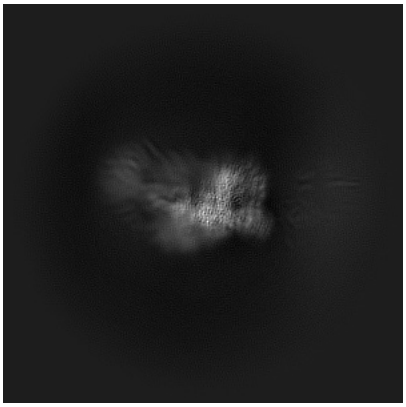
This section contains visualisations of the EMDB entry EMD-32074. These allow visual inspection of the internal detail of the map and identification of artifacts.

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.

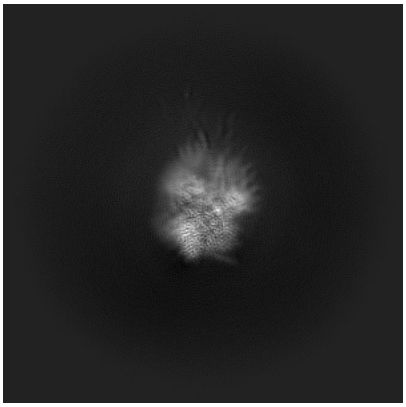
6.1 Orthogonal projections

6.1.1 Primary map

X



Y



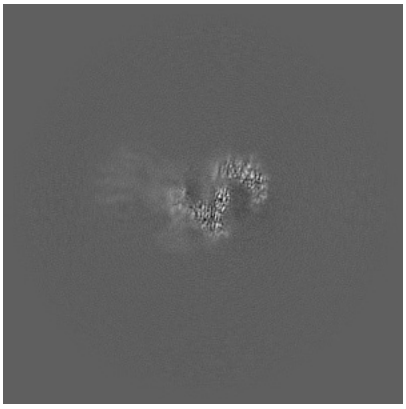
Z

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

6.2 Central slices

6.2.1 Primary map

X Index: 240



Y Index: 240

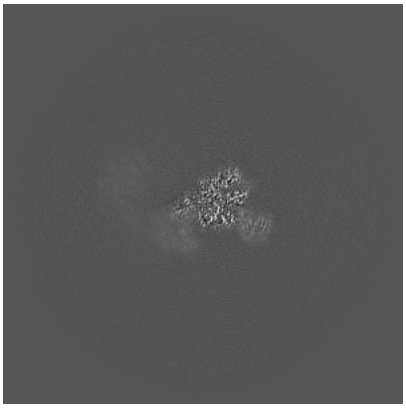
Z Index: 240

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

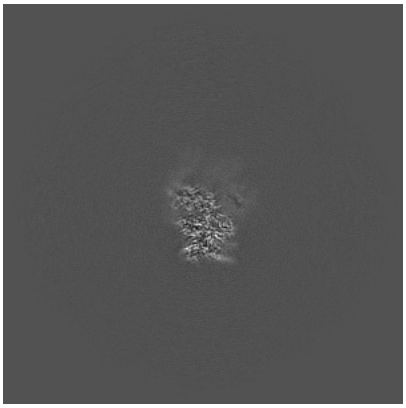
6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 226



Y Index: 211

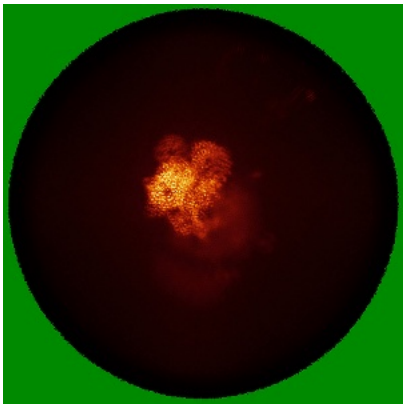


Z Index: 260

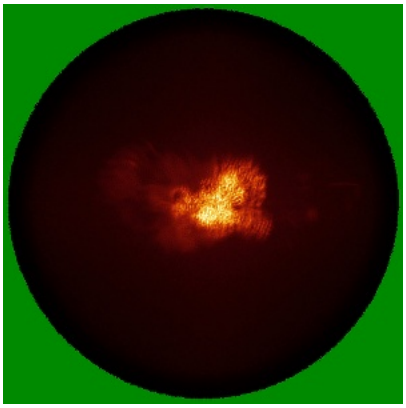
The images above show the largest variance slices of the map in three orthogonal directions.

