Full wwPDB EM Validation Report

Dec 9, 2024 – 02:58 PM EST

PDB ID	:	9EGX
EMDB ID	:	EMD-48039
Title	:	RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-hexasome, bp +27
Authors	:	Markert, J.; Farnung, L.
Deposited on	:	2024-11-21
Resolution	:	2.90 Å(reported)

This is a Full wwPDB EM Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

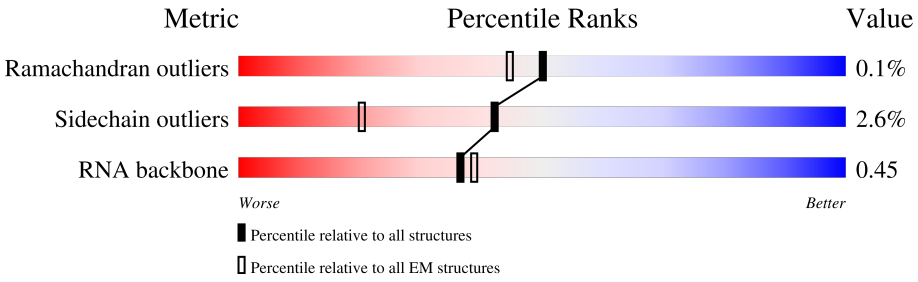
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.90 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.



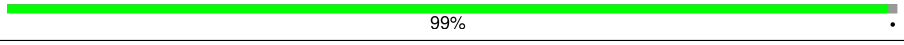
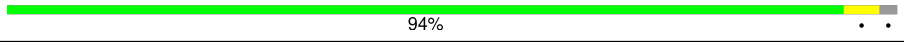
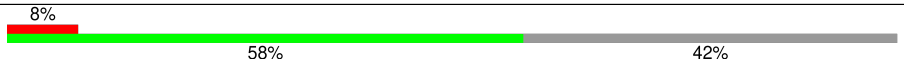
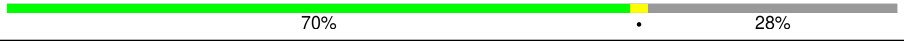
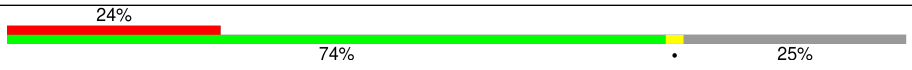
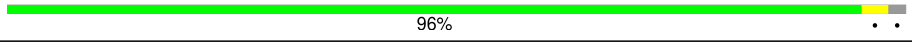
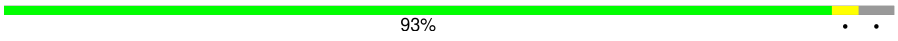
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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

Mol	Chain	Length	Quality of chain
1	A	1984	71% · 28%
2	B	1251	88% · 10%
3	C	275	92% · 6%
4	D	142	87% · 11%
5	E	210	100%
6	F	127	58% · 39%
7	G	172	97%
8	H	150	94% 5% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	
10	J	67	
11	K	117	
12	L	58	
13	M	1729	
14	N	206	
15	O	821	
16	P	21	
17	Q	1179	
18	R	713	
19	S	304	
20	T	215	
21	U	666	
22	V	531	
23	W	305	
24	X	531	
25	Y	121	
26	Z	1087	
27	a	136	
27	e	136	
28	b	103	
28	f	103	
29	c	130	
30	d	123	

2 Entry composition (i)

There are 32 unique types of molecules in this entry. The entry contains 65003 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 DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
1	A	1426	11210	7040	2013	2086	2	69	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1122	8980	5684	1576	1656	64	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	258	2072	1300	356	410	6	0	0

- Molecule 4 is a protein called RNA polymerase Rpb4/RPC9 core domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	126	1004	630	170	200	4	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	209	1720	1089	300	323	8	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	78	626	401	106	114	5	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	171	Total	C	N	O	S	0	0
			1333	866	214	245	8		

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	149	Total	C	N	O	S	0	0
			1197	759	195	238	5		

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	116	Total	C	N	O	S	0	0
			942	582	168	181	11		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	66	Total	C	N	O	S	0	0
			524	339	88	91	6		

- Molecule 11 is a protein called RNA polymerase II subunit J.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	0
			920	593	152	173	2		

- Molecule 12 is a protein called RNA polymerase II subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	47	Total	C	N	O	S	0	0
			397	246	77	68	6		

- Molecule 13 is a protein called Transcription elongation factor SPT6.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	M	1002	Total	C	N	O	0	0
			4309	2295	1003	1011		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-2	SER	-	expression tag	UNP Q7KZ85
M	-1	ASN	-	expression tag	UNP Q7KZ85
M	0	ALA	-	expression tag	UNP Q7KZ85

- Molecule 14 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	148	Total	C	N	O	P	0	0
			3048	1446	555	899	148		

- Molecule 15 is a protein called Protein IWS1 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	O	132	Total	C	N	O	0	0
			656	392	132	132		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	-1	SER	-	expression tag	UNP Q96ST2
O	0	ASN	-	expression tag	UNP Q96ST2
O	1	ALA	-	expression tag	UNP Q96ST2

- Molecule 16 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	P	21	Total	C	N	O	P	0	0
			432	193	59	159	21		

- Molecule 17 is a protein called RNA polymerase-associated protein CTR9 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	890	Total	C	N	O	S	0	0
			6427	4026	1164	1218	19		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	1174	GLU	-	expression tag	UNP Q6PD62
Q	1175	ASN	-	expression tag	UNP Q6PD62
Q	1176	LEU	-	expression tag	UNP Q6PD62
Q	1177	TYR	-	expression tag	UNP Q6PD62

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Q	1178	PHE	-	expression tag	UNP Q6PD62
Q	1179	GLN	-	expression tag	UNP Q6PD62

- Molecule 18 is a protein called RNA polymerase-associated protein RTF1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	244	Total	C	N	O	S	0	0
			1428	866	281	280	1		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	-2	SER	-	expression tag	UNP Q92541
R	-1	ASN	-	expression tag	UNP Q92541
R	0	ALA	-	expression tag	UNP Q92541

- Molecule 19 is a protein called Transcription elongation factor A protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	S	161	Total	C	N	O	0	0
			657	334	161	162		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	-2	SER	-	expression tag	UNP P23193
S	-1	ASN	-	expression tag	UNP P23193
S	0	ALA	-	expression tag	UNP P23193

- Molecule 20 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	159	Total	C	N	O	P	0	0
			3245	1540	608	939	158		

- Molecule 21 is a protein called RNA polymerase-associated protein LEO1.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	U	125	Total	C	N	O	0	0
			617	367	125	125		

- Molecule 22 is a protein called RNA polymerase II-associated factor 1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	V	244	Total	C	N	O	S	0	0
			1378	842	267	267	2		

- Molecule 23 is a protein called WDR61.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	300	Total	C	N	O	S	0	0
			2333	1483	392	454	4		

- Molecule 24 is a protein called Parafibromin.

Mol	Chain	Residues	Atoms				AltConf	Trace
24	X	43	Total	C	N	O	0	0
			353	220	69	64		

- Molecule 25 is a protein called Transcription elongation factor SPT4.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	Y	116	Total	C	N	O	S	0	0
			911	570	159	173	9		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Y	-3	GLY	-	expression tag	UNP P63272
Y	-2	PRO	-	expression tag	UNP P63272
Y	-1	GLY	-	expression tag	UNP P63272
Y	0	SER	-	expression tag	UNP P63272

- Molecule 26 is a protein called Transcription elongation factor SPT5.

Mol	Chain	Residues	Atoms						AltConf	Trace
26	Z	510	Total	C	N	O	P	S	0	0
			4025	2552	709	745	1	18		

- Molecule 27 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	a	77	Total	C	N	O	S	0	0
			627	397	116	111	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
27	e	97	Total	C	N	O	S	0	0
			801	504	155	139	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	36	MET	LYS	engineered mutation	UNP A0A310TTQ1
e	36	MET	LYS	engineered mutation	UNP A0A310TTQ1

- Molecule 28 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	b	83	Total	C	N	O	S	0	0
			662	418	129	114	1		
28	f	78	Total	C	N	O	S	0	0
			619	391	120	107	1		

- Molecule 29 is a protein called Histone H2A type 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
29	c	103	Total	C	N	O	0	0
			795	501	155	139		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	99	ARG	GLY	conflict	UNP P06897
c	123	SER	ALA	conflict	UNP P06897

- Molecule 30 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	d	95	Total	C	N	O	S	0	0
			745	469	134	140	2		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	0	MET	-	initiating methionine	UNP P02281
d	29	THR	SER	engineered mutation	UNP P02281

- Molecule 31 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
31	A	2	Total Zn 2 2	0
31	B	1	Total Zn 1 1	0
31	C	1	Total Zn 1 1	0
31	I	2	Total Zn 2 2	0
31	J	1	Total Zn 1 1	0
31	L	1	Total Zn 1 1	0
31	Y	1	Total Zn 1 1	0

- Molecule 32 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

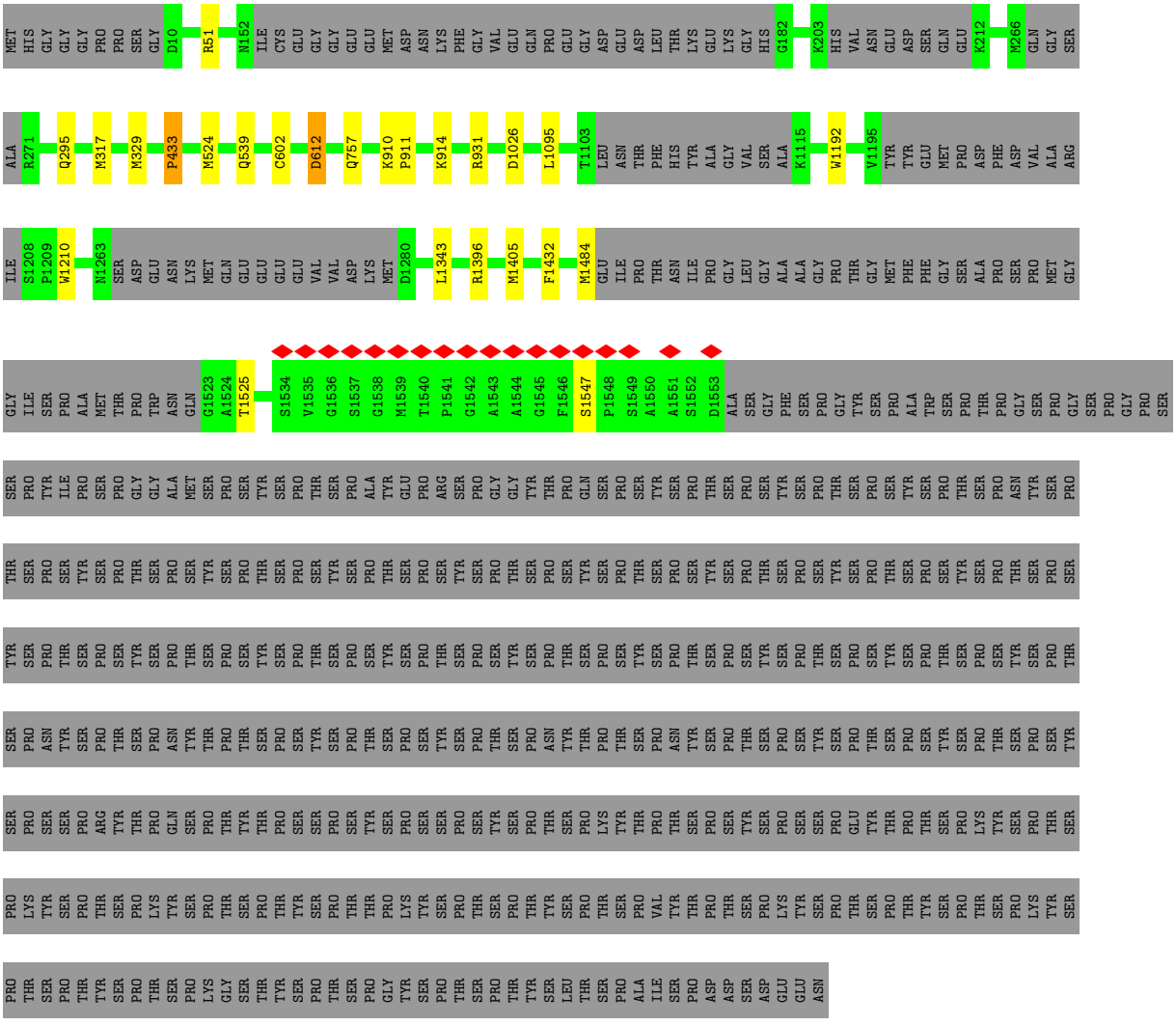
Mol	Chain	Residues	Atoms	AltConf
32	A	1	Total Mg 1 1	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

