wwPDB EM Validation Summary Report (i)

Aug 8, 2023 – 02:24 PM JST

PDB ID	:	8HK1
EMDB ID	:	EMD-34841
Title	:	The cryo-EM structure of human pre-17S U2 snRNP
Authors	:	Zhang, X.; Zhan, X.; Shi, Y.
Deposited on	:	2022-11-24
Resolution	:	2.70 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:

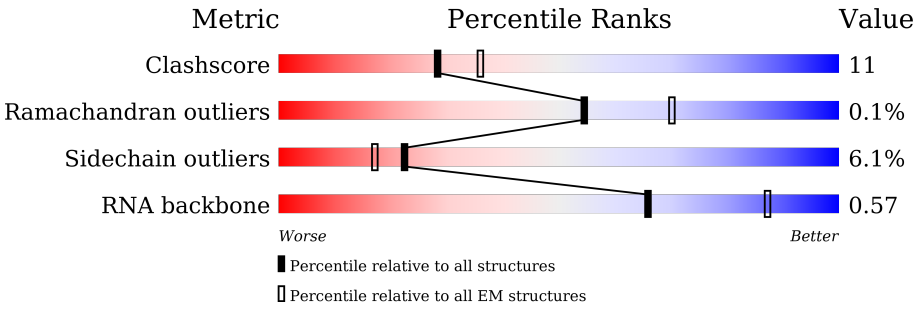
EMDB validation analysis	:	0.0.1.dev50
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.70 Å.

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



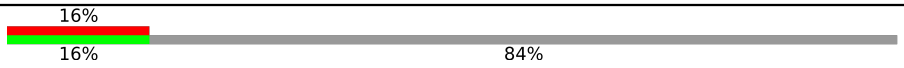
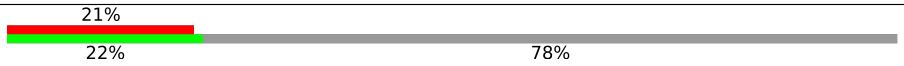
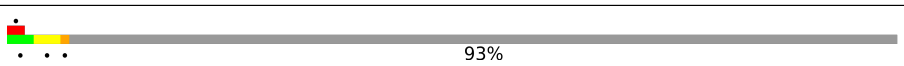
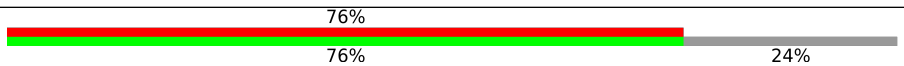
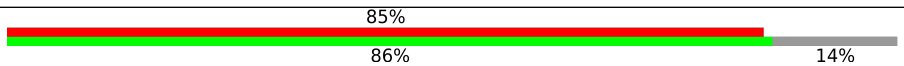
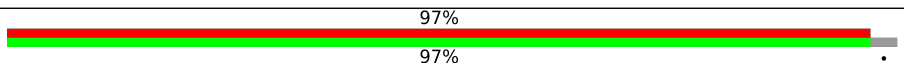
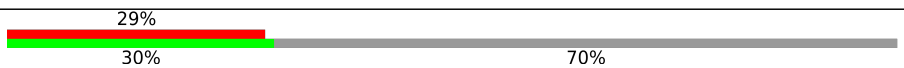
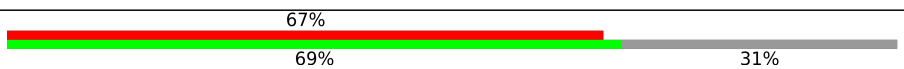
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=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	H	188	53% 45% 26% 7% • 21%
2	1	1304	46% 16% • 38%
3	2	895	7% 17% • 80%
4	3	1217	• 73% 22% • •
5	4	424	33% 38% 62%
6	5	86	66% 21% • 12%
7	6	110	63% 15% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	A	793	
9	B	464	
10	C	501	
11	D	755	
12	E	938	
13	F	255	
14	G	225	
15	a	118	
16	b	86	
17	c	92	
18	d	76	
19	e	126	
20	f	240	
21	g	119	

2 Entry composition (i)

There are 23 unique types of molecules in this entry. The entry contains 30236 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	H	148	3142	1406	551	1039	146	0	0

- Molecule 2 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	1	815	6487	4163	1121	1164	39	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	182	1210	754	229	225	2	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	1180	9247	5872	1571	1759	45	0	0

- Molecule 5 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	4	160	664	344	160	160	0	0

- Molecule 6 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	76	623	394	109	115	5	0	0

- Molecule 7 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	6	85	Total	C	N	O	S	0	0
			645	396	114	122	13		

- Molecule 8 is a protein called Splicing factor 3A subunit 1.

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

- Molecule 9 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	B	100	Total	C	N	O	0	0
			421	221	100	100		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	C	422	Total	C	N	O	S	0	0
			2103	1174	466	459	4		

- Molecule 11 is a protein called HIV Tat-specific factor 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	D	113	Total	C	N	O	S	0	0
			930	592	155	178	5		

- Molecule 12 is a protein called ATP-dependent RNA helicase DDX42.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	E	68	Total	C	N	O	S	0	0
			579	351	94	131	3		

- 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 685	C 353	N 166	O 166	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
15	a	90	Total 372	C 192	N 90	O 90	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
16	b	74	Total 308	C 160	N 74	O 74	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
17	c	79	Total 319	C 161	N 79	O 79	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
18	d	74	Total 305	C 157	N 74	O 74	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
19	e	83	Total 341	C 175	N 83	O 83	0	0

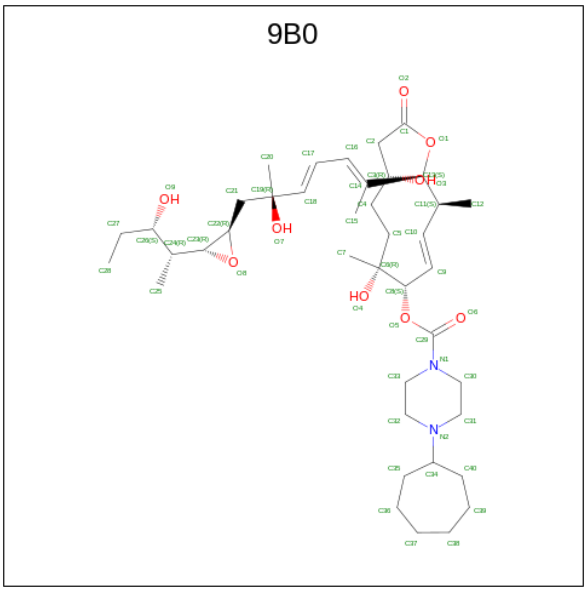
- Molecule 20 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	f	71	Total 293	C 151	N 71	O 71	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	0	0
			337	173	82	82		

- Molecule 22 is [(2 {S},3 {S},4 {E},6 {S},7 {R},10 {R})-3,7-dimethyl-2-[(2 {E},4 {E},6 {R})-6-methyl-6-oxidanyl-7-[(2 {R},3 {R})-3-[(2 {R},3 {S})-3-oxidanylpentan-2-yl]oxiran-2-yl]hepta-2,4-dien-2-yl]-7,10-bis(oxidanyl)-12-oxidanylidene-1-oxacyclododec-4-en-6-yl] 4-cycloheptylpiperazine-1-carboxylate (three-letter code: 9B0) (formula: $C_{40}H_{66}N_2O_9$).



Mol	Chain	Residues	Atoms				AltConf
22	1	1	Total	C	N	O	0
			51	40	2	9	

- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

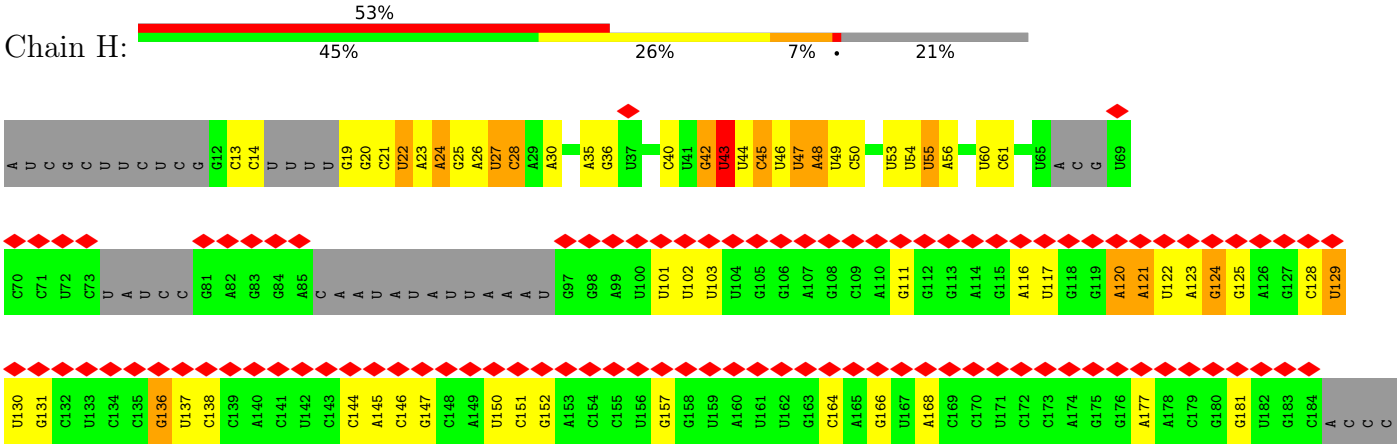
Mol	Chain	Residues	Atoms		AltConf
23	6	3	Total	Zn	0
			3	3	
23	C	1	Total	Zn	0
			1	1	

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.

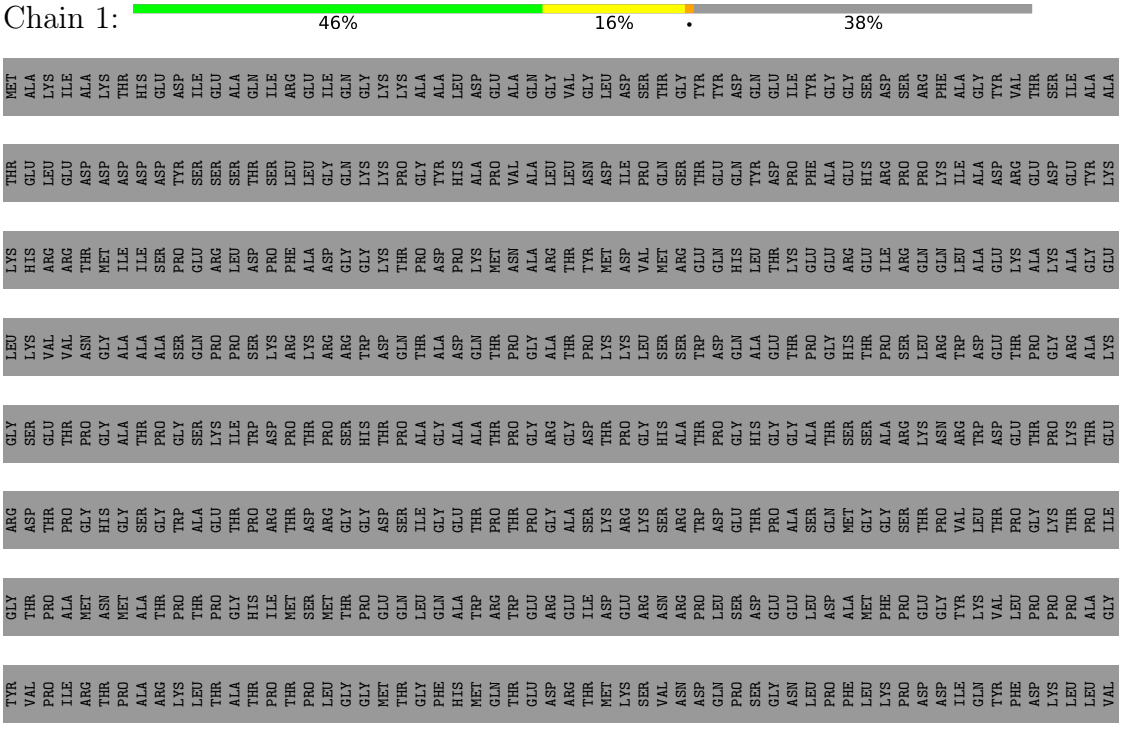
• Molecule 1: U2 snRNA

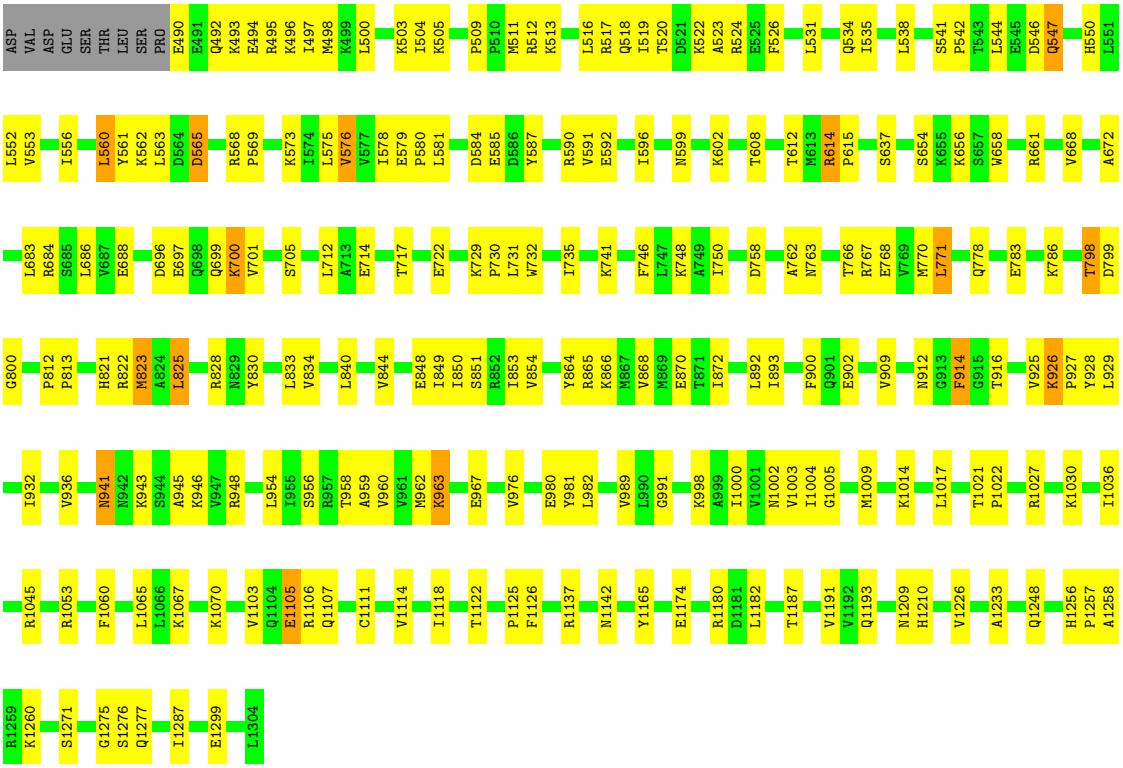
Chain H:



• Molecule 2: Splicing factor 3B subunit 1

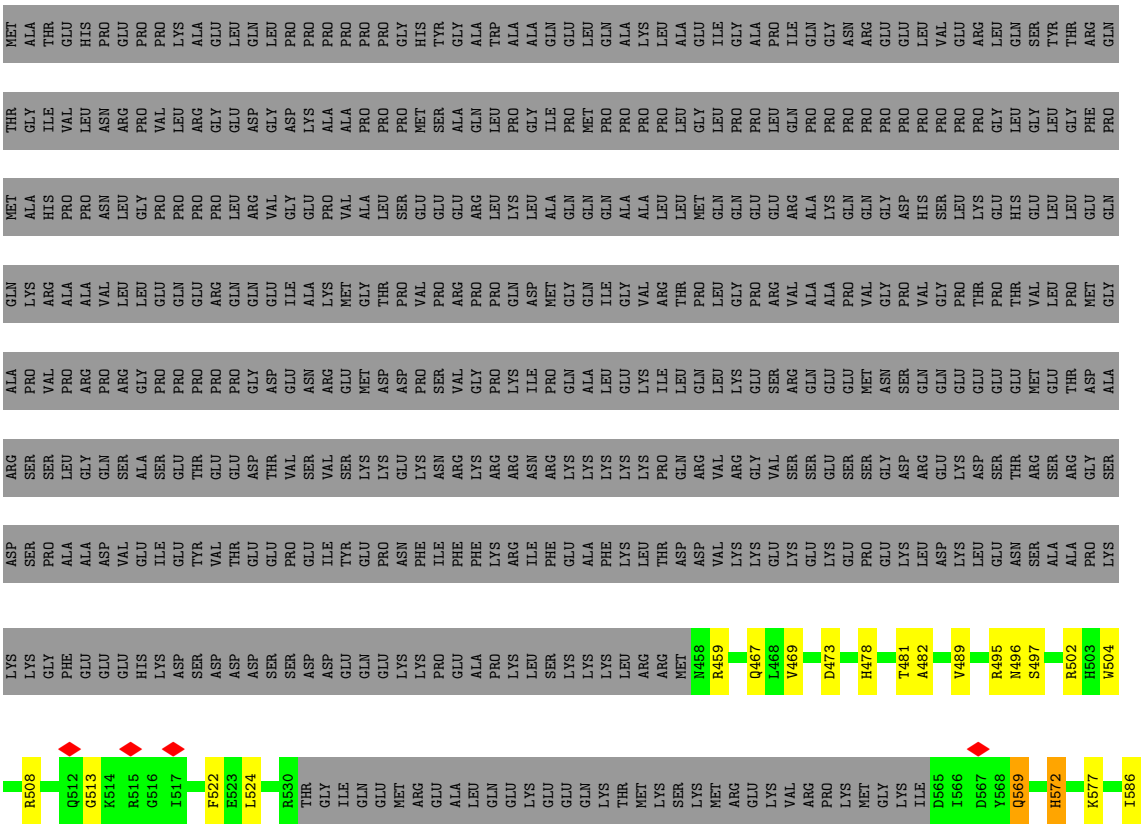
Chain 1:

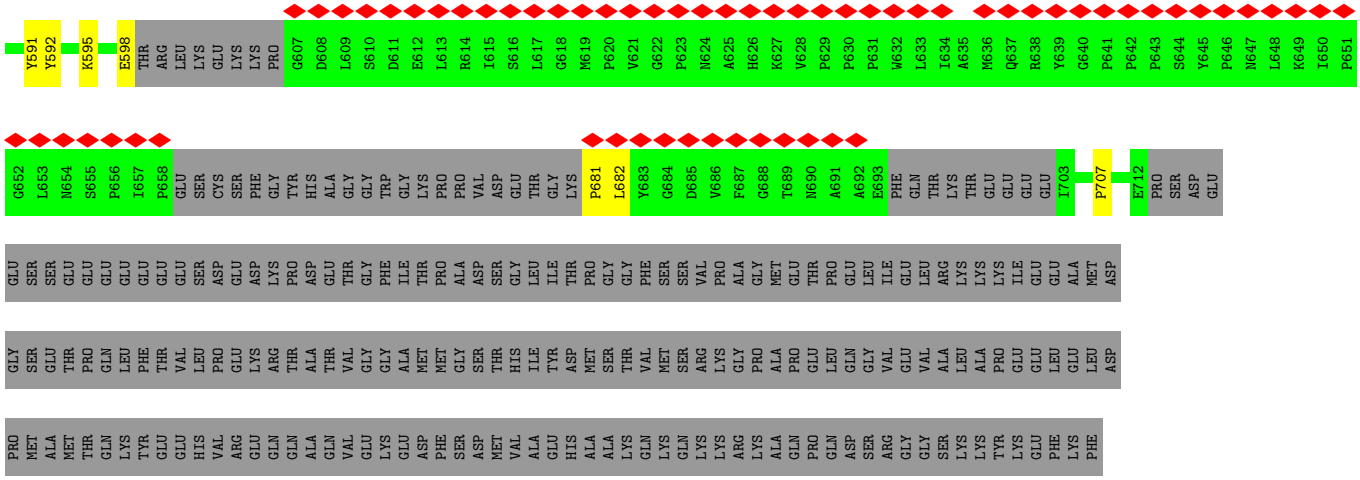




• Molecule 3: Splicing factor 3B subunit 2

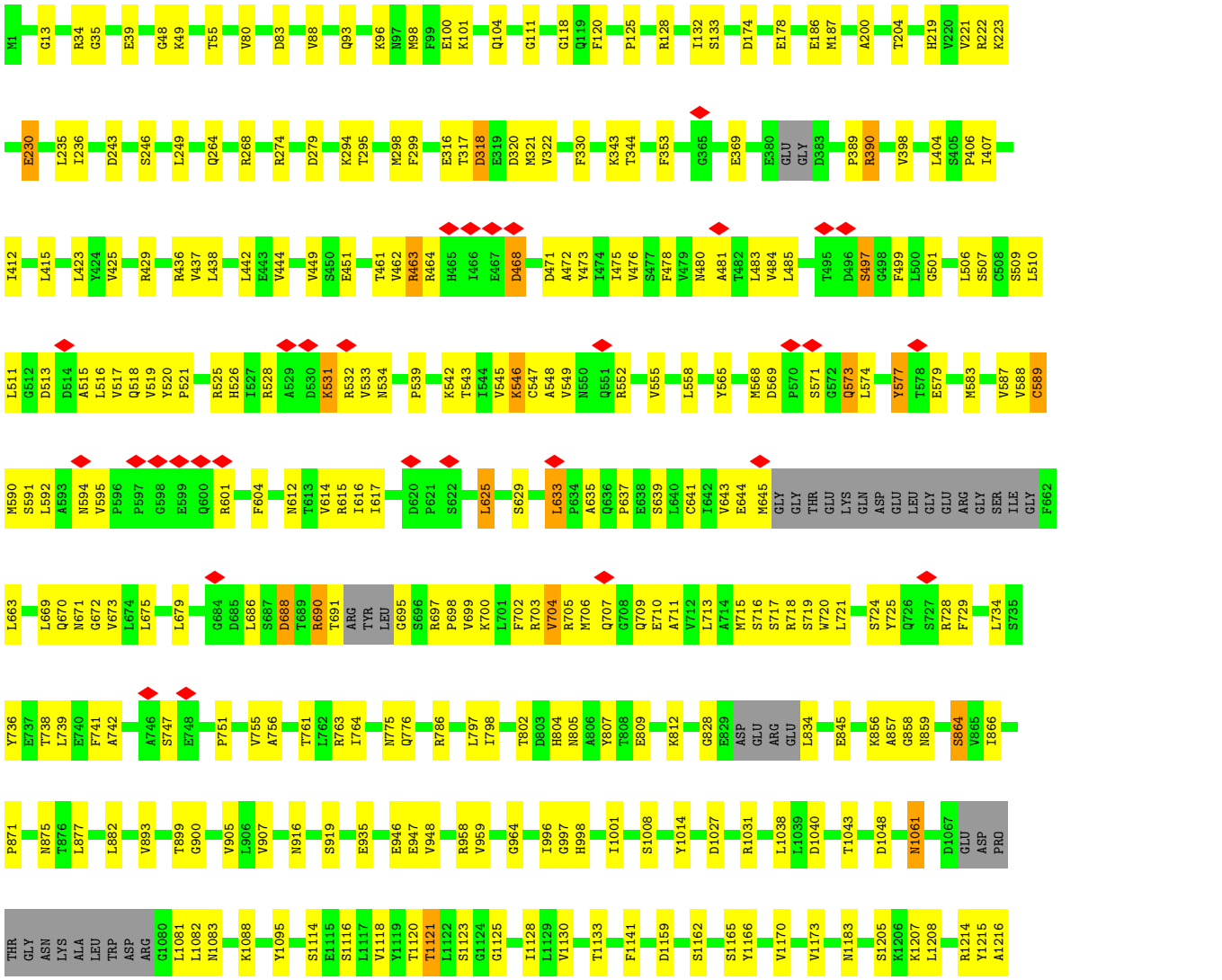
Chain 2:





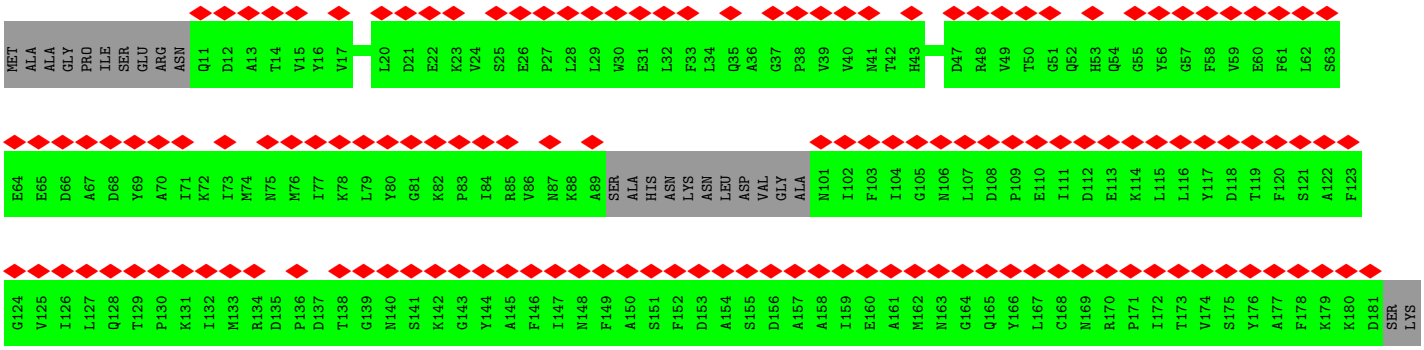
• Molecule 4: Splicing factor 3B subunit 3

Chain 3:



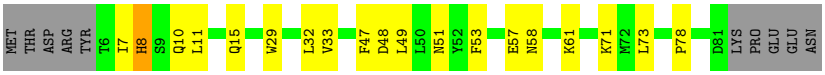
• Molecule 5: Splicing factor 3B subunit 4

Chain 4:



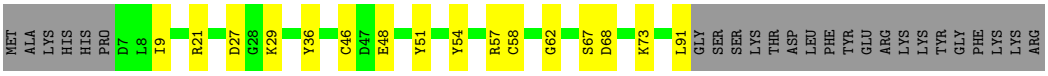
• Molecule 6: Splicing factor 3B subunit 5