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

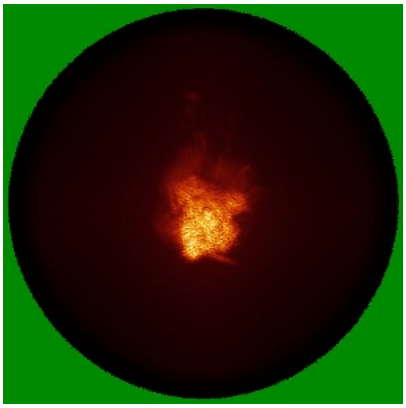
6.4.1 Primary map



X



Y



Z

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

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.25. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

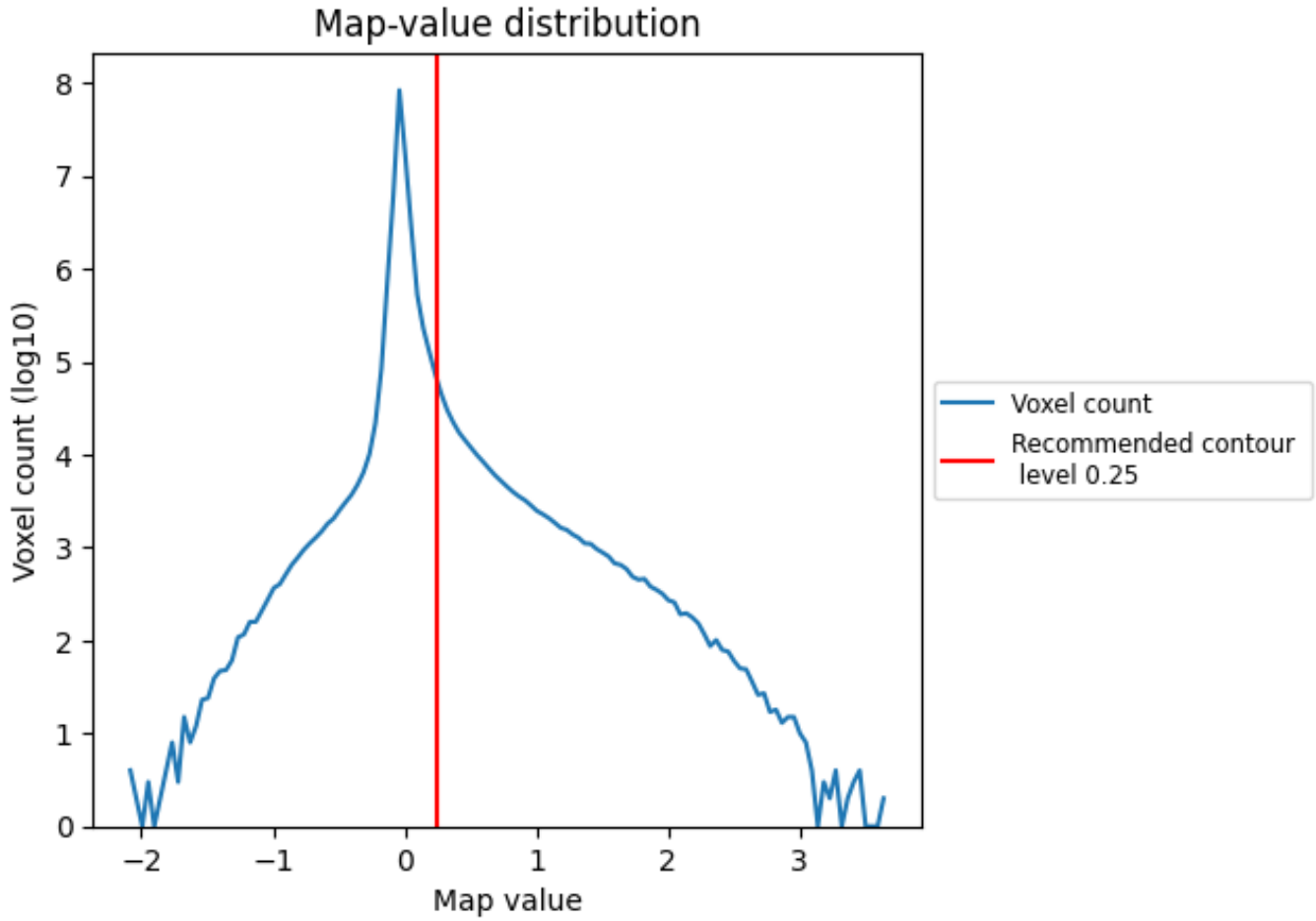
6.6 Mask visualisation (i)