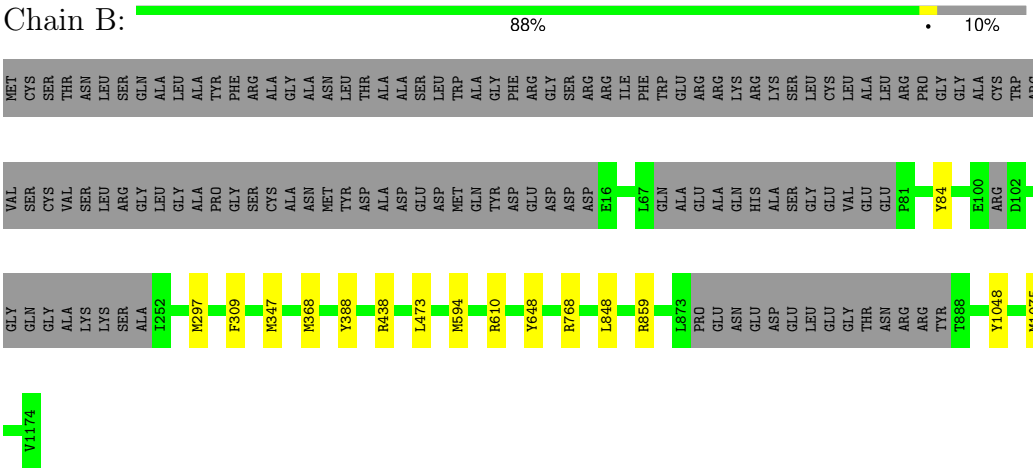
- Molecule 1: DNA-directed RNA polymerase subunit

Chain A:



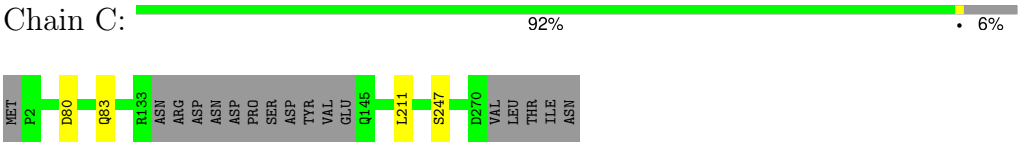
- Molecule 2: DNA-directed RNA polymerase subunit beta

Chain B:



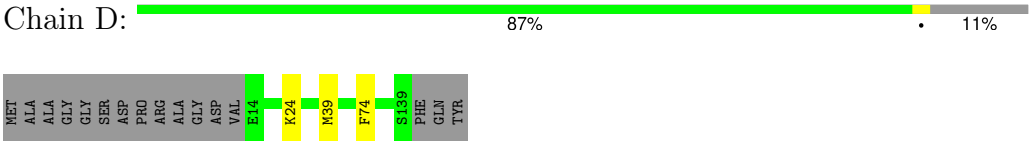
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



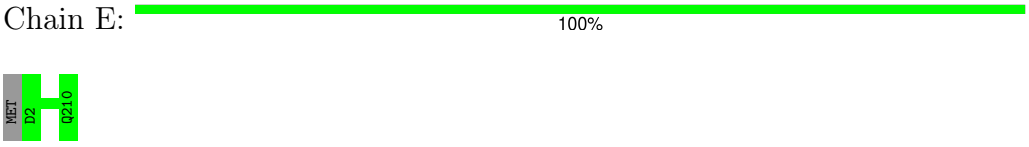
• Molecule 4: RNA polymerase Rpb4/RPC9 core domain-containing protein

Chain D:



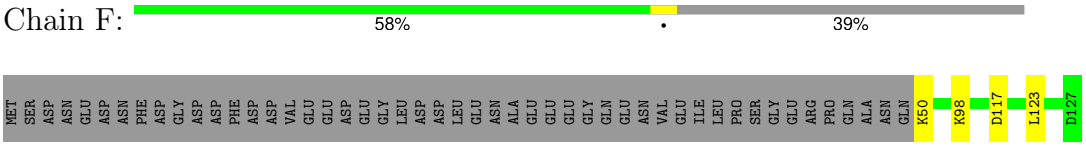
• Molecule 5: DNA-directed RNA polymerase II subunit E

Chain E:



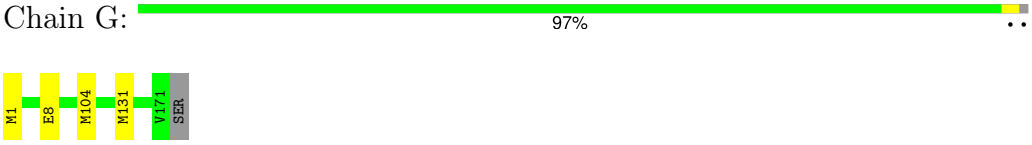
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



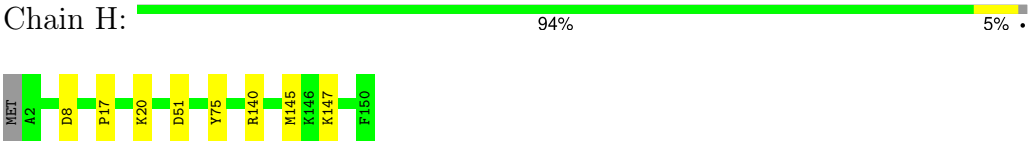
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:



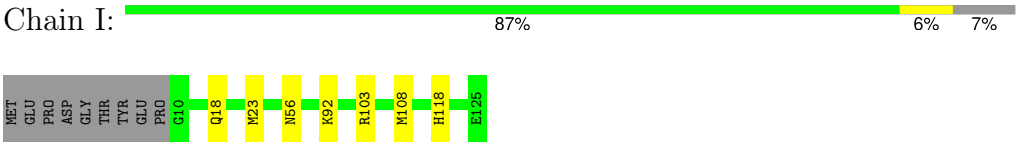
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



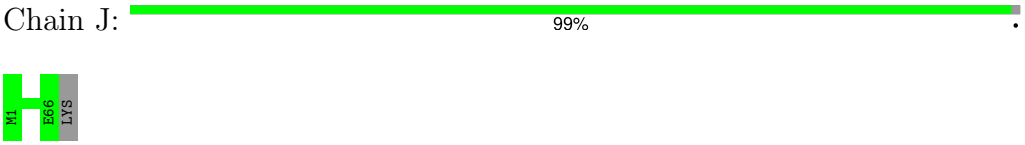
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



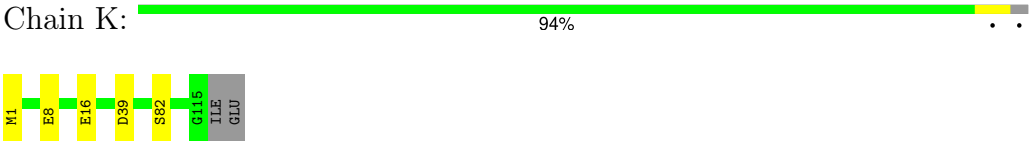
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



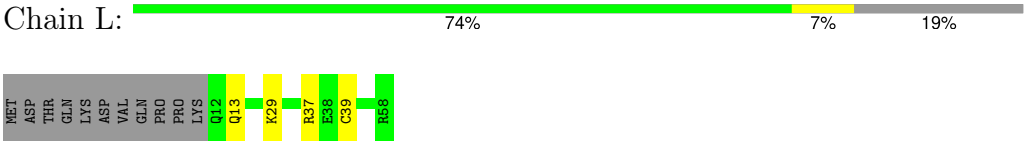
• Molecule 11: RNA polymerase II subunit J

Chain K:



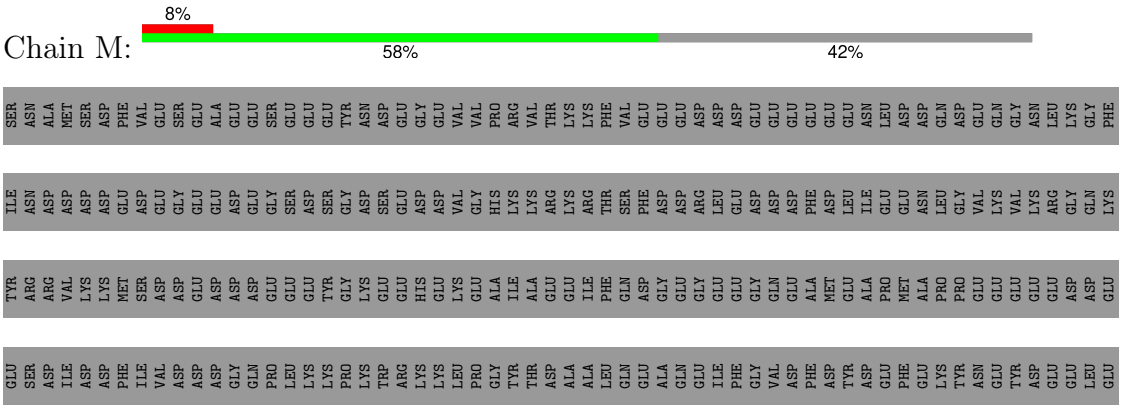
• Molecule 12: RNA polymerase II subunit K