Chain 5:



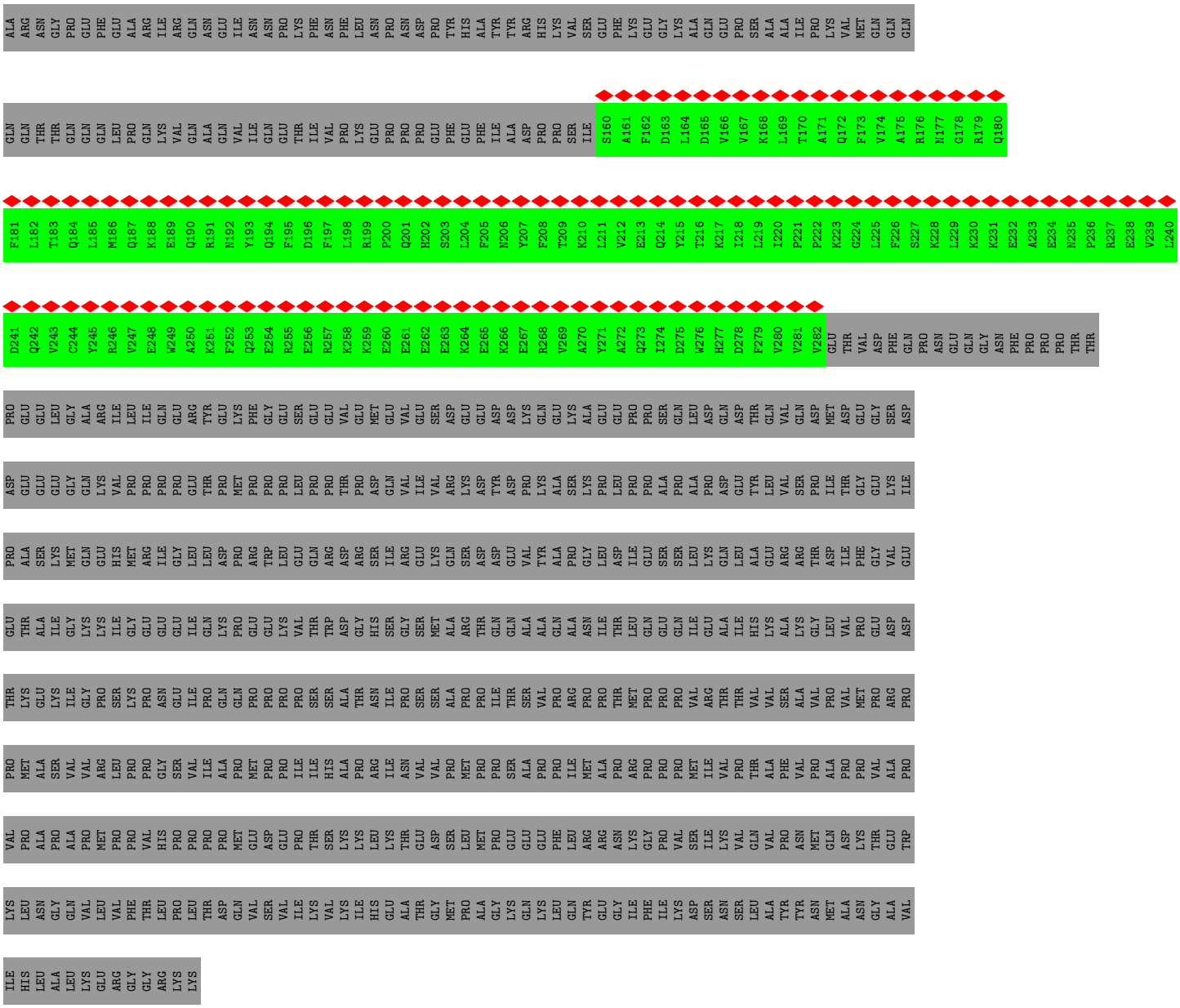
• Molecule 7: PHD finger-like domain-containing protein 5A

Chain 6:



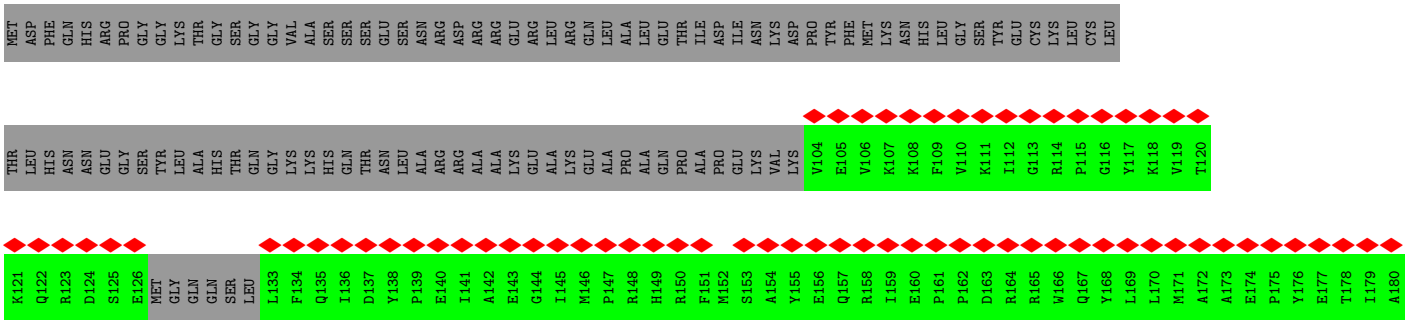
• Molecule 8: Splicing factor 3A subunit 1

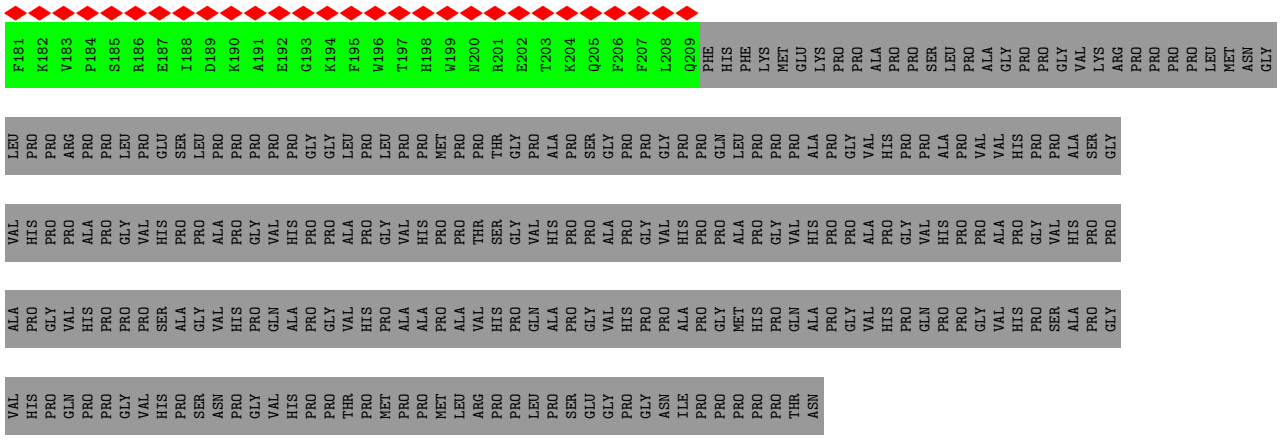
Chain A:



• Molecule 9: Splicing factor 3A subunit 2

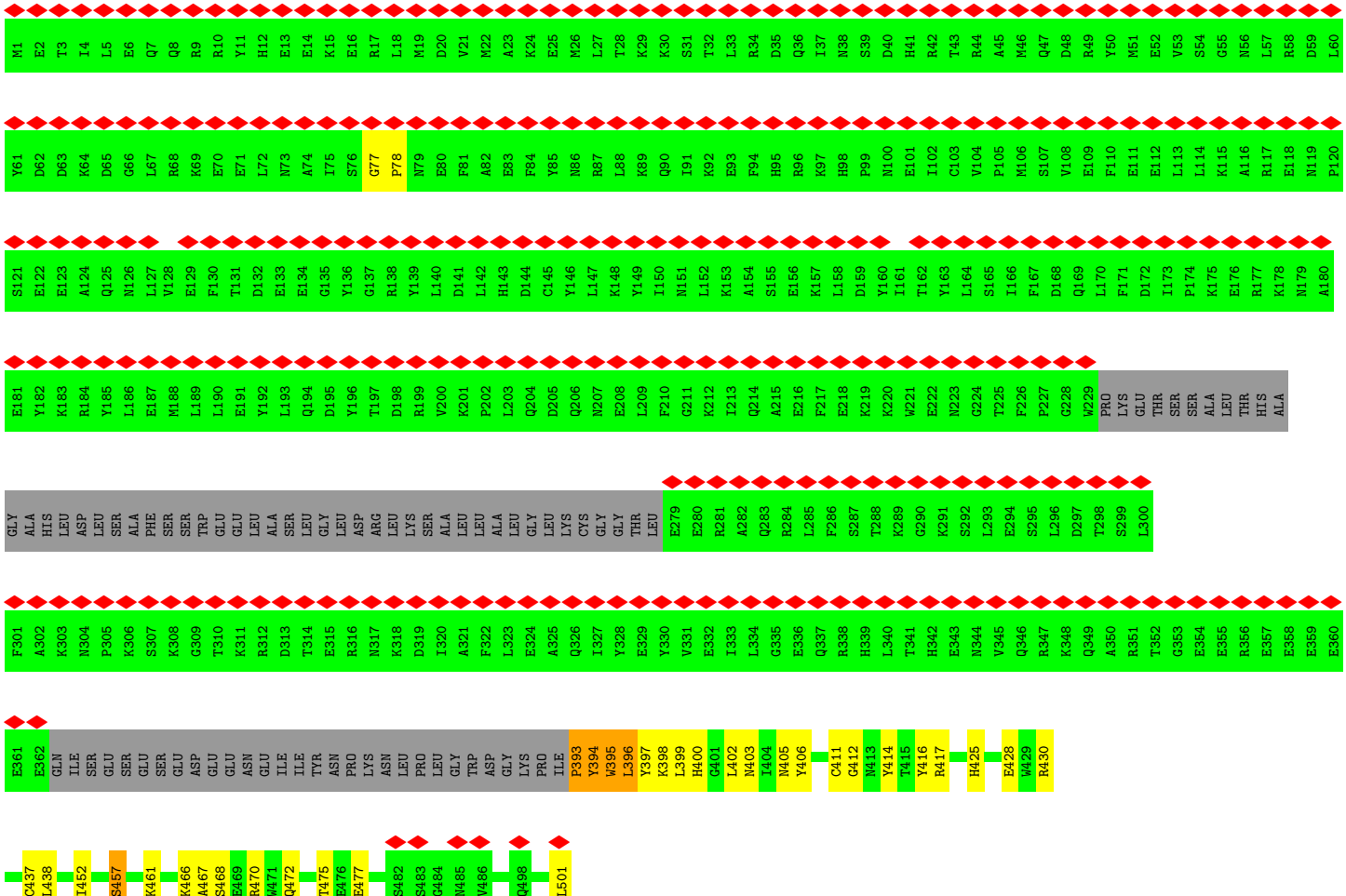
Chain B:





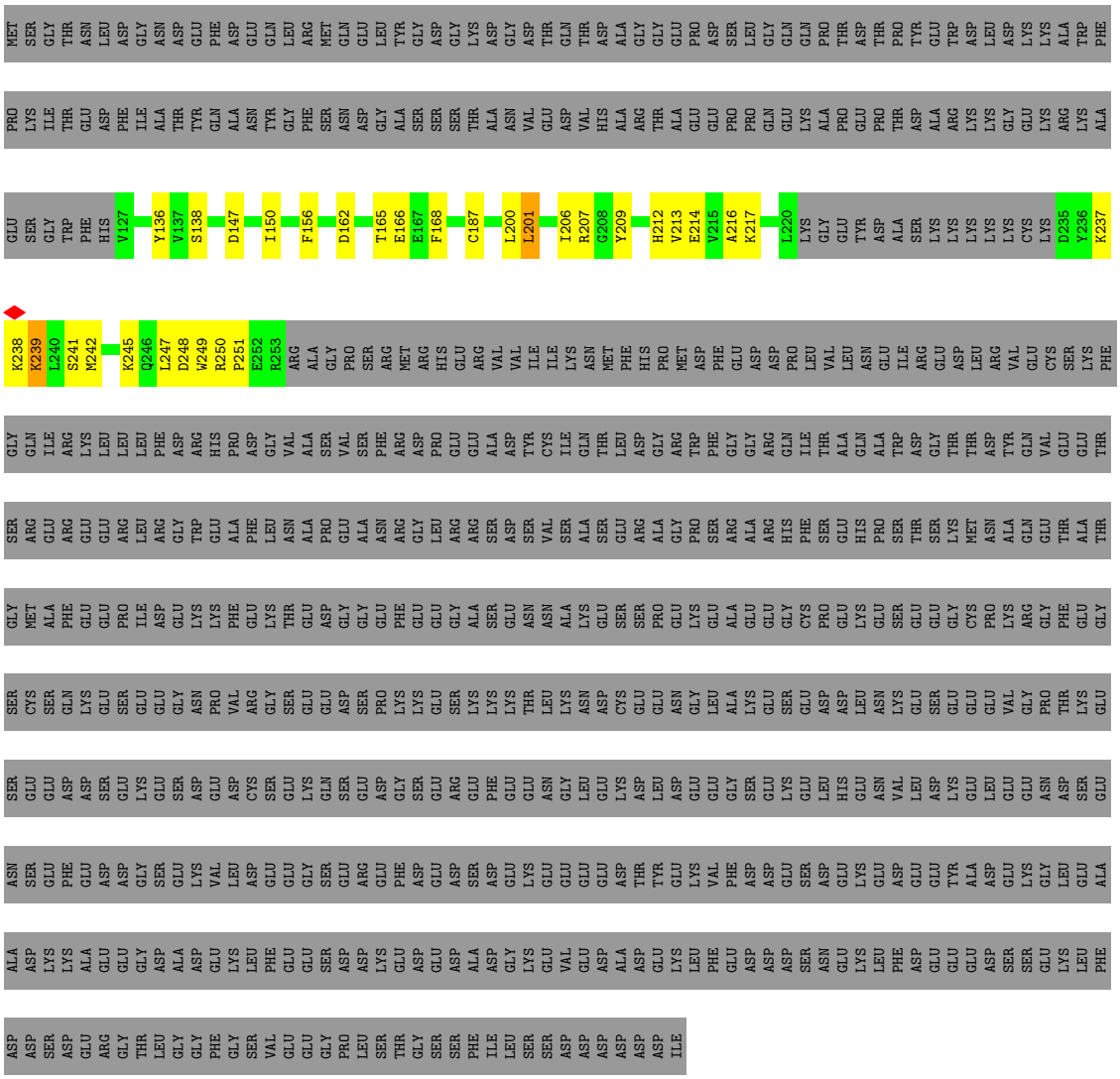
• Molecule 10: Splicing factor 3A subunit 3

Chain C:



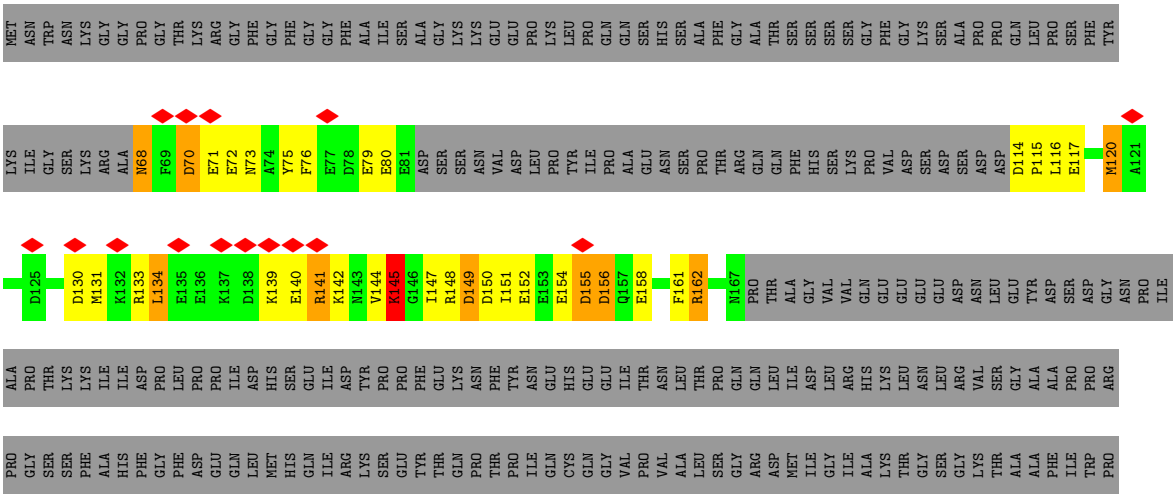
• Molecule 11: HIV Tat-specific factor 1

Chain D:



• Molecule 12: ATP-dependent RNA helicase DDX42

Chain E:



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

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

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

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

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

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

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

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

ALA SER GLY ASN ASN SER ARG GLU GLY THR GLY GLY SER ASN GLY LYS ARG GLU ARG TYR THR GLU ASN ARG GLY SER SER ARG HIS SER HIS GLY LYS GLU THR GLY ASN ARG HIS SER ASP SER PRO ARG HIS GLY ASP GLY GLY ARG HIS GLY ASP GLY GLY TYR ARG HIS PRO GLU SER SER

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

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

• Molecule 13: U2 small nuclear ribonucleoprotein A'

Chain F:

MET V2 K3 L4 T5 A6 E7 L8 I9 E10 Q11 A12 A13 Q14 Y15 T16 N17 A18 V19 R20 D21 R22 E23 L24 D25 L26 R27 G28 Y29 K30 I31 P32 V33 I34 E35 N36 L37 G38 A39 T40 L41 D42 Q43 F44 D45 A46 I47 D48 F49 S50 D51 N52 E53 I54 R55 K56 L57 D58 G59 F60

P61 L62 L63 R64 R65 L66 K67 T68 L69 L70 V71 N72 N73 N74 R75 I76 C77 R78 I79 G80 E81 G82 L83 D84 Q85 A86 L87 P88 C89 L90 T91 E92 L93 I94 L95 T96 N97 N98 S99 L100 V101 E102 L103 G104 D105 L106 D107 P108 L109 A110 S111 L112 K113 S114 L115 T116 Y117 L118 S119 I120

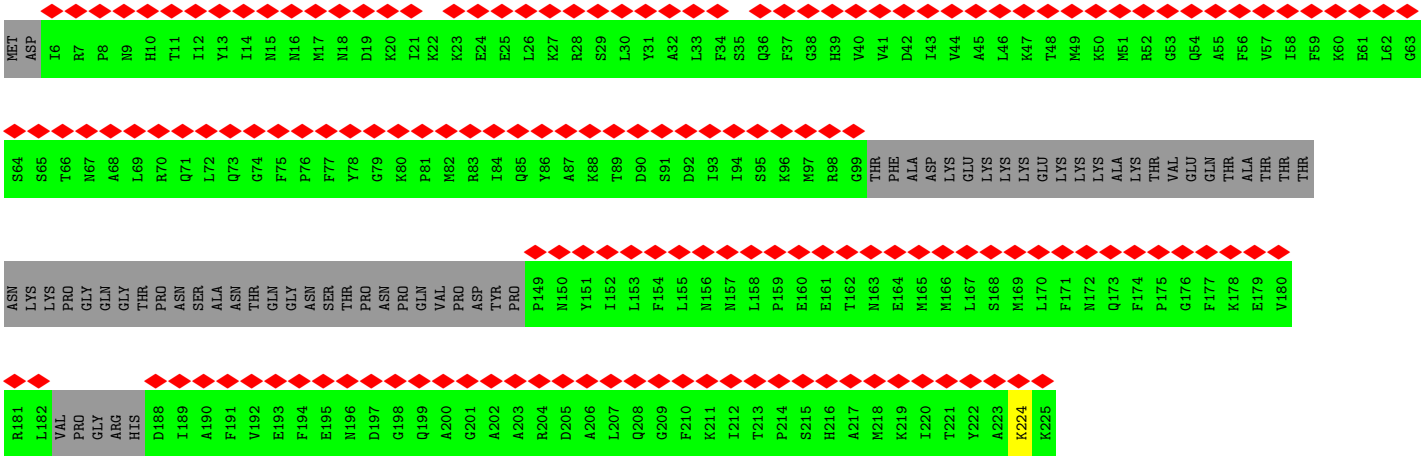
L121 R122 N123 P124 V125 T126 N127 K128 K129 H130 Y131 R132 L133 Y134 V135 I136 Y137 K138 V139 P140 Q141 V142 R143 V144 L145 D146 F147 Q148 K149 V150 K151 L152 K153 E154 R155 Q156 E157 A158 E159 K160 M161 F162 K163 GLY LYS ARG GLY ALA GLN LEU ALA LYS ASP ILE ALA ARG ARG SER LYS THR

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

ASP GLY GLU GLU GLU MET GLU GLU ASP THR VAL THR ASN GLY SER

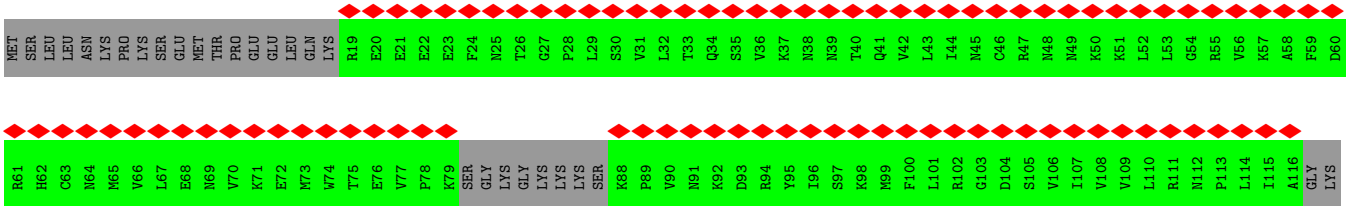
• Molecule 14: U2 small nuclear ribonucleoprotein B"

Chain G:



• Molecule 15: Small nuclear ribonucleoprotein Sm D2

Chain a:



• Molecule 16: Small nuclear ribonucleoprotein F

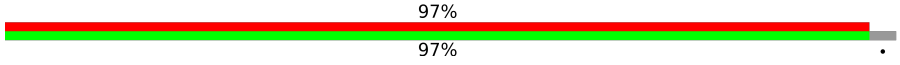
Chain b:

• Molecule 17: Small nuclear ribonucleoprotein E

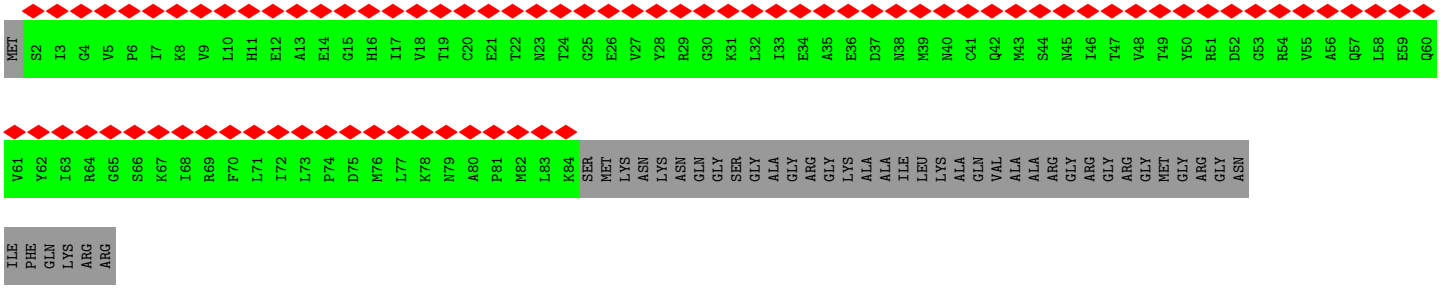
Chain c:

• Molecule 18: Small nuclear ribonucleoprotein G

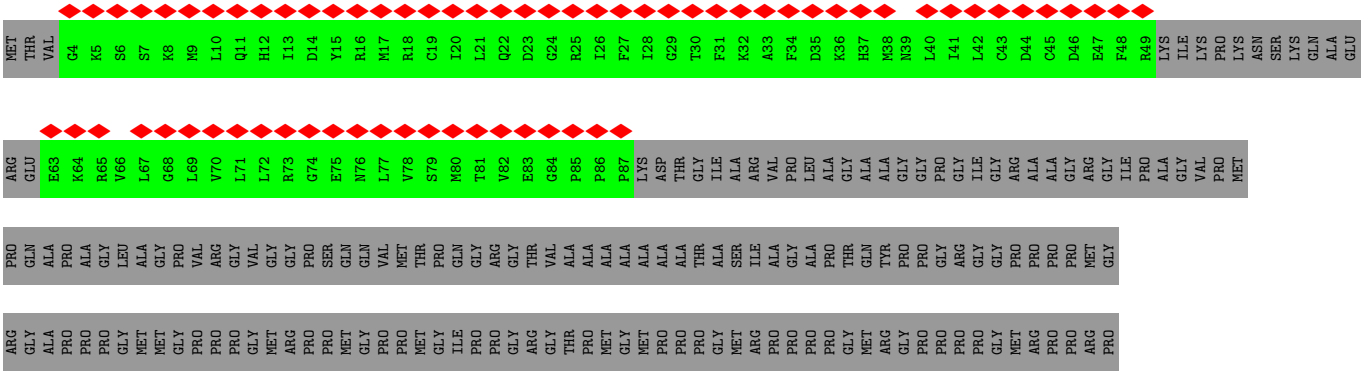
Chain d:



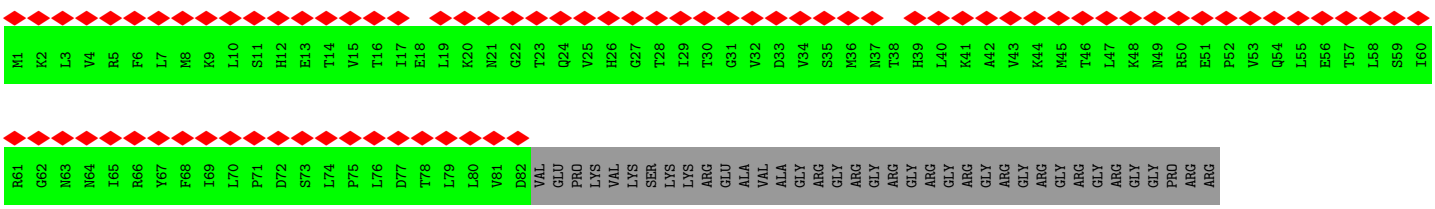
• Molecule 19: Small nuclear ribonucleoprotein Sm D3



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



• Molecule 21: Small nuclear ribonucleoprotein Sm D1



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	430443	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	6.598	Depositor
Minimum map value	-3.978	Depositor
Average map value	0.004	Depositor
Map value standard deviation	0.116	Depositor
Recommended contour level	0.22	Depositor
Map size (Å)	347.84, 347.84, 347.84	wwPDB
Map dimensions	320, 320, 320	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: ZN, 9B0

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	H	0.41	2/3507 (0.1%)	0.95	17/5454 (0.3%)
2	1	0.28	0/6609	0.46	1/8946 (0.0%)
3	2	0.25	0/1248	0.45	0/1672
4	3	0.27	0/9435	0.49	1/12802 (0.0%)
5	4	0.24	0/670	0.45	0/850
6	5	0.51	1/641 (0.2%)	0.54	0/867
7	6	0.26	0/653	0.45	0/877
8	A	0.24	0/507	0.41	0/640
9	B	0.24	0/426	0.50	0/543
10	C	0.26	0/2133	0.45	0/2773
11	D	0.27	0/944	0.46	0/1265
12	E	0.62	0/585	0.62	0/779
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
16	b	0.26	0/311	0.51	0/395
17	c	0.24	0/319	0.49	0/399
18	d	0.25	0/307	0.51	0/388
19	e	0.25	0/343	0.52	0/433
20	f	0.24	0/294	0.49	0/370
21	g	0.24	0/339	0.51	0/428
All	All	0.30	3/31005 (0.0%)	0.56	19/42071 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	H	0	1
10	C	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	5	78	PRO	C-N	8.68	1.50	1.34
1	H	45	C	O3'-P	-7.10	1.52	1.61
1	H	55	U	P-OP2	5.60	1.58	1.49

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	28	C	N1-C1'-C2'	-11.57	98.96	114.00
1	H	120	A	C6-N1-C2	-8.10	113.74	118.60
1	H	42	G	C4'-C3'-O3'	7.34	127.67	113.00
1	H	27	U	C3'-C2'-C1'	7.04	107.14	101.50
1	H	55	U	OP1-P-OP2	6.81	129.82	119.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
10	C	393	PRO	Mainchain
1	H	43	U	Sidechain

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	H	3142	0	1594	72	0
2	1	6487	0	6698	214	0
3	2	1210	0	1010	26	0
4	3	9247	0	9179	167	0
5	4	664	0	230	0	0
6	5	623	0	586	15	0
7	6	645	0	626	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	A	504	0	150	0	0
9	B	421	0	148	0	0
10	C	2103	0	1108	45	0
11	D	930	0	933	36	0
12	E	579	0	507	116	0
13	F	666	0	209	0	0
14	G	685	0	226	7	0
15	a	372	0	118	0	0
16	b	308	0	115	0	0
17	c	319	0	92	0	0
18	d	305	0	105	0	0
19	e	341	0	112	0	0
20	f	293	0	99	0	0
21	g	337	0	110	0	0
22	1	51	0	0	6	0
23	6	3	0	0	0	0
23	C	1	0	0	0	0
All	All	30236	0	23955	567	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 567 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:H:25:G:H1'	10:C:399:LEU:CD2	1.34	1.52
2:1:828:ARG:NH2	12:E:71:GLU:HB3	1.08	1.38
2:1:828:ARG:NH2	12:E:71:GLU:CB	1.87	1.36
2:1:524:ARG:HG2	2:1:562:LYS:O	1.21	1.31
2:1:828:ARG:HH22	12:E:71:GLU:CB	1.41	1.29

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	
2	1	813/1304 (62%)	782 (96%)	31 (4%)	0	100	100
3	2	172/895 (19%)	156 (91%)	15 (9%)	1 (1%)	25	50
4	3	1168/1217 (96%)	1121 (96%)	46 (4%)	1 (0%)	51	78
5	4	156/424 (37%)	153 (98%)	3 (2%)	0	100	100
6	5	74/86 (86%)	73 (99%)	1 (1%)	0	100	100
7	6	83/110 (76%)	79 (95%)	4 (5%)	0	100	100
8	A	121/793 (15%)	121 (100%)	0	0	100	100
9	B	96/464 (21%)	94 (98%)	2 (2%)	0	100	100
10	C	416/501 (83%)	398 (96%)	17 (4%)	1 (0%)	47	73
11	D	109/755 (14%)	94 (86%)	15 (14%)	0	100	100
12	E	64/938 (7%)	57 (89%)	6 (9%)	1 (2%)	9	24
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
16	b	72/86 (84%)	71 (99%)	1 (1%)	0	100	100
17	c	77/92 (84%)	77 (100%)	0	0	100	100
18	d	72/76 (95%)	72 (100%)	0	0	100	100
19	e	81/126 (64%)	79 (98%)	2 (2%)	0	100	100
20	f	67/240 (28%)	67 (100%)	0	0	100	100
21	g	80/119 (67%)	79 (99%)	1 (1%)	0	100	100
All	All	4127/8824 (47%)	3975 (96%)	148 (4%)	4 (0%)	54	78

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
10	C	412	GLY
3	2	496	ASN
12	E	145	LYS
4	3	997	GLY

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
2	1	701/1104 (64%)	665 (95%)	36 (5%)	24 50
3	2	104/776 (13%)	98 (94%)	6 (6%)	20 43
4	3	1022/1051 (97%)	962 (94%)	60 (6%)	19 43
5	4	8/336 (2%)	8 (100%)	0	100 100
6	5	67/77 (87%)	64 (96%)	3 (4%)	27 55
7	6	73/95 (77%)	72 (99%)	1 (1%)	67 86
8	A	4/709 (1%)	4 (100%)	0	100 100
9	B	7/382 (2%)	7 (100%)	0	100 100
10	C	80/446 (18%)	69 (86%)	11 (14%)	3 8
11	D	104/661 (16%)	96 (92%)	8 (8%)	13 30
12	E	60/771 (8%)	48 (80%)	12 (20%)	1 3
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
16	b	4/74 (5%)	4 (100%)	0	100 100
17	c	1/84 (1%)	1 (100%)	0	100 100
18	d	3/66 (4%)	3 (100%)	0	100 100
19	e	3/101 (3%)	3 (100%)	0	100 100
20	f	3/177 (2%)	3 (100%)	0	100 100
21	g	3/101 (3%)	3 (100%)	0	100 100
All	All	2264/7534 (30%)	2127 (94%)	137 (6%)	22 41

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	C	475	THR
11	D	187	CYS
12	E	141	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	3	298	MET
4	3	246	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
4	3	775	ASN
4	3	983	ASN
4	3	776	GLN
4	3	870	ASN
4	3	1052	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	H	143/188 (76%)	27 (18%)	0

5 of 27 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	H	22	U
1	H	23	A
1	H	24	A
1	H	43	U
1	H	47	U

There are no RNA pucker outliers to report.

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)

Of 5 ligands modelled in this entry, 4 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	$\#\|Z\| > 2$
22	9B0	1	1500	-	52,54,54	0.75	2 (3%)	61,77,77	0.66	1 (1%)

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
22	9B0	1	1500	-	-	30/67/91/91	0/3/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
22	1	1500	9B0	C17-C16	-3.50	1.32	1.43
22	1	1500	9B0	C21-C19	-2.50	1.51	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
22	1	1500	9B0	C13-C14-C16	3.69	122.65	119.13

There are no chirality outliers.

5 of 30 torsion outliers are listed below:

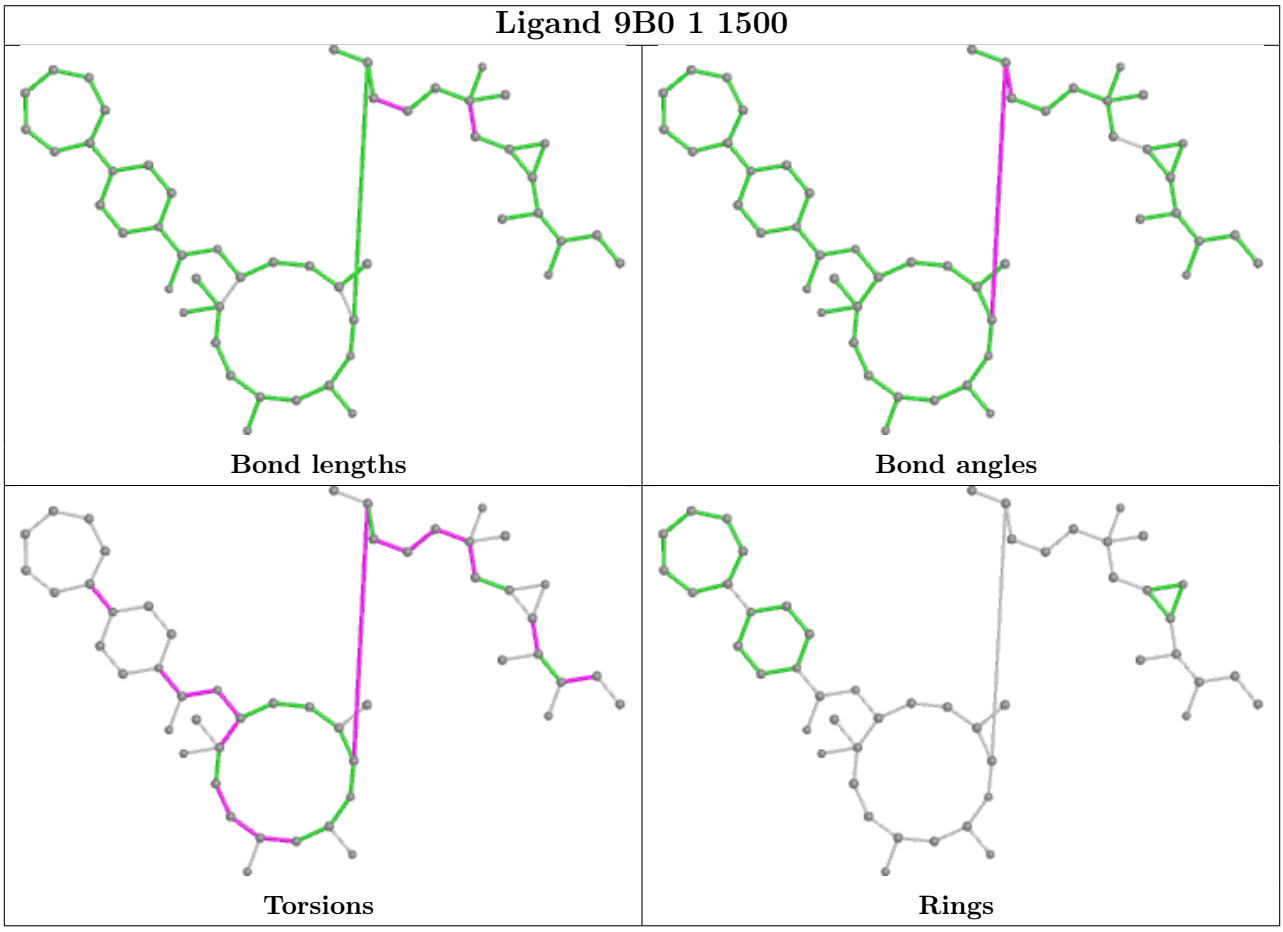
Mol	Chain	Res	Type	Atoms
22	1	1500	9B0	O6-C29-N1-C33
22	1	1500	9B0	O5-C29-N1-C33
22	1	1500	9B0	O6-C29-N1-C30
22	1	1500	9B0	O5-C29-N1-C30
22	1	1500	9B0	N1-C29-O5-C8

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
22	1	1500	9B0	6	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.

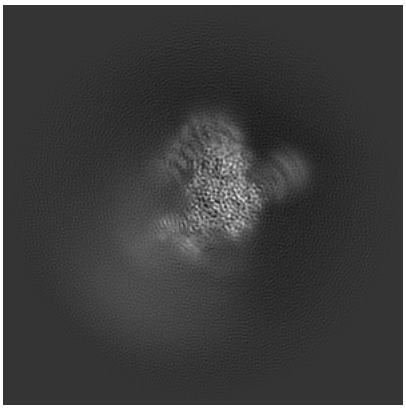
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-34841. These allow visual inspection of the internal detail of the map and identification of artifacts.