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

7 Map analysis

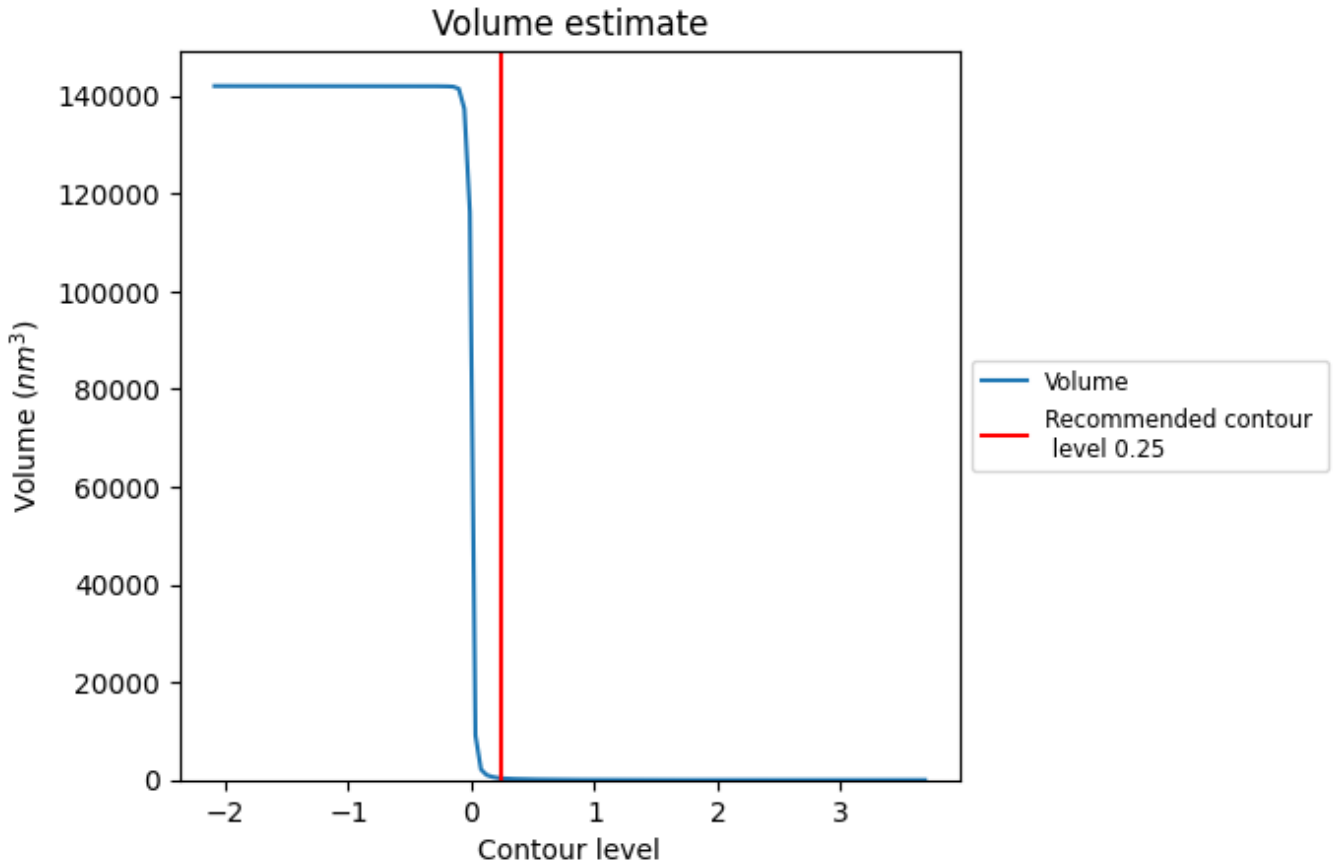
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