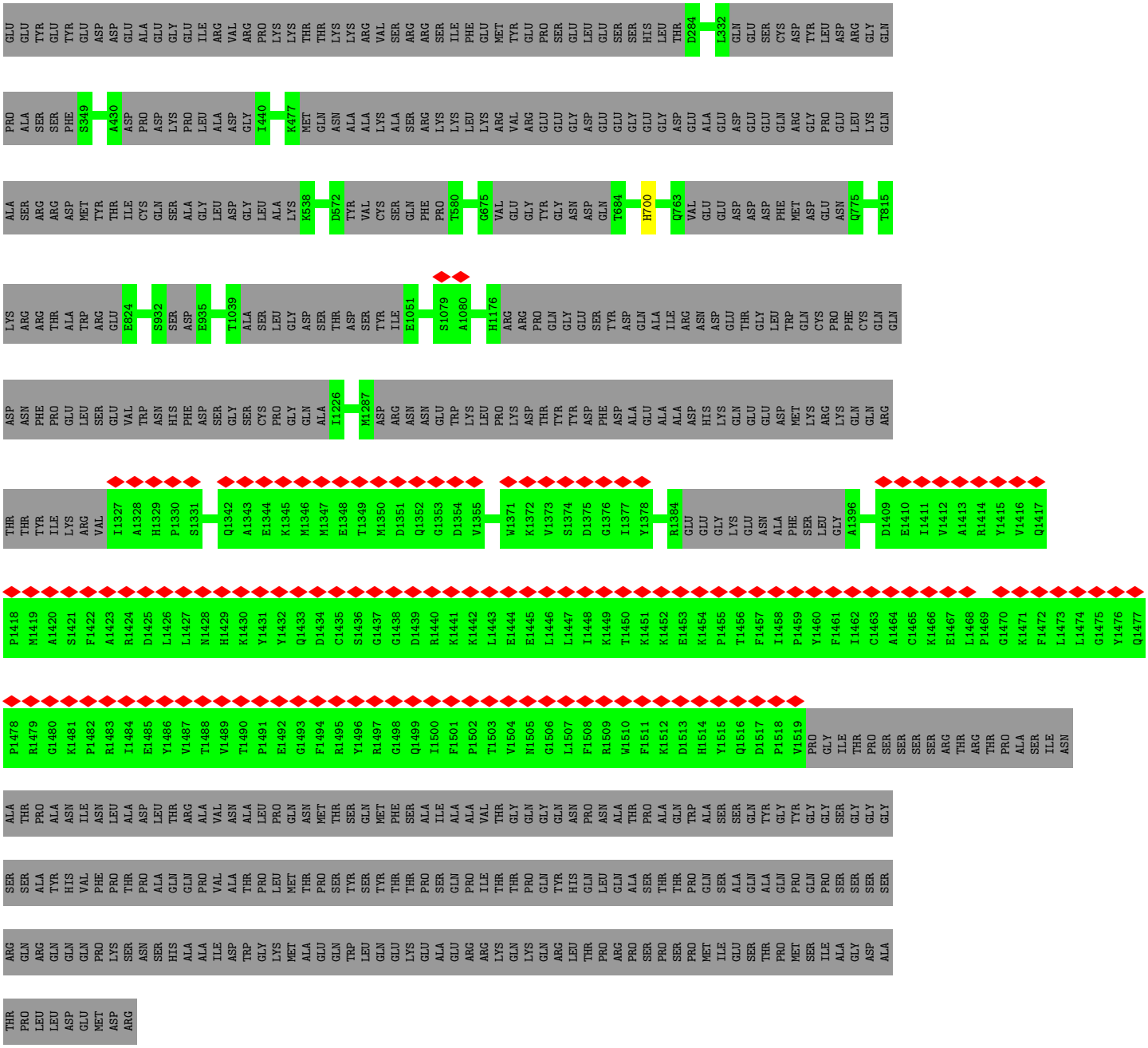
Chain L:



• Molecule 13: Transcription elongation factor SPT6

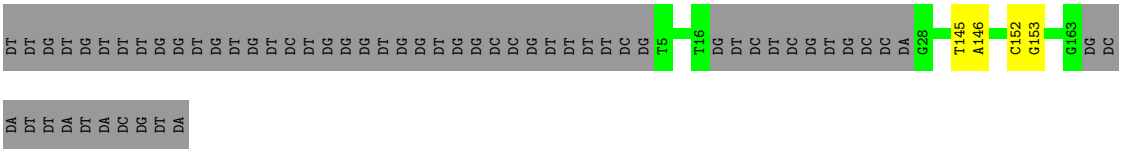
Chain M:





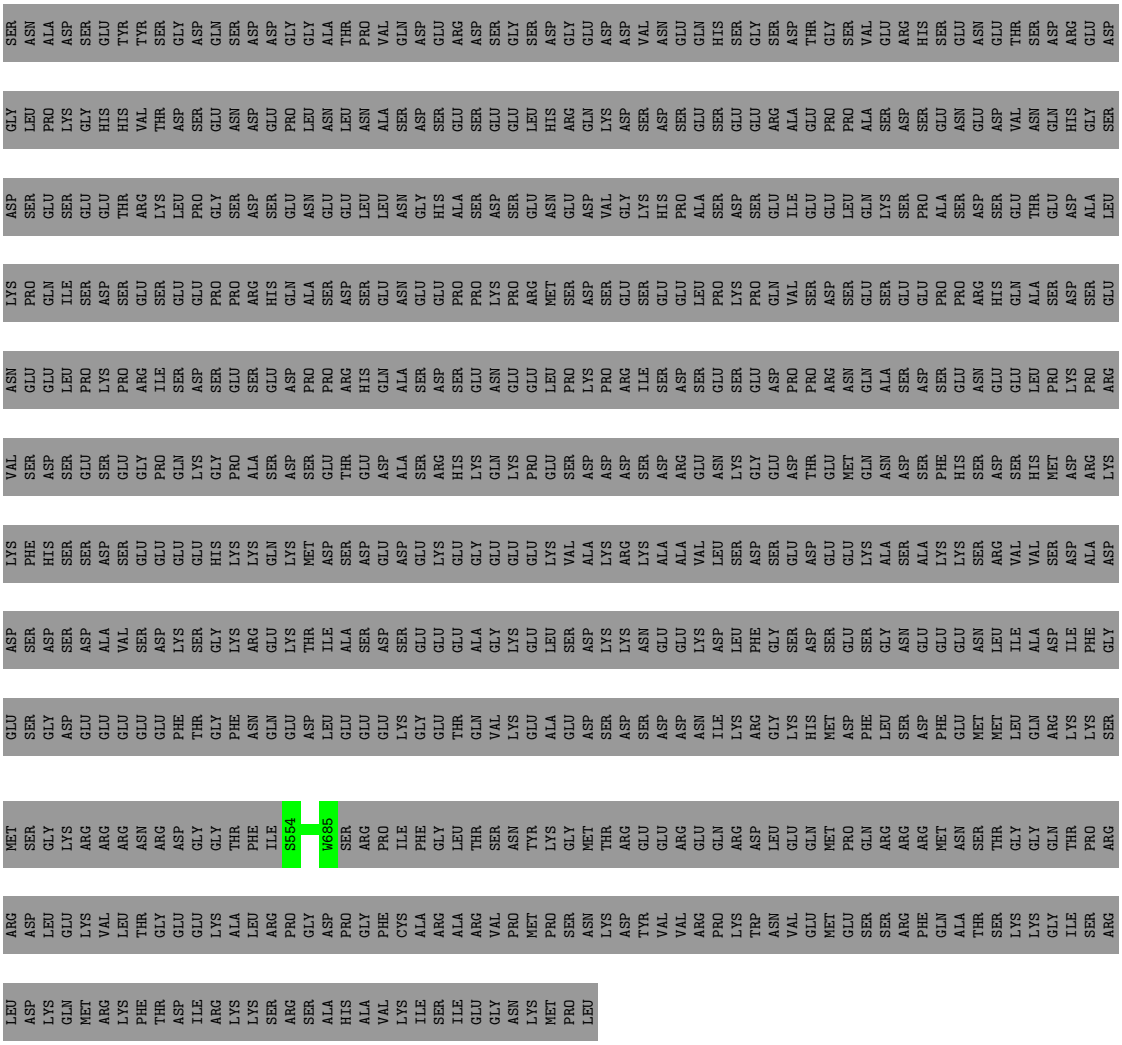
• Molecule 14: Non-template DNA

Chain N:



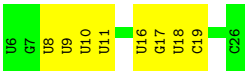
• Molecule 15: Protein IWS1 homolog

Chain O:



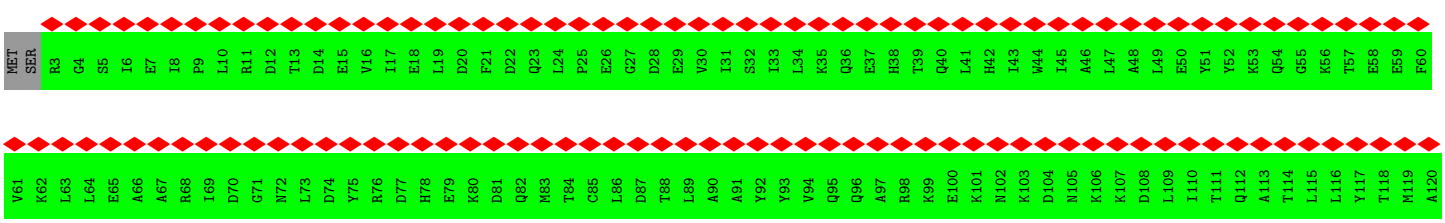
• Molecule 16: RNA

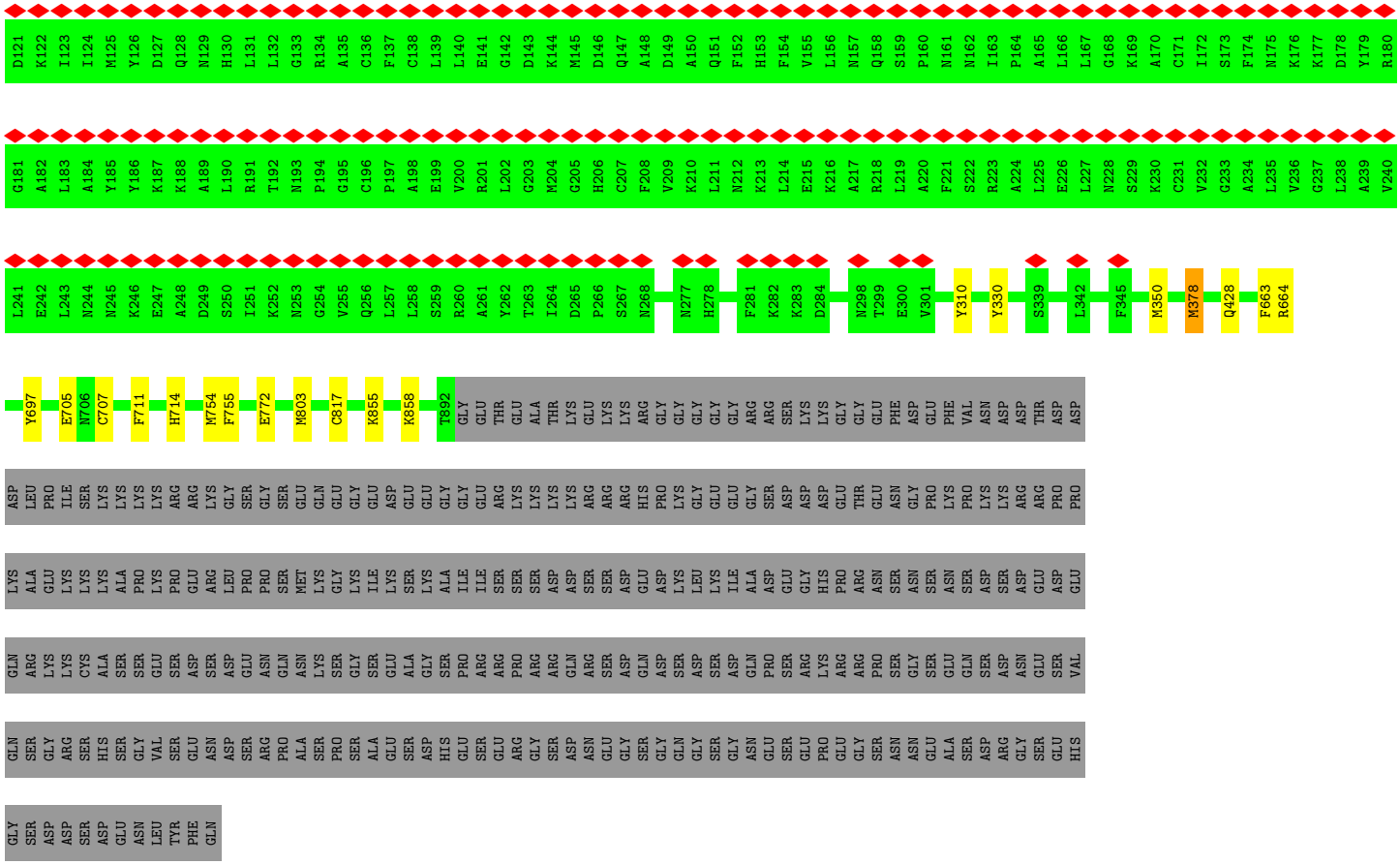
Chain P:



• Molecule 17: RNA polymerase-associated protein CTR9 homolog

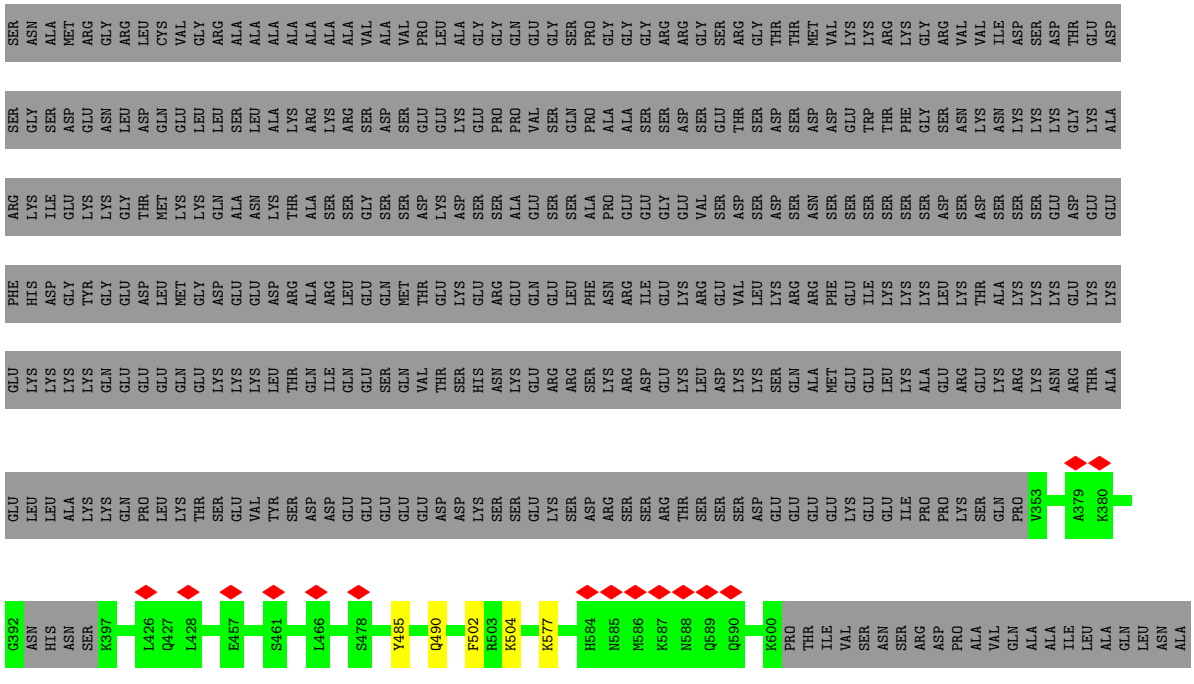
Chain Q:

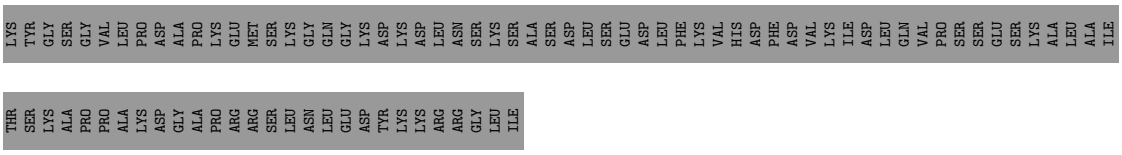




• Molecule 18: RNA polymerase-associated protein RTF1 homolog

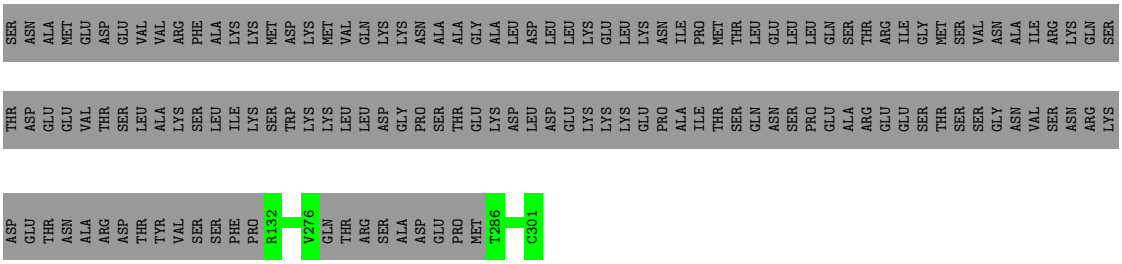
Chain R:





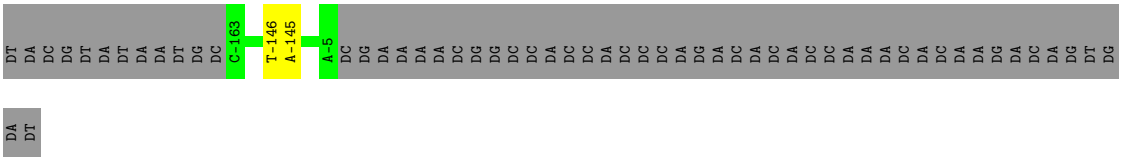
• Molecule 19: Transcription elongation factor A protein 1

Chain S:



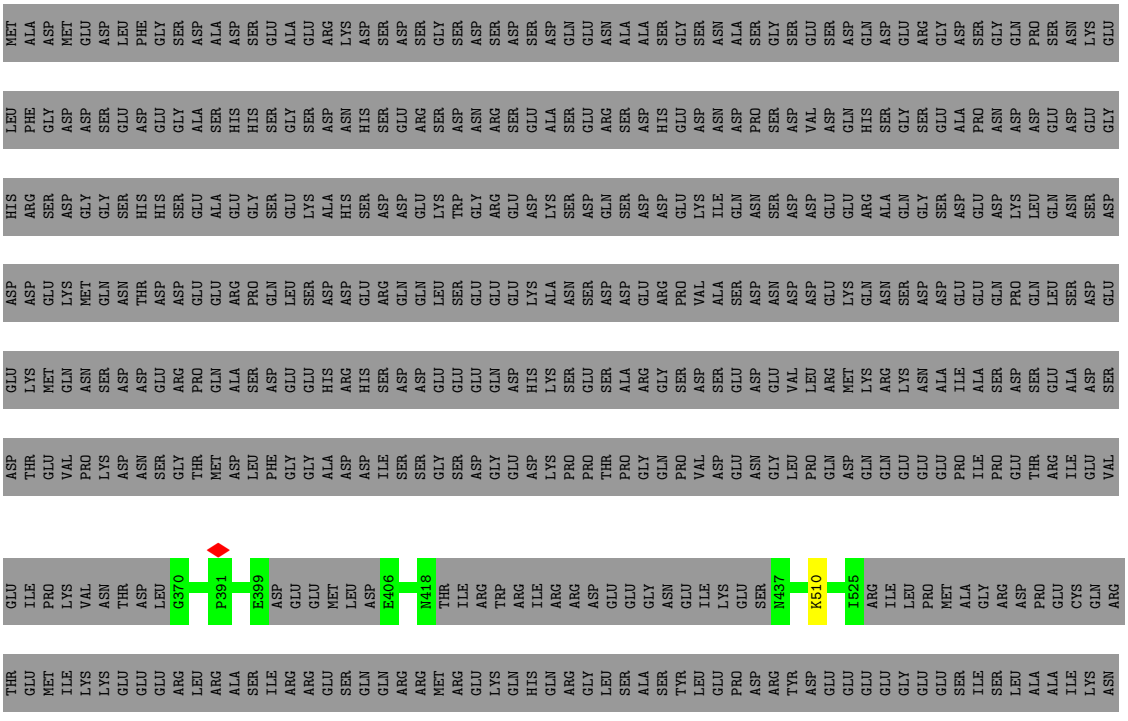
• Molecule 20: Template DNA

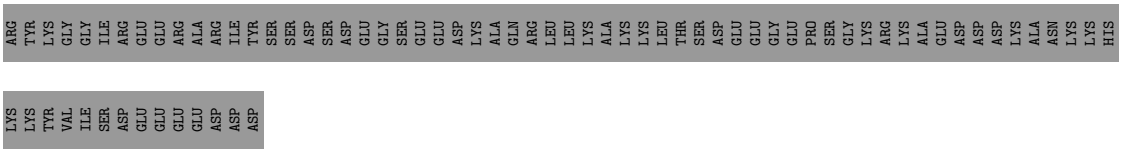
Chain T:



• Molecule 21: RNA polymerase-associated protein LEO1

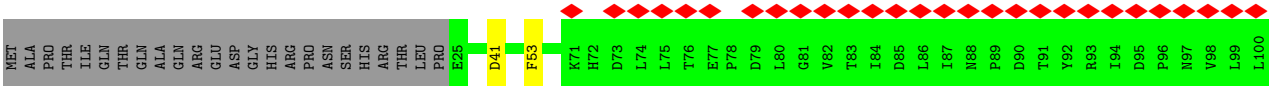
Chain U:

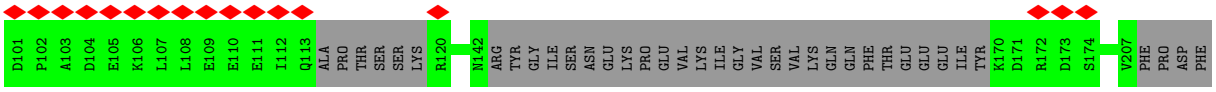


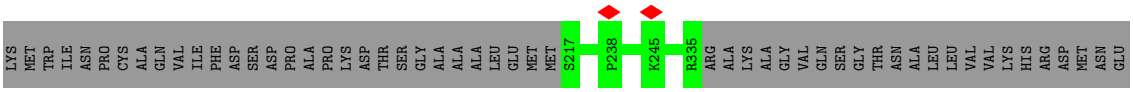


• Molecule 22: RNA polymerase II-associated factor 1 homolog

Chain V:

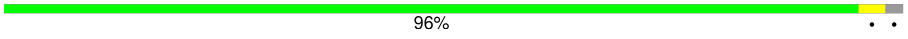






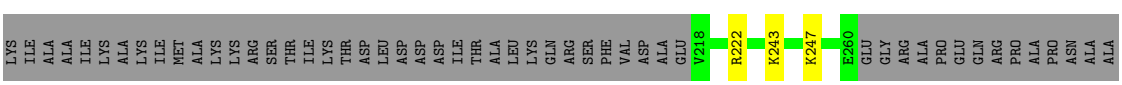
• Molecule 23: WDR61

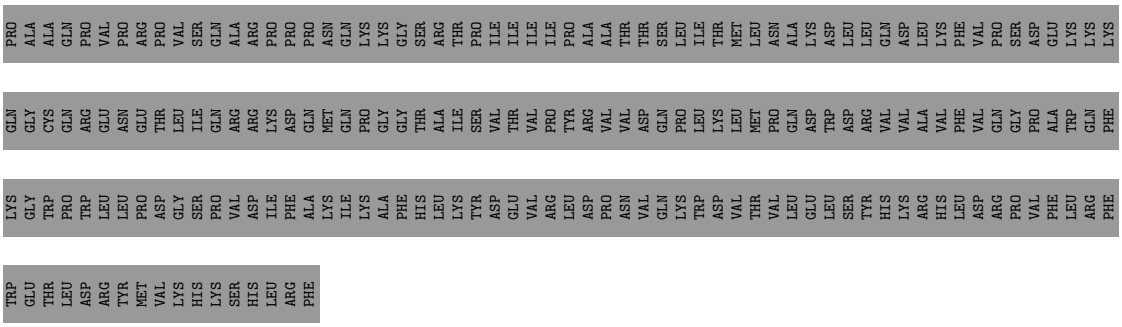
Chain W:



• Molecule 24: Parafibromin

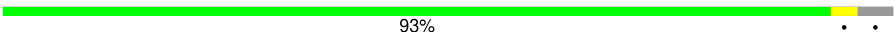
Chain X:

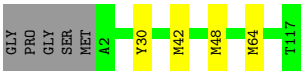




• Molecule 25: Transcription elongation factor SPT4

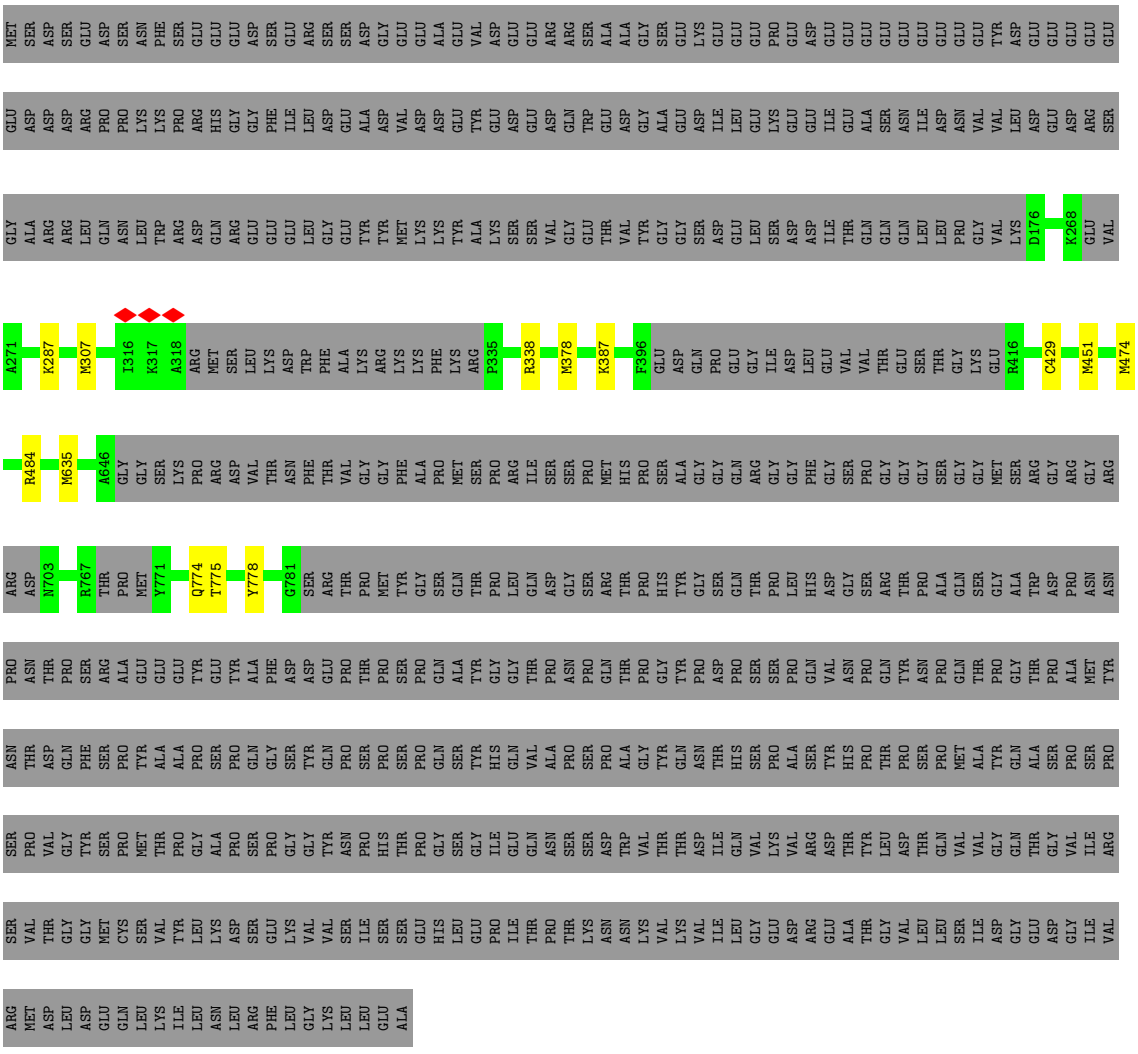
Chain Y:





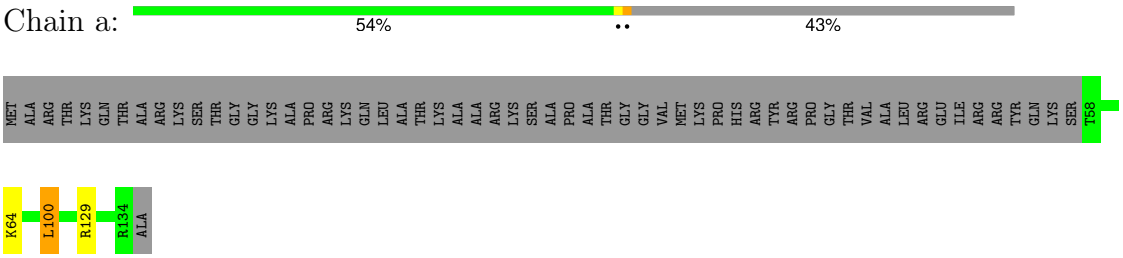
• Molecule 26: Transcription elongation factor SPT5

Chain Z:



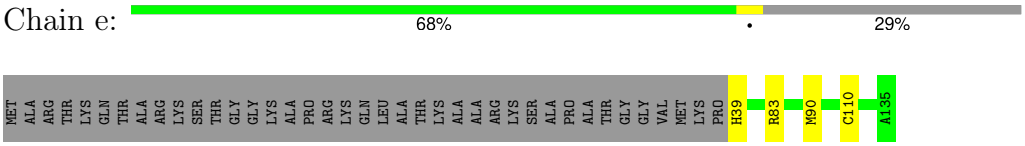
• Molecule 27: Histone H3

Chain a:



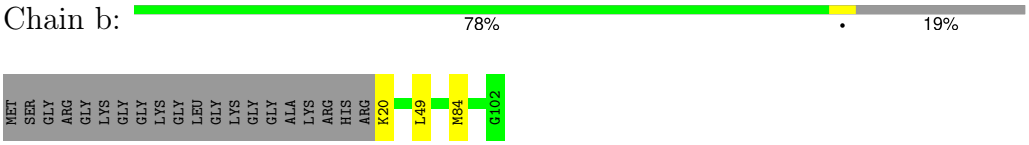
• Molecule 27: Histone H3

Chain e:



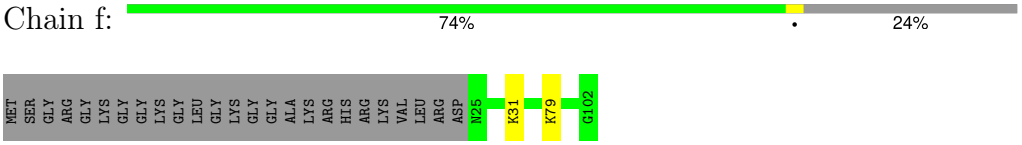
• Molecule 28: Histone H4

Chain b:



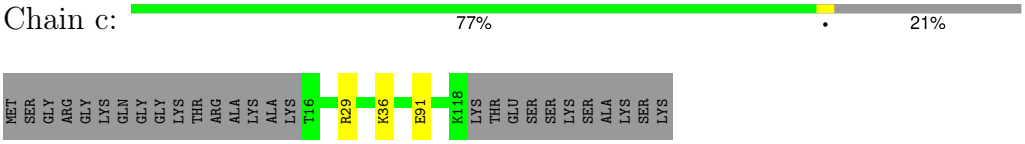
• Molecule 28: Histone H4

Chain f:



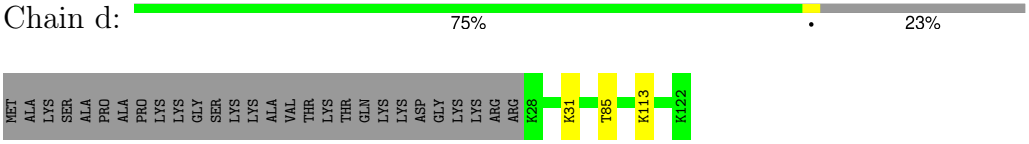
• Molecule 29: Histone H2A type 1

Chain c:



• Molecule 30: Histone H2B 1.1

Chain d:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	1139653	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.708	Depositor
Minimum map value	-0.178	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.014	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	549.46, 549.46, 549.46	wwPDB
Map dimensions	500, 500, 500	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.09892, 1.09892, 1.09892	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, TPO, SEP, MG

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	A	0.27	0/11384	0.57	6/15361 (0.0%)
2	B	0.26	0/9158	0.54	1/12360 (0.0%)
3	C	0.26	0/2115	0.54	0/2873
4	D	0.25	0/1017	0.51	0/1368
5	E	0.26	0/1751	0.54	0/2366
6	F	0.29	0/636	0.59	0/859
7	G	0.28	0/1364	0.55	0/1853
8	H	0.30	0/1219	0.61	1/1644 (0.1%)
9	I	0.31	0/964	0.58	0/1305
10	J	0.30	0/533	0.54	0/719
11	K	0.29	0/939	0.51	0/1271
12	L	0.32	0/403	0.70	0/536
13	M	0.22	0/4330	0.42	0/5591
14	N	0.54	0/3417	0.93	4/5275 (0.1%)
15	O	0.23	0/655	0.36	0/913
16	P	0.17	0/477	0.73	0/738
17	Q	0.27	0/6531	0.52	2/8861 (0.0%)
18	R	0.25	0/1437	0.48	0/1972
19	S	0.22	0/659	0.41	0/827
20	T	0.52	0/3642	0.86	2/5614 (0.0%)
21	U	0.24	0/613	0.46	0/847
22	V	0.25	0/1386	0.48	0/1909
23	W	0.32	1/2392 (0.0%)	0.58	3/3257 (0.1%)
24	X	0.30	0/356	0.65	0/478
25	Y	0.25	0/927	0.57	0/1250
26	Z	0.25	0/4084	0.52	0/5498
27	a	0.29	0/634	0.67	1/851 (0.1%)
27	e	0.29	0/812	0.62	0/1088
28	b	0.30	0/669	0.64	1/894 (0.1%)
28	f	0.29	0/626	0.66	0/837
29	c	0.28	0/805	0.58	0/1088
30	d	0.31	0/756	0.60	0/1015