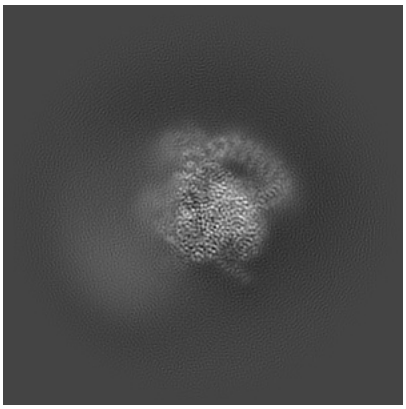
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

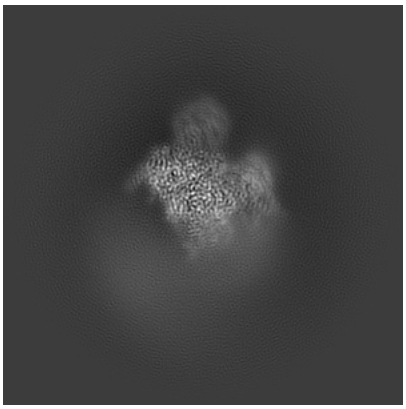
6.1.1 Primary map



X

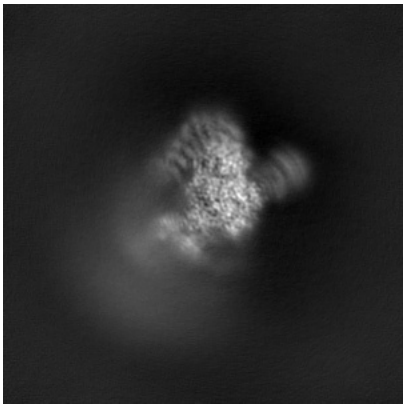


Y

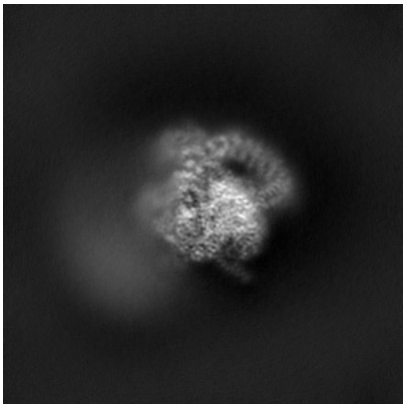


Z

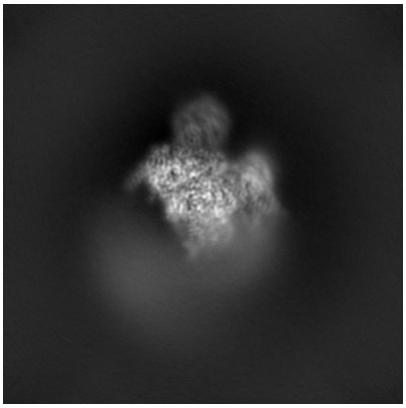
6.1.2 Raw map



X



Y

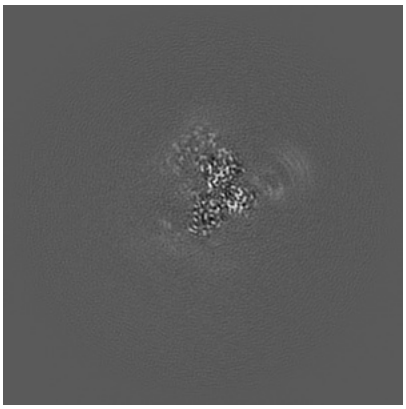


Z

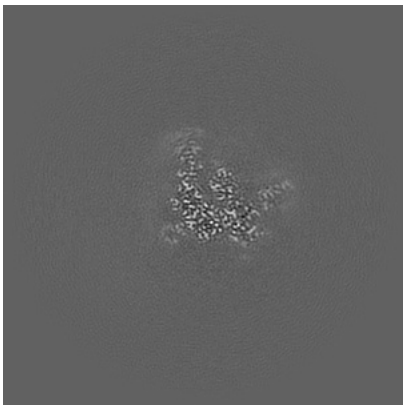
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

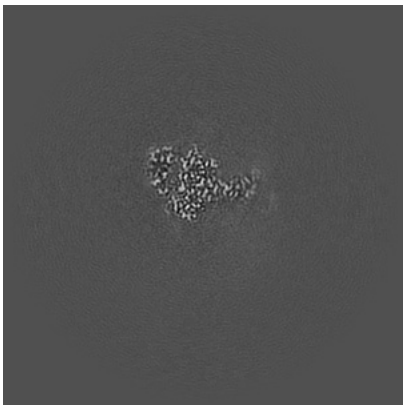
6.2.1 Primary map



X Index: 160

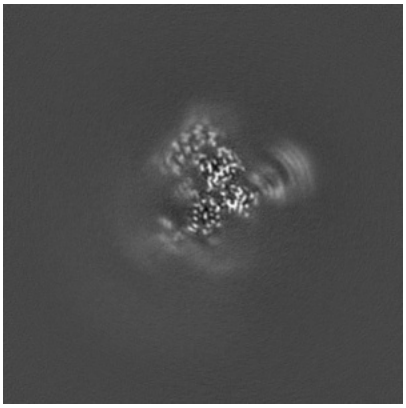


Y Index: 160

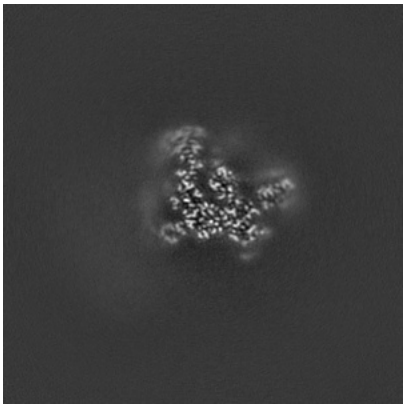


Z Index: 160

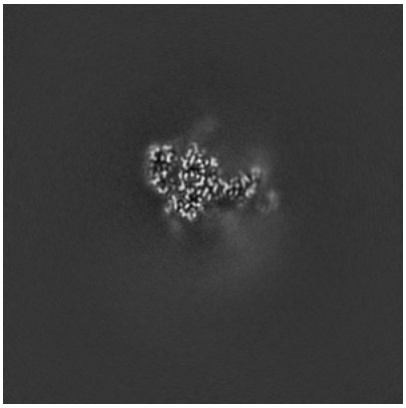
6.2.2 Raw map



X Index: 160



Y Index: 160

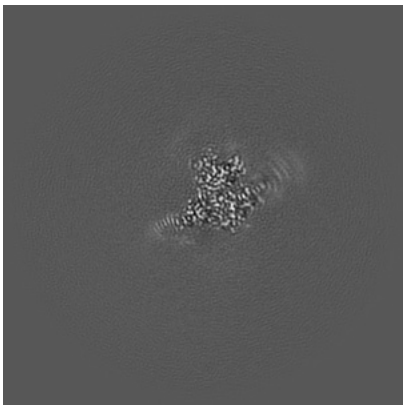


Z Index: 160

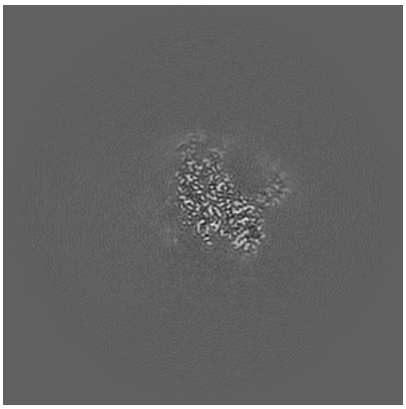
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

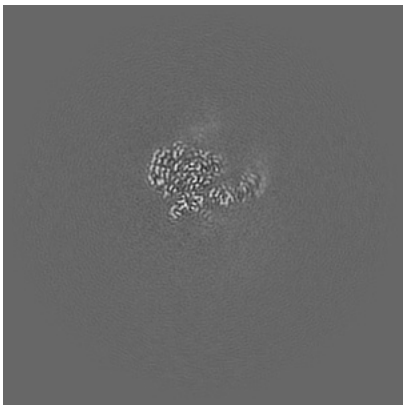
6.3.1 Primary map



X Index: 151

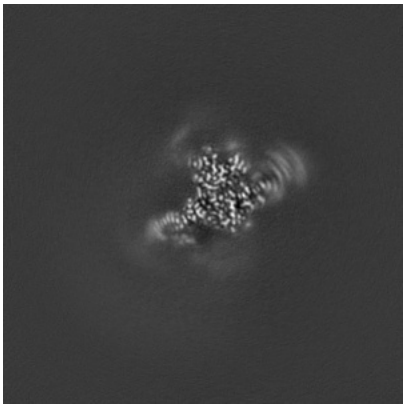


Y Index: 166

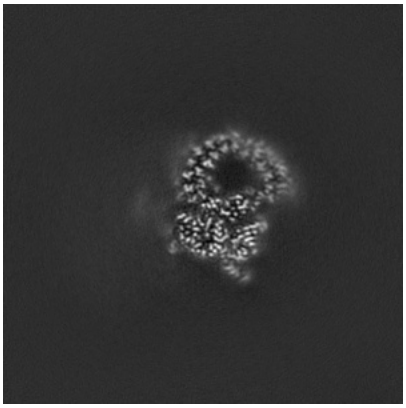


Z Index: 166

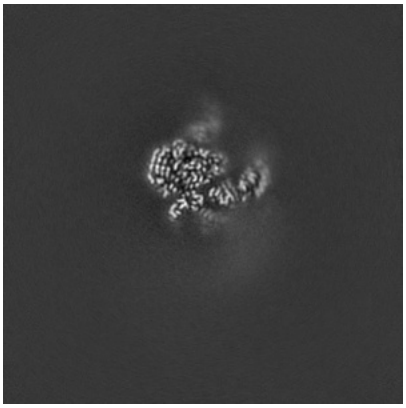
6.3.2 Raw map



X Index: 151



Y Index: 177

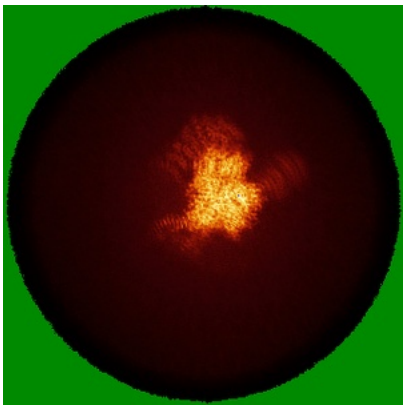


Z Index: 166

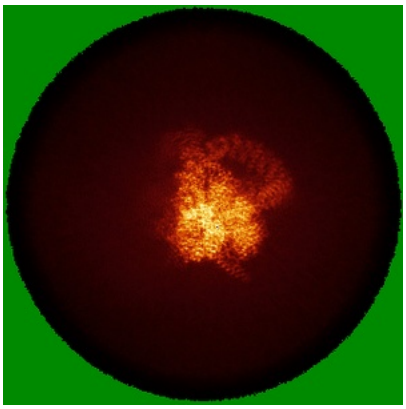
The images above show the largest variance slices of the map in three orthogonal directions.