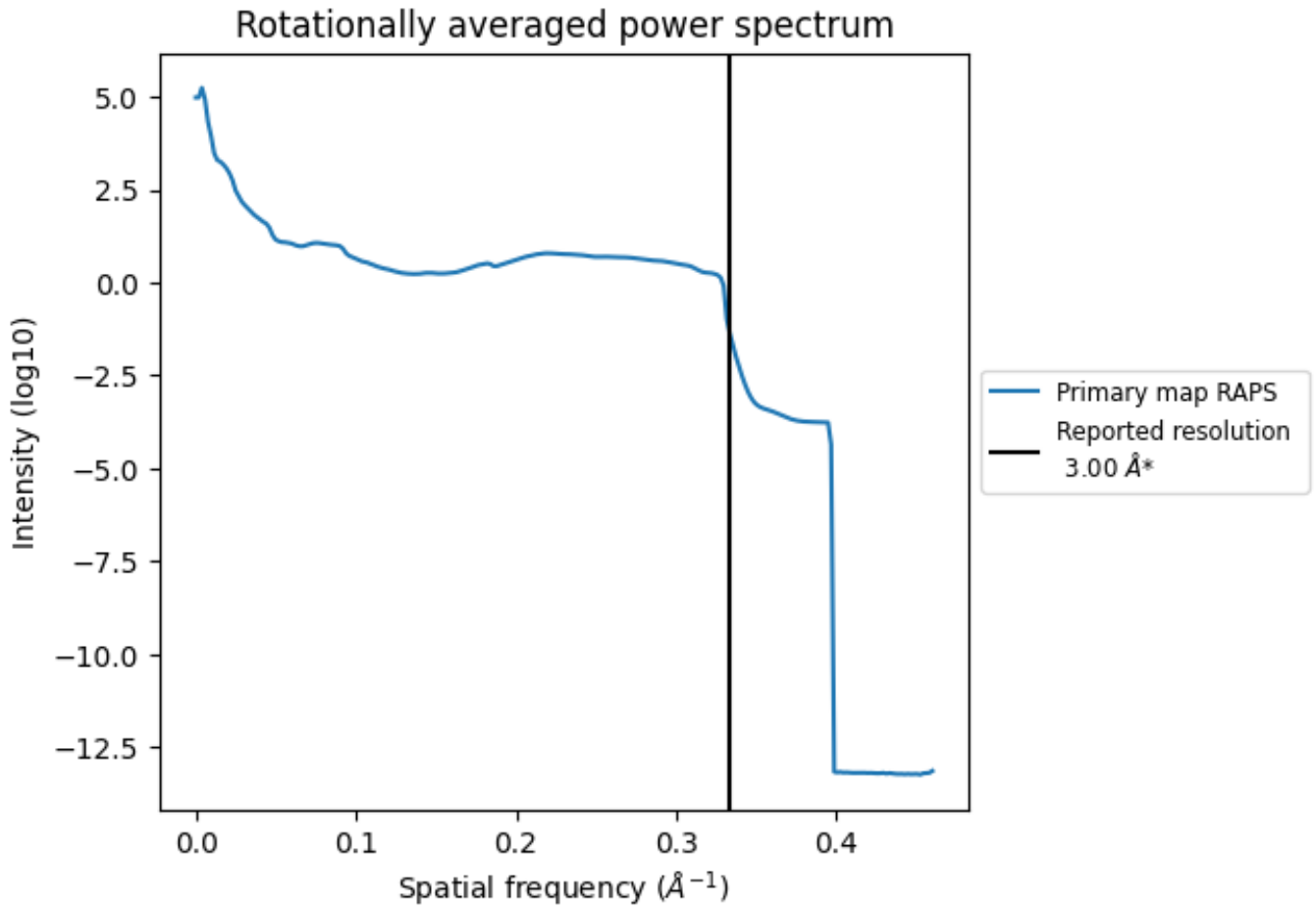
7.2 Volume estimate (i)



The volume at the recommended contour level is 341 nm^3; this corresponds to an approximate mass of 308 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 0.333 $Å^{-1}$

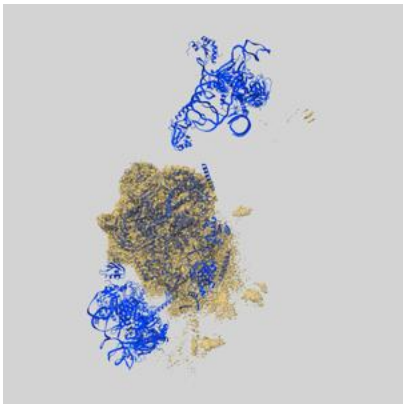
8 Fourier-Shell correlation

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

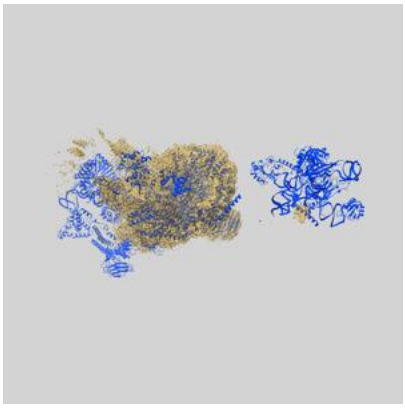
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-32074 and PDB model 7VPX. Per-residue inclusion information can be found in section 3 on page 11.