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
All	All	0.31	1/66691 (0.0%)	0.60	21/91318 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
23	W	106	PRO	CG-CD	-6.58	1.28	1.50

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
23	W	106	PRO	CA-N-CD	-13.85	92.11	111.50
1	A	911	PRO	CA-N-CD	-10.19	97.24	111.50
1	A	433	PRO	CA-N-CD	-8.17	100.06	111.50
23	W	106	PRO	N-CD-CG	-7.77	91.55	103.20
14	N	146	DA	OP1-P-OP2	-7.07	109.00	119.60
14	N	153	DG	OP1-P-OP2	-6.78	109.42	119.60
20	T	-145	DA	OP1-P-OP2	-6.65	109.62	119.60
8	H	17	PRO	CA-N-CD	-6.48	102.43	111.50
14	N	145	DT	OP1-P-O3'	6.34	119.14	105.20
1	A	612	ASP	CB-CG-OD1	6.23	123.90	118.30
23	W	52	GLU	CA-CB-CG	6.20	127.03	113.40
28	b	49	LEU	CA-CB-CG	5.93	128.95	115.30
1	A	910	LYS	C-N-CD	-5.75	107.94	120.60
14	N	152	DC	OP1-P-O3'	5.52	117.35	105.20
17	Q	772	GLU	C-N-CA	5.46	135.34	121.70
20	T	-146	DT	OP2-P-O3'	5.45	117.19	105.20
27	a	100	LEU	CA-CB-CG	5.40	127.72	115.30
1	A	1095	LEU	CA-CB-CG	5.37	127.66	115.30
2	B	473	LEU	CA-CB-CG	5.07	126.97	115.30
17	Q	378	MET	CA-CB-CG	5.01	121.81	113.30
1	A	1026	ASP	CB-CG-OD1	5.00	122.80	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	A	1408/1984 (71%)	1330 (94%)	76 (5%)	2 (0%)	48 77
2	B	1112/1251 (89%)	1042 (94%)	70 (6%)	0	100 100
3	C	254/275 (92%)	236 (93%)	18 (7%)	0	100 100
4	D	124/142 (87%)	120 (97%)	4 (3%)	0	100 100
5	E	207/210 (99%)	200 (97%)	7 (3%)	0	100 100
6	F	76/127 (60%)	74 (97%)	2 (3%)	0	100 100
7	G	169/172 (98%)	162 (96%)	7 (4%)	0	100 100
8	H	147/150 (98%)	137 (93%)	10 (7%)	0	100 100
9	I	114/125 (91%)	104 (91%)	10 (9%)	0	100 100
10	J	64/67 (96%)	58 (91%)	6 (9%)	0	100 100
11	K	113/117 (97%)	111 (98%)	2 (2%)	0	100 100
12	L	45/58 (78%)	39 (87%)	5 (11%)	1 (2%)	5 21
13	M	976/1729 (56%)	935 (96%)	40 (4%)	1 (0%)	48 77
15	O	130/821 (16%)	127 (98%)	3 (2%)	0	100 100
17	Q	888/1179 (75%)	863 (97%)	25 (3%)	0	100 100
18	R	240/713 (34%)	230 (96%)	10 (4%)	0	100 100
19	S	157/304 (52%)	156 (99%)	1 (1%)	0	100 100
21	U	117/666 (18%)	104 (89%)	12 (10%)	1 (1%)	14 43
22	V	234/531 (44%)	220 (94%)	14 (6%)	0	100 100
23	W	298/305 (98%)	282 (95%)	16 (5%)	0	100 100
24	X	41/531 (8%)	40 (98%)	1 (2%)	0	100 100
25	Y	114/121 (94%)	106 (93%)	8 (7%)	0	100 100
26	Z	497/1087 (46%)	477 (96%)	19 (4%)	1 (0%)	44 73
27	a	75/136 (55%)	75 (100%)	0	0	100 100
27	e	95/136 (70%)	95 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
28	b	81/103 (79%)	79 (98%)	2 (2%)	0	100	100
28	f	76/103 (74%)	74 (97%)	2 (3%)	0	100	100
29	c	101/130 (78%)	100 (99%)	1 (1%)	0	100	100
30	d	93/123 (76%)	90 (97%)	3 (3%)	0	100	100
All	All	8046/13396 (60%)	7666 (95%)	374 (5%)	6 (0%)	50	77

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1343	LEU
12	L	39	CYS
13	M	700	HIS
21	U	510	LYS
26	Z	774	GLN
1	A	433	PRO

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	A	1229/1761 (70%)	1212 (99%)	17 (1%)	62	86
2	B	986/1084 (91%)	968 (98%)	18 (2%)	54	82
3	C	235/252 (93%)	231 (98%)	4 (2%)	56	83
4	D	109/126 (86%)	106 (97%)	3 (3%)	38	73
5	E	191/192 (100%)	191 (100%)	0	100	100
6	F	68/111 (61%)	64 (94%)	4 (6%)	16	45
7	G	146/153 (95%)	142 (97%)	4 (3%)	40	73
8	H	130/131 (99%)	123 (95%)	7 (5%)	18	49
9	I	104/112 (93%)	97 (93%)	7 (7%)	13	39
10	J	55/56 (98%)	55 (100%)	0	100	100
11	K	104/106 (98%)	99 (95%)	5 (5%)	21	54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
12	L	44/55 (80%)	41 (93%)	3 (7%)	13 38
13	M	41/1524 (3%)	41 (100%)	0	100 100
17	Q	533/1011 (53%)	515 (97%)	18 (3%)	32 67
18	R	57/625 (9%)	52 (91%)	5 (9%)	8 26
19	S	4/268 (2%)	4 (100%)	0	100 100
22	V	46/462 (10%)	44 (96%)	2 (4%)	25 57
23	W	255/260 (98%)	249 (98%)	6 (2%)	44 76
24	X	40/467 (9%)	37 (92%)	3 (8%)	11 33
25	Y	102/105 (97%)	98 (96%)	4 (4%)	27 62
26	Z	435/939 (46%)	424 (98%)	11 (2%)	42 75
27	a	67/111 (60%)	64 (96%)	3 (4%)	23 56
27	e	84/111 (76%)	80 (95%)	4 (5%)	21 54
28	b	68/79 (86%)	66 (97%)	2 (3%)	37 72
28	f	63/79 (80%)	61 (97%)	2 (3%)	34 69
29	c	82/102 (80%)	79 (96%)	3 (4%)	29 64
30	d	81/103 (79%)	78 (96%)	3 (4%)	29 64
All	All	5359/10385 (52%)	5221 (97%)	138 (3%)	42 74

All (138) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	51	ARG
1	A	295	GLN
1	A	317	MET
1	A	329	MET
1	A	524	MET
1	A	539	GLN
1	A	602	CYS
1	A	612	ASP
1	A	757	GLN
1	A	914	LYS
1	A	931	ARG
1	A	1192	TRP
1	A	1210	TRP
1	A	1396	ARG
1	A	1405	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1432	PHE
1	A	1484	MET
2	B	84	TYR
2	B	170	ASP
2	B	297	MET
2	B	309	PHE
2	B	347	MET
2	B	368	MET
2	B	388	TYR
2	B	438	ARG
2	B	594	MET
2	B	610	ARG
2	B	648	TYR
2	B	768	ARG
2	B	848	LEU
2	B	859	ARG
2	B	1048	TYR
2	B	1075	MET
2	B	1163	MET
2	B	1165	MET
3	C	80	ASP
3	C	83	GLN
3	C	211	LEU
3	C	247	SER
4	D	24	LYS
4	D	39	MET
4	D	74	PHE
6	F	50	LYS
6	F	98	LYS
6	F	117	ASP
6	F	123	LEU
7	G	1	MET
7	G	8	GLU
7	G	104	MET
7	G	131	MET
8	H	8	ASP
8	H	20	LYS
8	H	51	ASP
8	H	75	TYR
8	H	140	ARG
8	H	145	MET
8	H	147	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
9	I	18	GLN
9	I	23	MET
9	I	56	ASN
9	I	92	LYS
9	I	103	ARG
9	I	108	MET
9	I	118	HIS
11	K	1	MET
11	K	8	GLU
11	K	16	GLU
11	K	39	ASP
11	K	82	SER
12	L	13	GLN
12	L	29	LYS
12	L	37	ARG
17	Q	310	TYR
17	Q	330	TYR
17	Q	350	MET
17	Q	378	MET
17	Q	428	GLN
17	Q	663	PHE
17	Q	664	ARG
17	Q	697	TYR
17	Q	705	GLU
17	Q	707	CYS
17	Q	711	PHE
17	Q	714	HIS
17	Q	754	MET
17	Q	755	PHE
17	Q	803	MET
17	Q	817	CYS
17	Q	855	LYS
17	Q	858	LYS
18	R	485	TYR
18	R	490	GLN
18	R	502	PHE
18	R	504	LYS
18	R	577	LYS
22	V	41	ASP
22	V	53	PHE
23	W	16	ASP
23	W	145	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
23	W	147	ARG
23	W	191	MET
23	W	194	ARG
23	W	271	PHE
24	X	222	ARG
24	X	243	LYS
24	X	247	LYS
25	Y	30	TYR
25	Y	42	MET
25	Y	48	MET
25	Y	64	MET
26	Z	287	LYS
26	Z	307	MET
26	Z	338	ARG
26	Z	378	MET
26	Z	387	LYS
26	Z	429	CYS
26	Z	451	MET
26	Z	474	MET
26	Z	484	ARG
26	Z	635	MET
26	Z	778	TYR
27	a	64	LYS
27	a	100	LEU
27	a	129	ARG
28	b	20	LYS
28	b	84	MET
29	c	29	ARG
29	c	36	LYS
29	c	91	GLU
30	d	31	LYS
30	d	85	THR
30	d	113	LYS
27	e	39	HIS
27	e	83	ARG
27	e	90	MET
27	e	110	CYS
28	f	31	LYS
28	f	79	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (53) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	96	HIS
1	A	122	ASN
1	A	188	GLN
1	A	320	ASN
1	A	372	ASN
1	A	387	ASN
1	A	539	GLN
1	A	700	GLN
1	A	780	ASN
1	A	790	GLN
1	A	809	HIS
1	A	991	GLN
1	A	1036	ASN
1	A	1044	HIS
1	A	1129	ASN
1	A	1194	ASN
1	A	1417	HIS
1	A	1457	ASN
1	A	1462	GLN
2	B	197	GLN
2	B	227	ASN
2	B	319	ASN
2	B	420	GLN
2	B	471	ASN
2	B	582	GLN
2	B	649	ASN
2	B	725	GLN
2	B	1094	GLN
2	B	1142	ASN
3	C	111	GLN
8	H	29	HIS
9	I	56	ASN
9	I	91	HIS
12	L	26	ASN
17	Q	305	GLN
17	Q	311	GLN
17	Q	373	ASN
17	Q	527	HIS
17	Q	616	HIS
17	Q	860	GLN
22	V	56	ASN
22	V	69	GLN
23	W	27	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
23	W	173	ASN
23	W	223	ASN
23	W	273	HIS
25	Y	12	HIS
26	Z	244	ASN
26	Z	272	ASN
28	b	93	GLN
30	d	64	ASN
27	e	68	GLN
27	e	108	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
16	P	20/21 (95%)	7 (35%)	3 (15%)