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

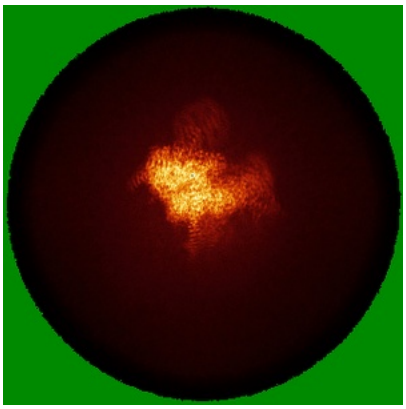
6.4.1 Primary map



X

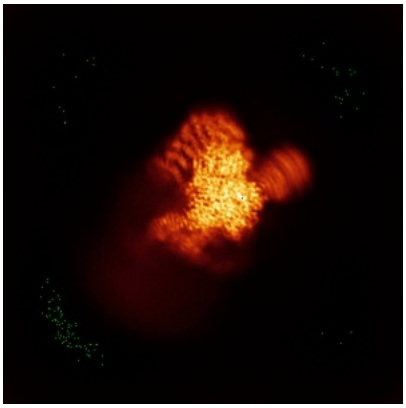


Y

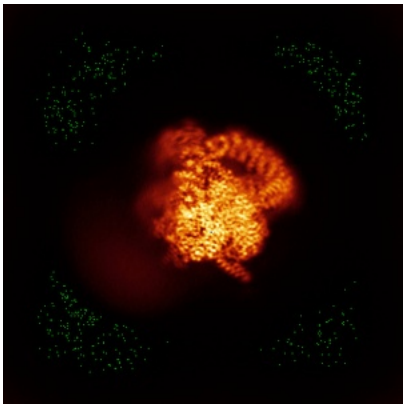


Z

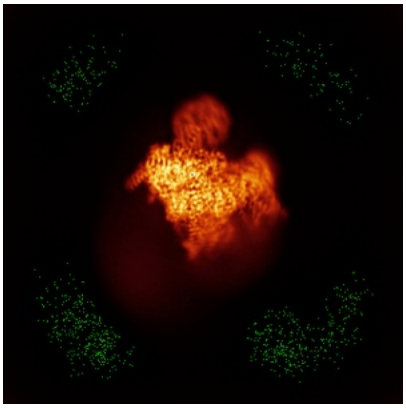
6.4.2 Raw map



X



Y

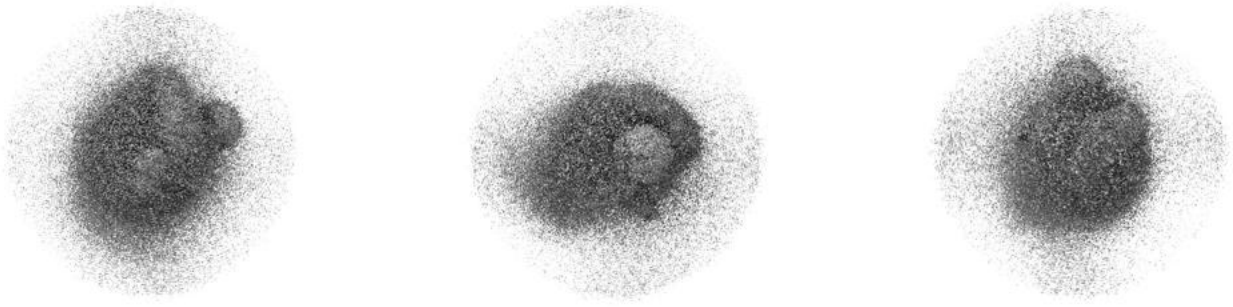


Z

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

6.5.2 Raw map

X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

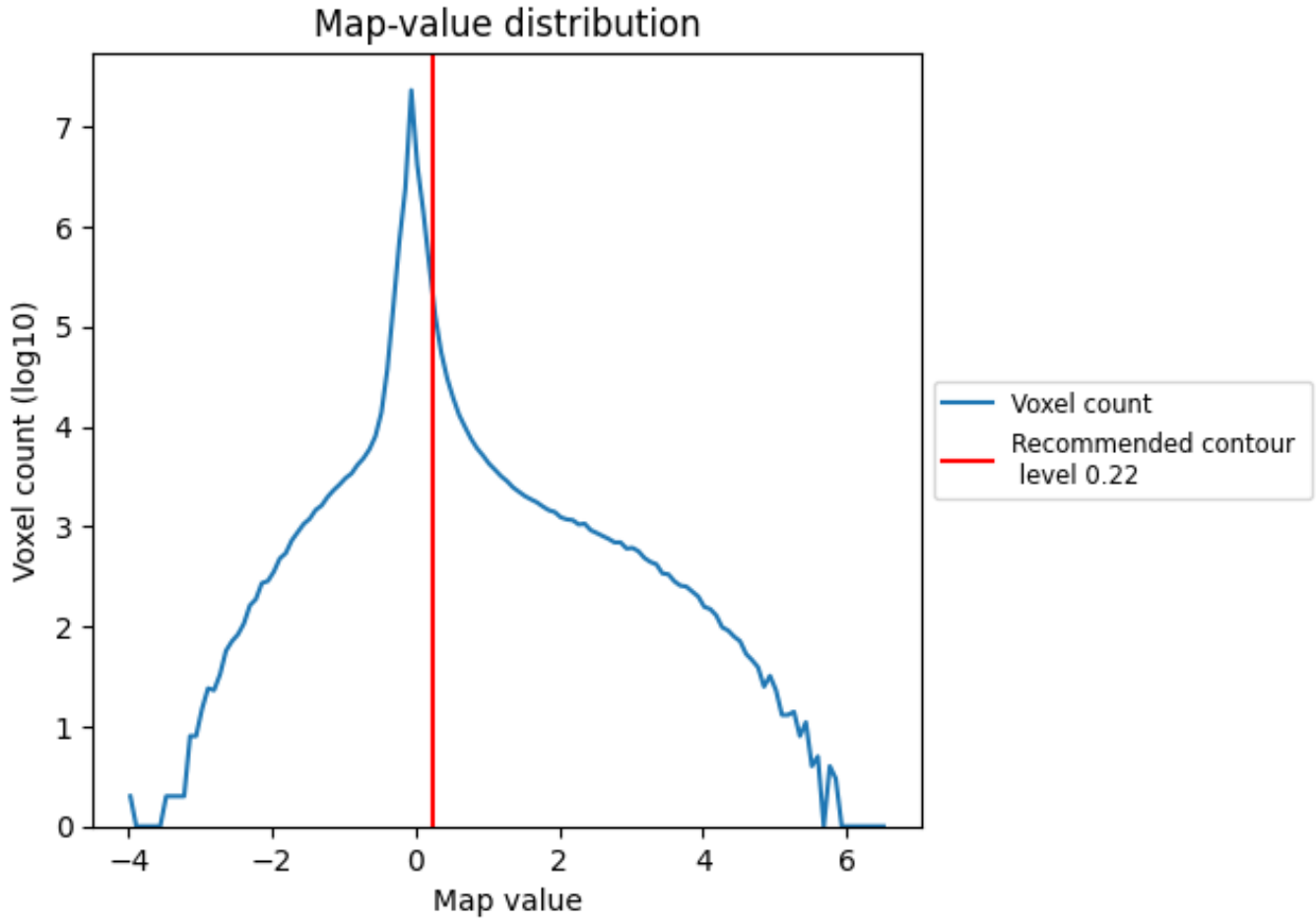
6.6 Mask visualisation (i)

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

7 Map analysis

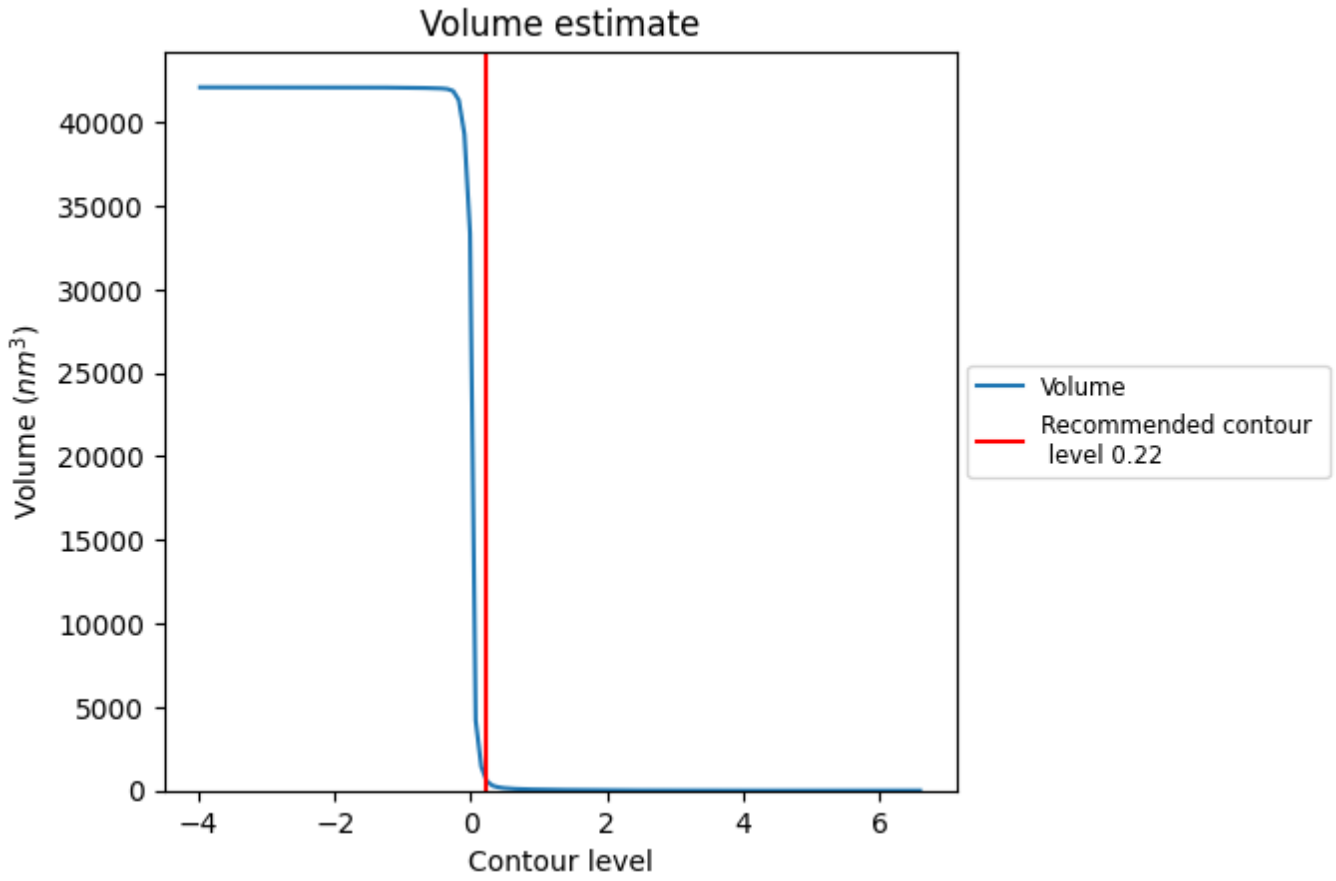
This section contains the results of statistical analysis of the map.

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.

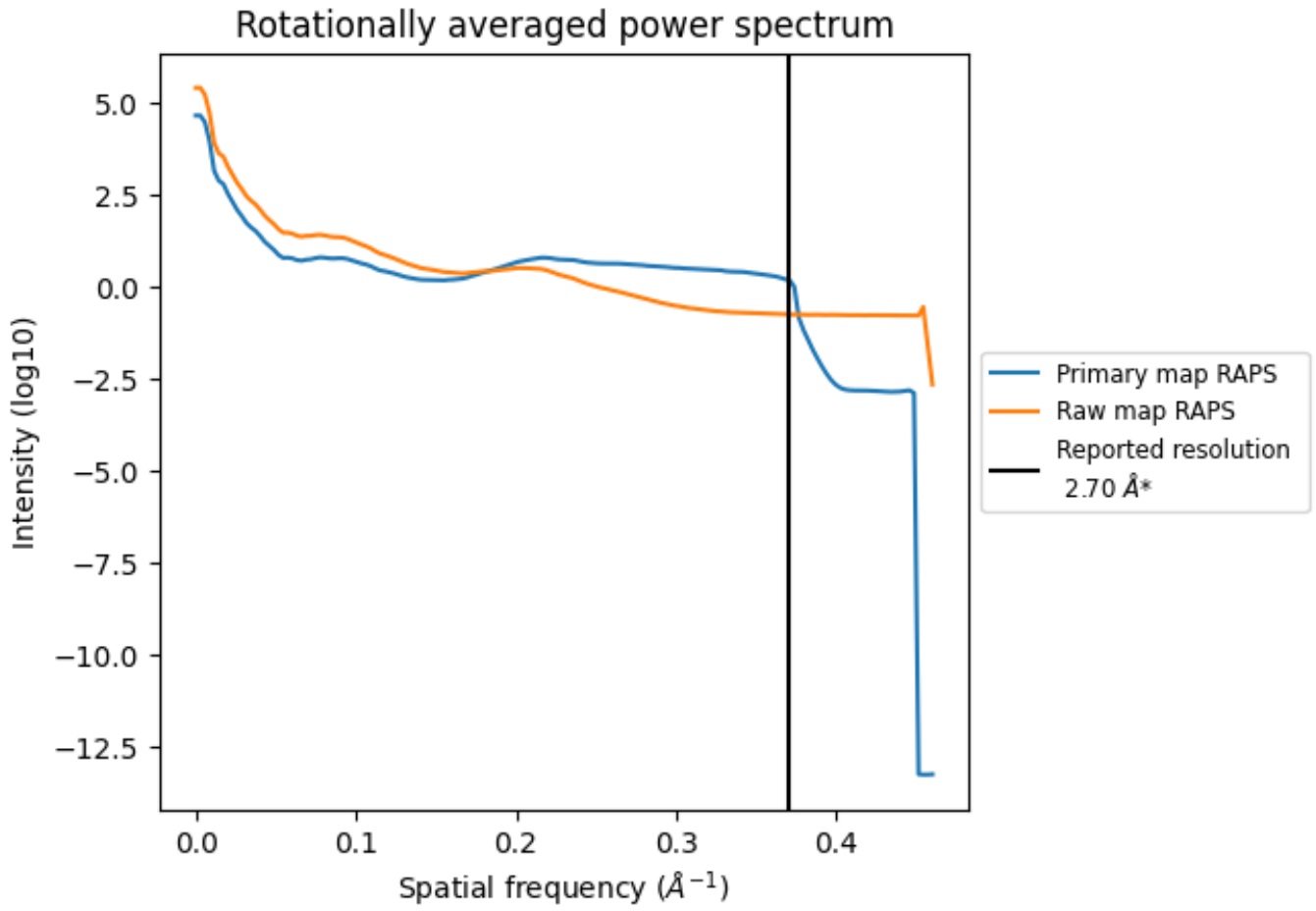
7.2 Volume estimate



The volume at the recommended contour level is 715 nm^3; this corresponds to an approximate mass of 646 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)