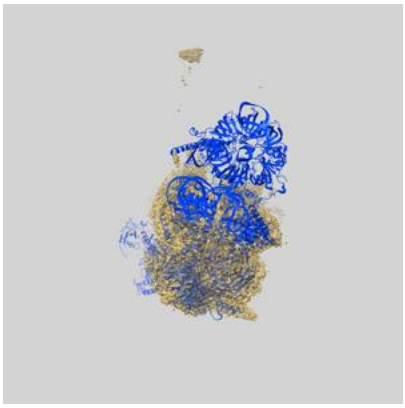
9.1 Map-model overlay



X



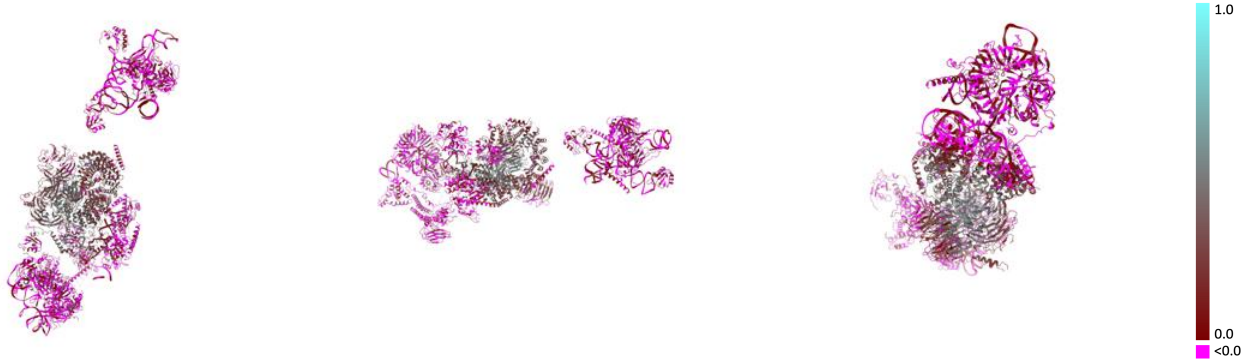
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.25 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.

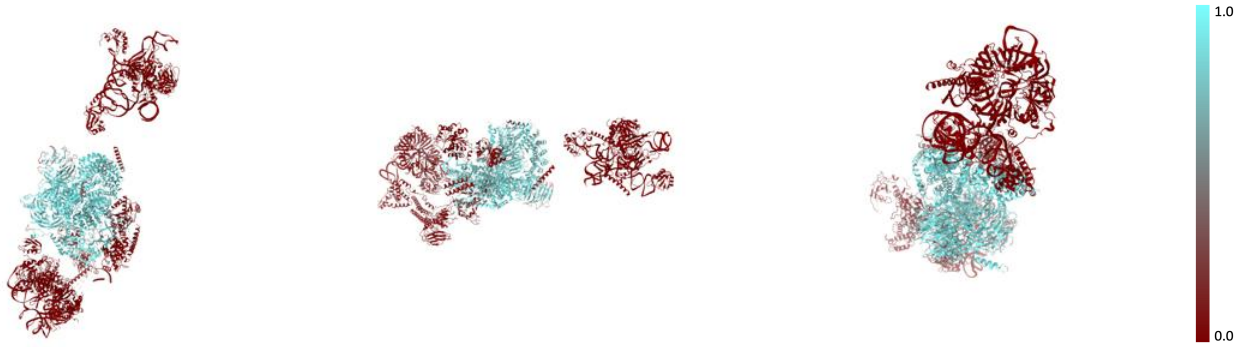
9.2 Q-score mapped to coordinate model (i)



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.