All (7) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
16	P	8	U
16	P	9	U
16	P	10	U
16	P	11	U
16	P	16	U
16	P	17	G
16	P	19	C

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
16	P	8	U
16	P	16	U
16	P	18	U

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

3 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	SEP	A	1547	1	8,9,10	1.62	1 (12%)	7,12,14	1.29	1 (14%)
1	TPO	A	1525	1	8,10,11	1.11	0	10,14,16	2.14	1 (10%)
26	TPO	Z	775	26	8,10,11	1.12	0	10,14,16	1.97	1 (10%)

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
1	SEP	A	1547	1	-	0/6/8/10	-
1	TPO	A	1525	1	-	0/9/11/13	-
26	TPO	Z	775	26	-	2/9/11/13	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1547	SEP	P-O1P	3.53	1.61	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1525	TPO	P-OG1-CB	-6.14	106.65	123.33
26	Z	775	TPO	P-OG1-CB	-5.70	107.84	123.33
1	A	1547	SEP	OG-CB-CA	2.81	110.88	108.14

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
26	Z	775	TPO	C-CA-CB-CG2
26	Z	775	TPO	CB-OG1-P-O2P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
21	U	1
22	V	1
13	M	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	U	497:ASP	C	505:SER	N	27.06
1	V	299:GLU	C	310:ASN	N	12.67
1	M	1334:ASN	C	1338:ILE	N	5.50

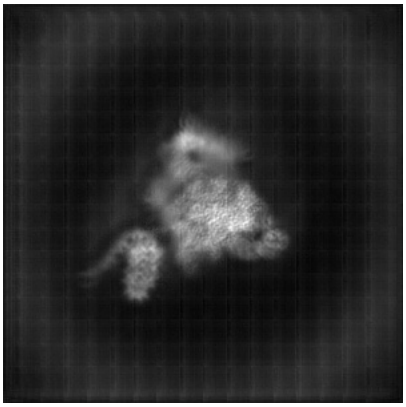
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-48039. These allow visual inspection of the internal detail of the map and identification of artifacts.

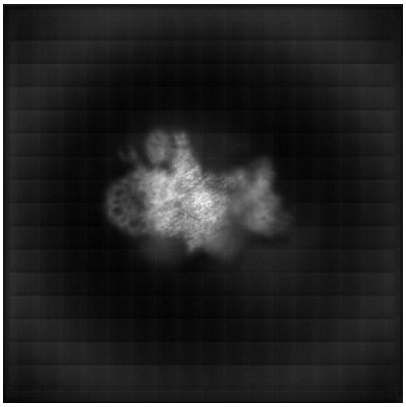
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

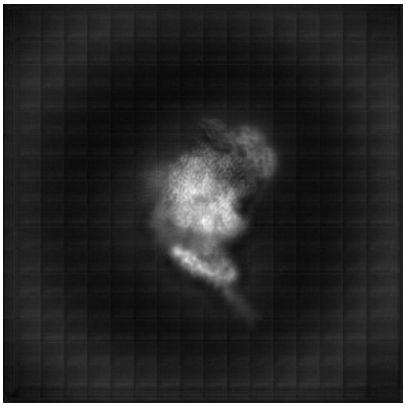
6.1.1 Primary map



X



Y

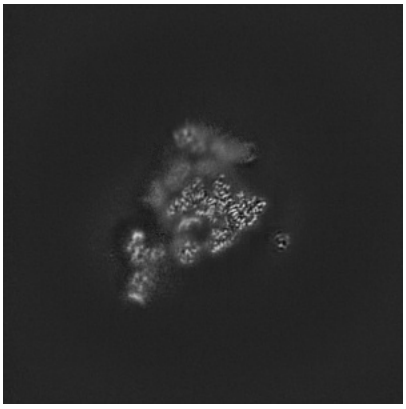


Z

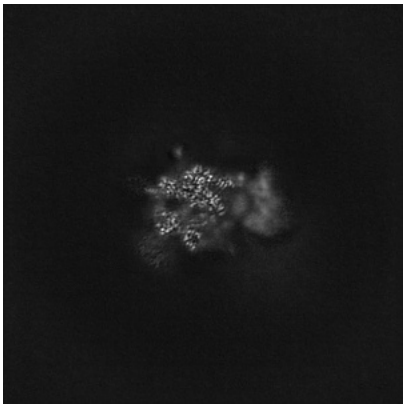
The images above show the map projected in three orthogonal directions.

6.2 Central slices

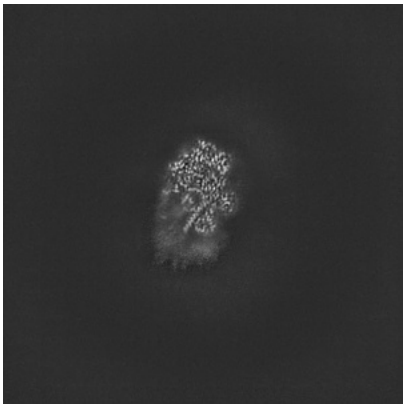
6.2.1 Primary map



X Index: 250



Y Index: 250

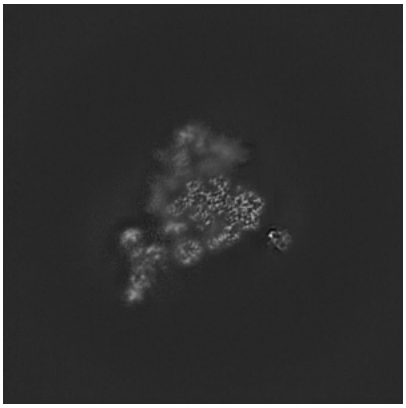


Z Index: 250

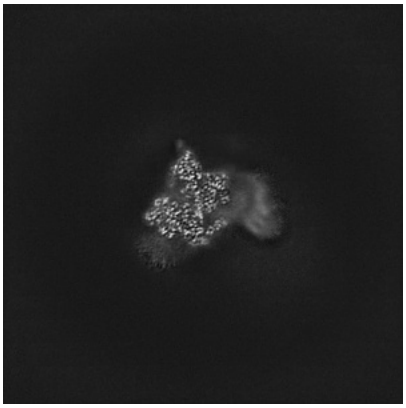
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

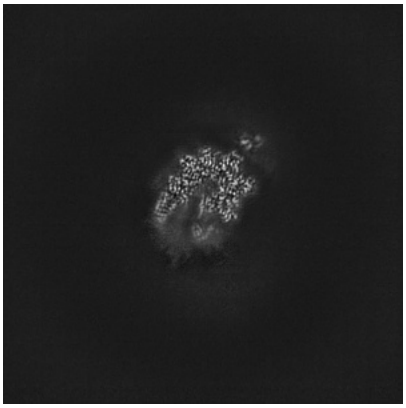
6.3.1 Primary map



X Index: 257



Y Index: 264

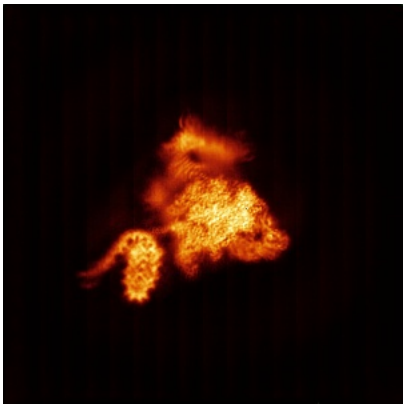


Z Index: 236

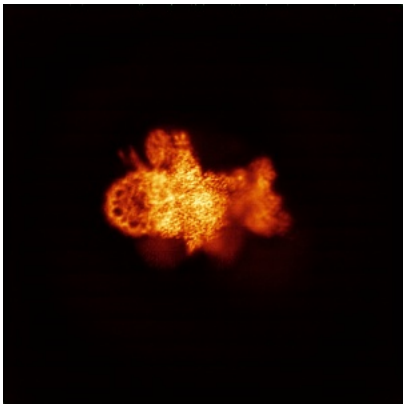
The images above show the largest variance slices of the map in three orthogonal directions.

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

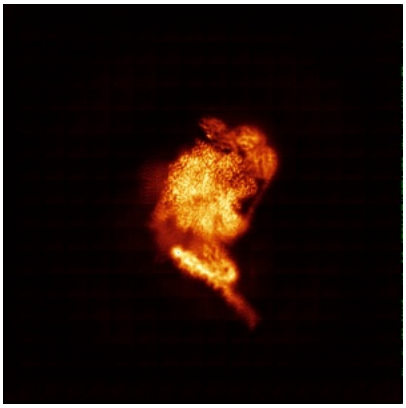
6.4.1 Primary map



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

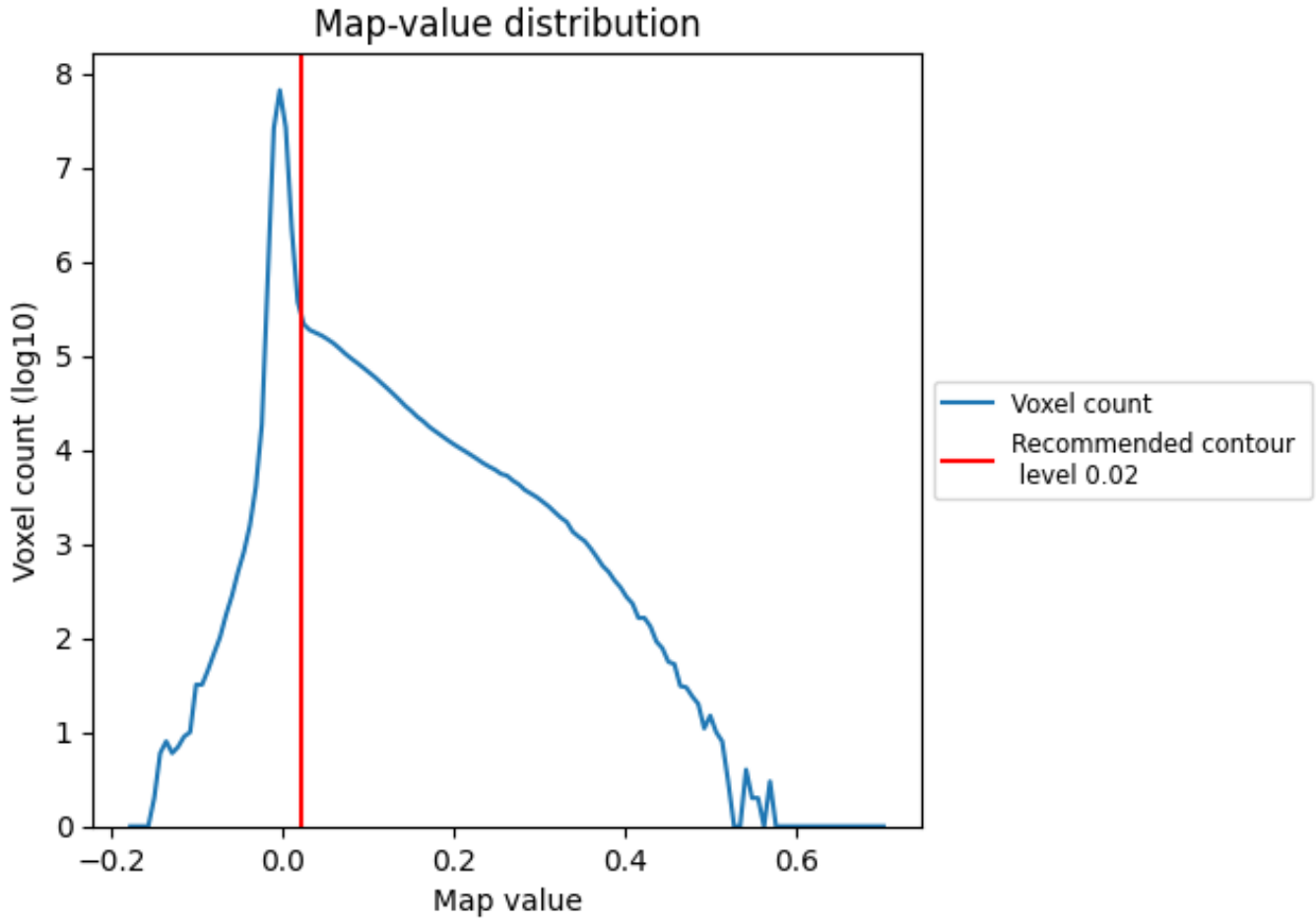
6.6 Mask visualisation (i)