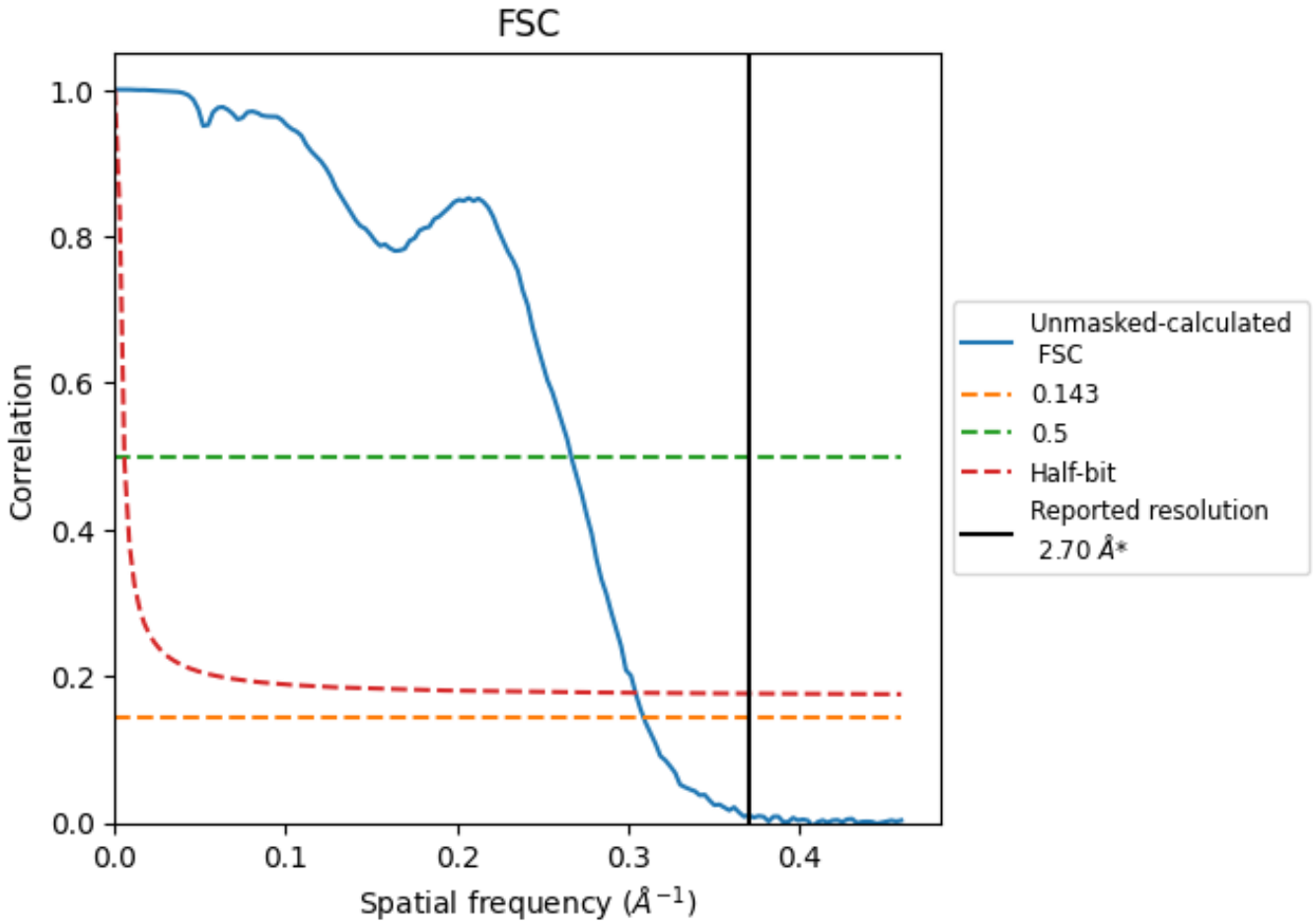


*Reported resolution corresponds to spatial frequency of 0.370 $Å^{-1}$

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.370 Å^{-1}

8.2 Resolution estimates (i)

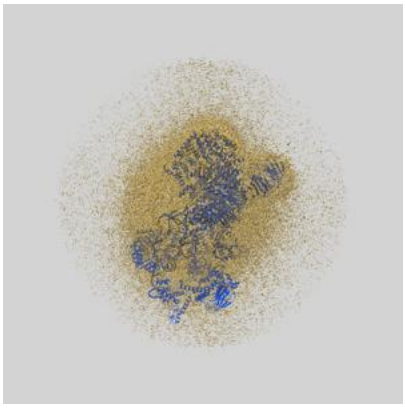
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.70	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.23	3.75	3.28

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.23 differs from the reported value 2.7 by more than 10 %

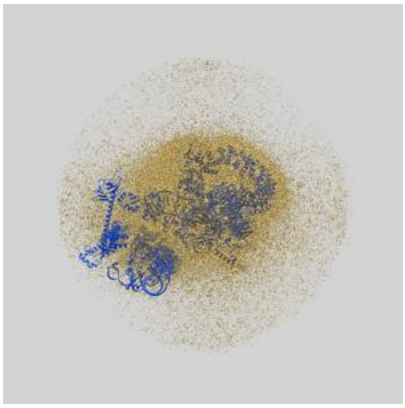
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-34841 and PDB model 8HK1. Per-residue inclusion information can be found in section 3 on page 8.

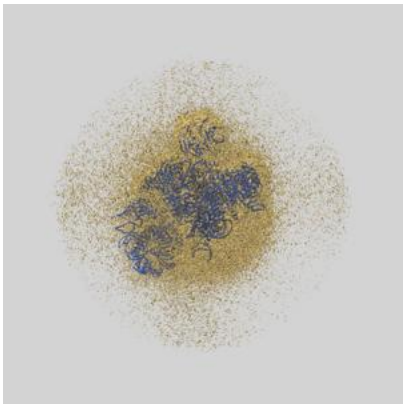
9.1 Map-model overlay



X



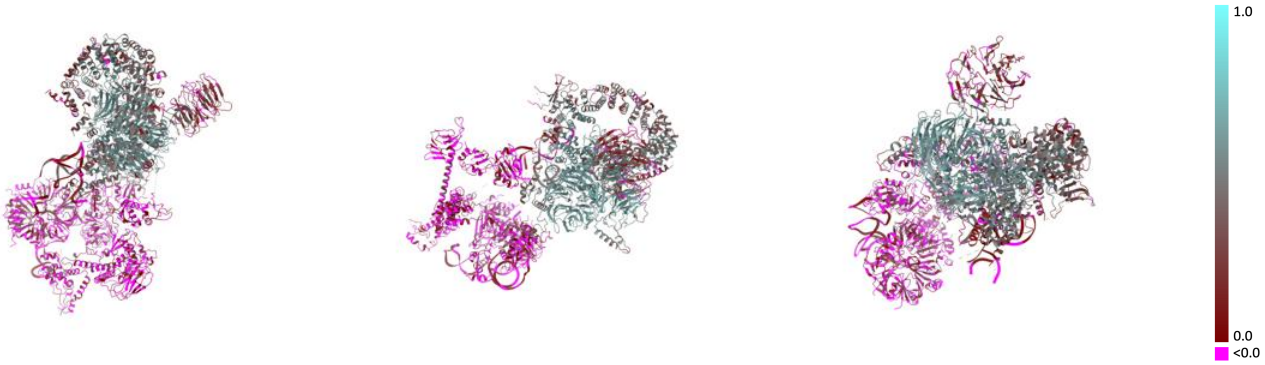
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.22 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.

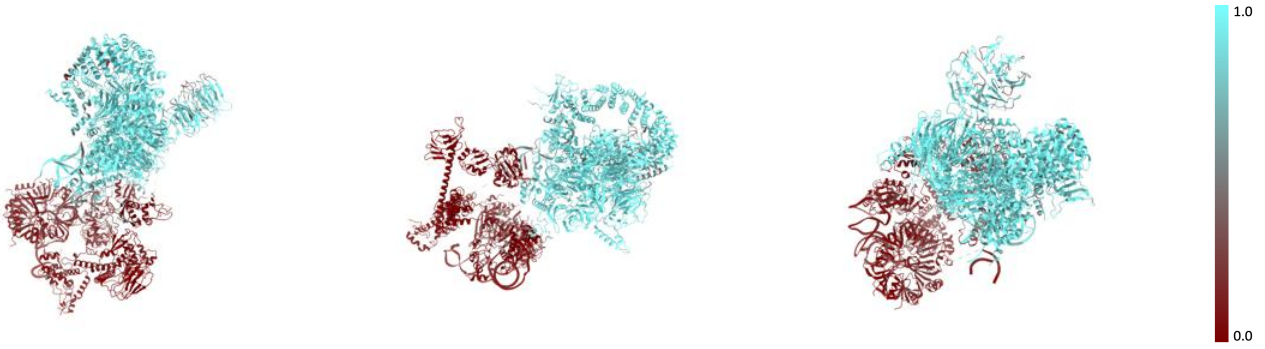
9.2 Q-score mapped to coordinate model



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.

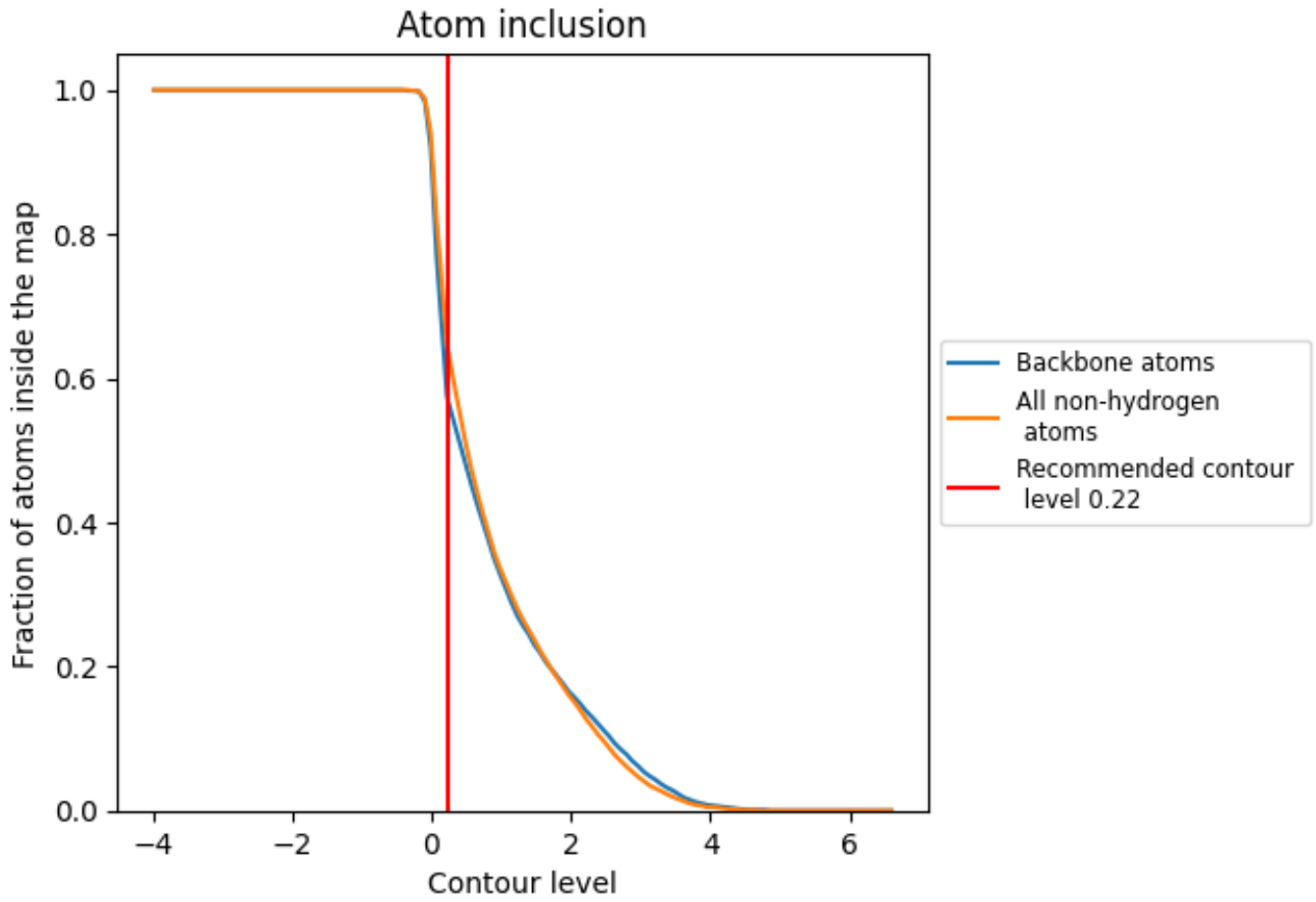
9.3 Atom inclusion mapped to coordinate model



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 (0.22).

9.4 Atom inclusion (i)



At the recommended contour level, 58% of all backbone atoms, 65% 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 (0.22) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6490	0.3170
1	0.9420	0.4610
2	0.6950	0.3390
3	0.9240	0.4850
4	0.1220	-0.0090
5	0.9610	0.5880
6	0.9700	0.5760
A	0.0060	-0.0030
B	0.0070	-0.0200
C	0.3430	0.1520
D	0.8790	0.3320
E	0.6180	0.2010
F	0.0190	-0.0280
G	0.0250	0.0070
H	0.2810	0.0690
a	0.0270	-0.0030
b	0.0230	0.0070
c	0.0090	-0.0060
d	0.0230	0.0050
e	0.0260	0.0120
f	0.0440	0.0470
g	0.0300	-0.0190