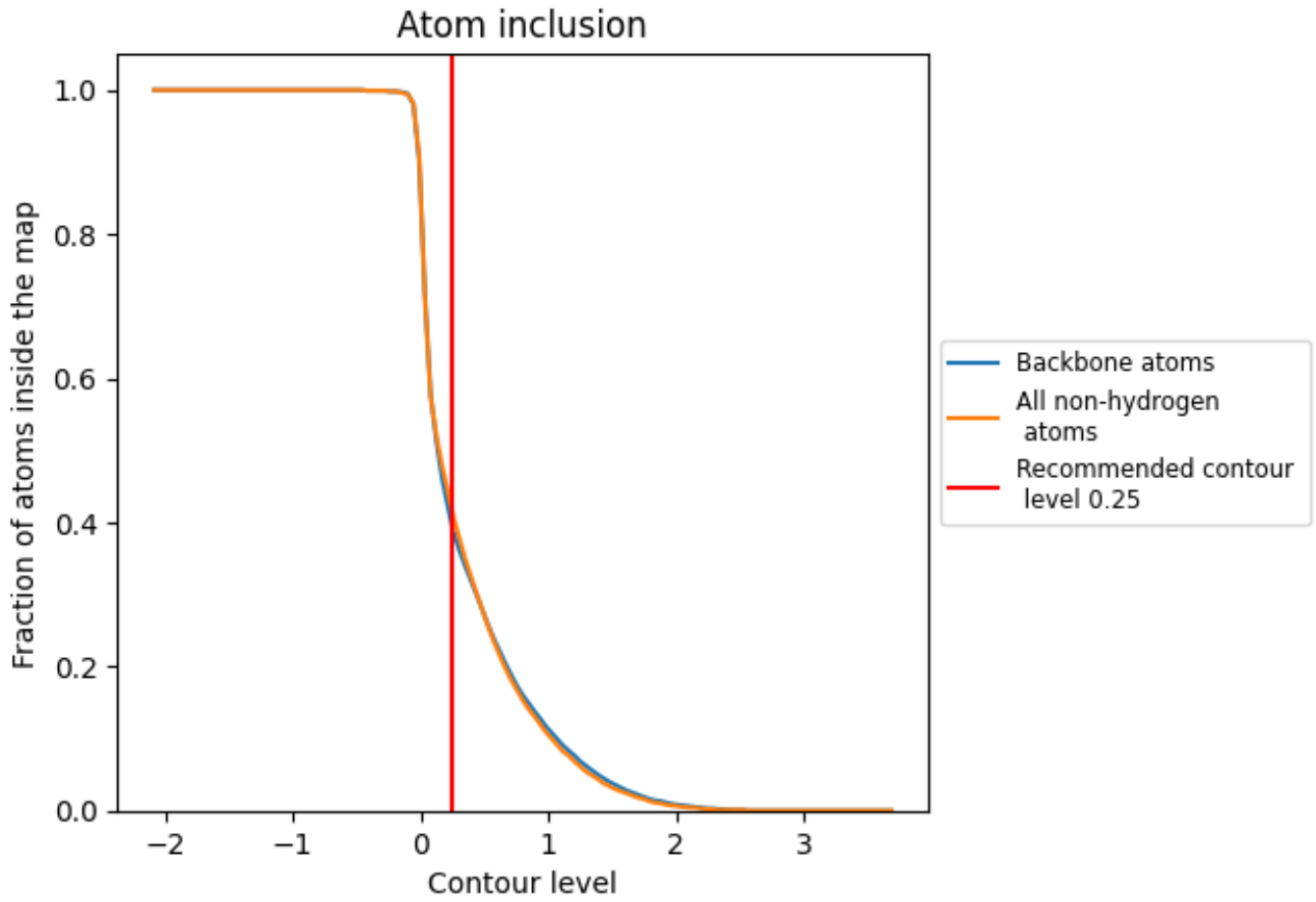
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.25).

9.4 Atom inclusion (i)



At the recommended contour level, 39% of all backbone atoms, 41% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.25) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.4130	0.1550
1	0.8560	0.3180
2	0.6680	0.2780
3	0.8360	0.3320
4	0.1820	-0.0130
5	0.9150	0.4810
6	0.9070	0.4080
A	0.0000	-0.0220
B	0.3050	0.0340
C	0.3410	0.1190
D	0.3080	0.0390
E	0.1050	0.0210
F	0.0000	0.0170
G	0.0000	-0.0200
H	0.2320	0.0690
I	0.8540	0.1490
J	0.7150	0.1100
K	0.0000	0.0440
L	0.0000	0.0100
M	0.0000	0.0120
N	0.0000	0.0110
O	0.0000	0.0030
a	0.0000	0.0040
b	0.0000	0.0060
c	0.0000	-0.0130
d	0.0000	0.0090
e	0.0000	0.0120
f	0.0030	-0.0260
g	0.0030	-0.0450
h	0.0000	-0.0010
i	0.0000	0.0050
j	0.0000	0.0260
k	0.0000	0.0060
l	0.0000	0.0120
m	0.0000	-0.0060
n	0.0000	-0.0180