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

7 Map analysis

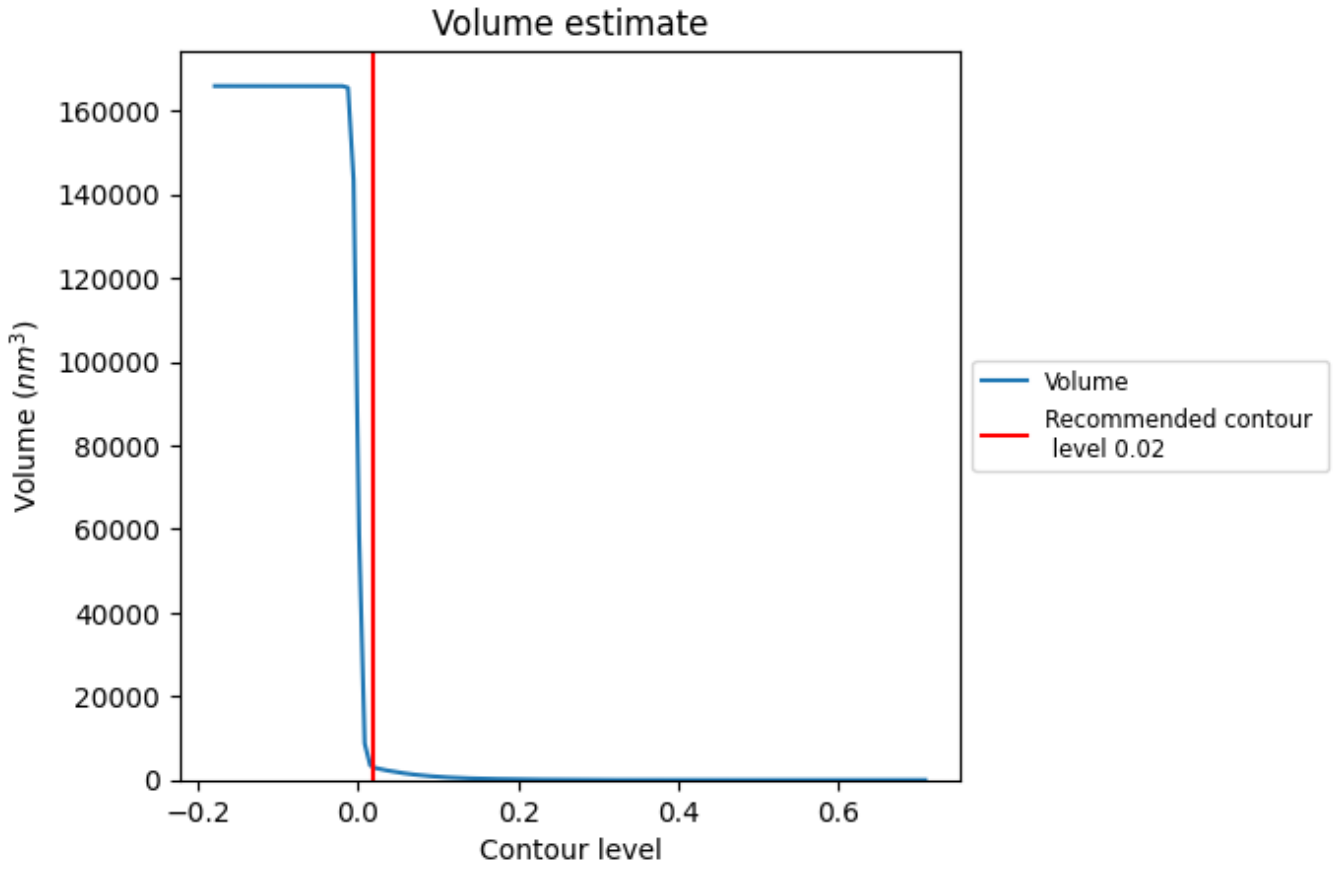
This section contains the results of statistical analysis of the map.

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.

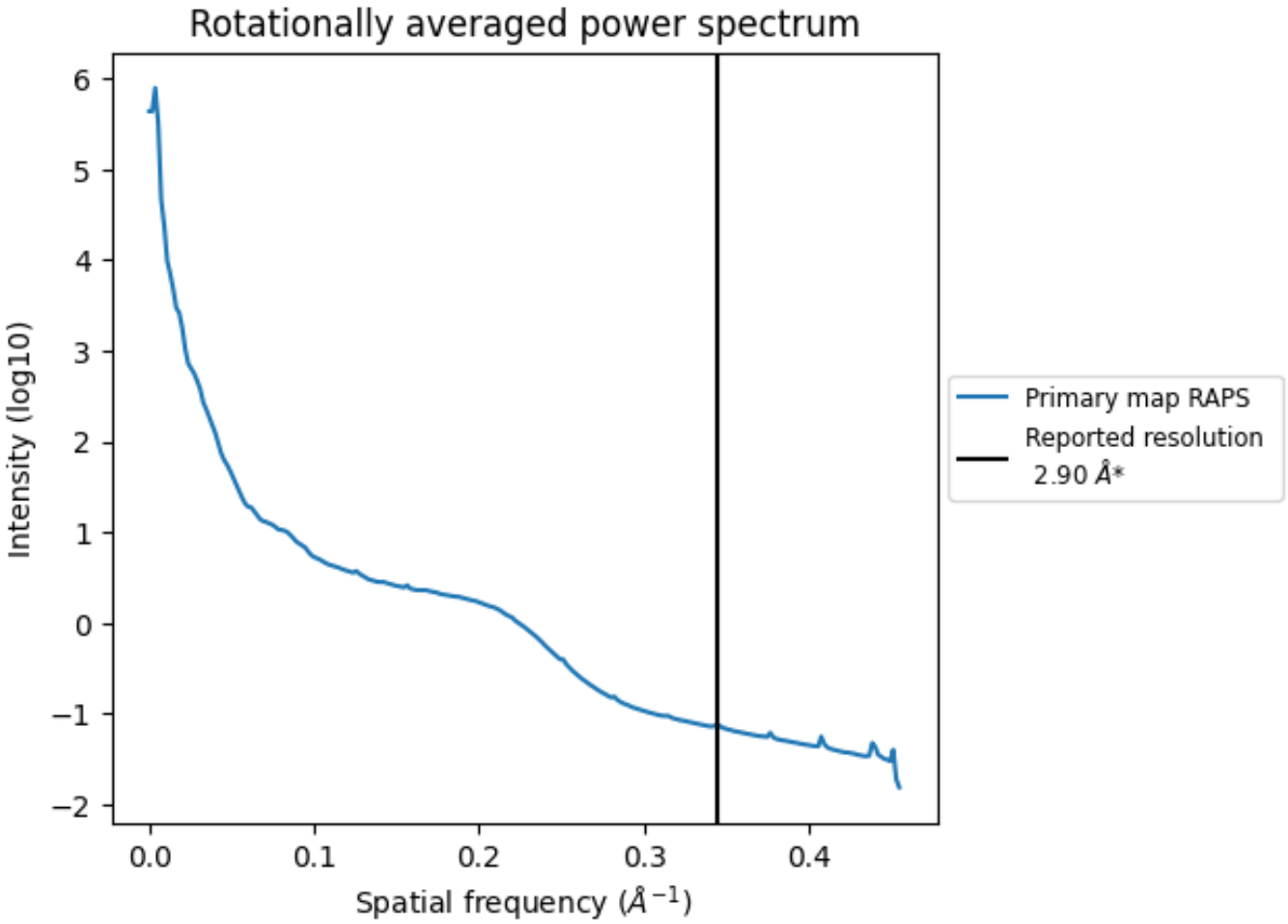
7.2 Volume estimate



The volume at the recommended contour level is 3060 nm^3; this corresponds to an approximate mass of 2764 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 0.345 Å^{-1}

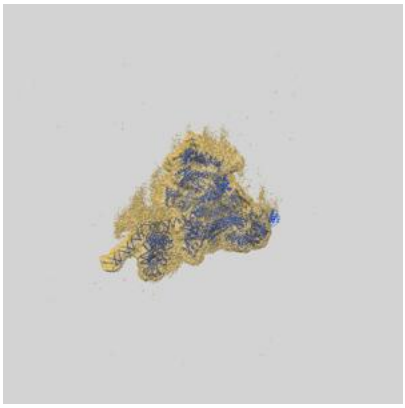
8 Fourier-Shell correlation (i)

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

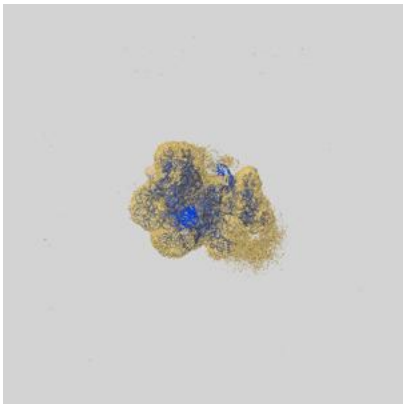
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-48039 and PDB model 9EGX. Per-residue inclusion information can be found in section 3 on page 11.

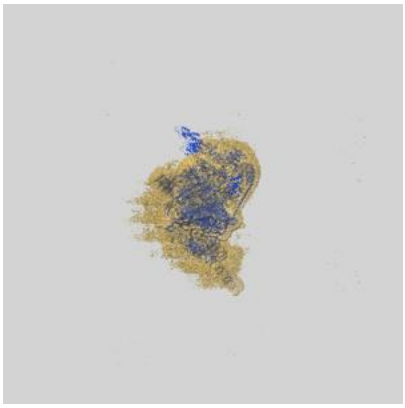
9.1 Map-model overlay



X



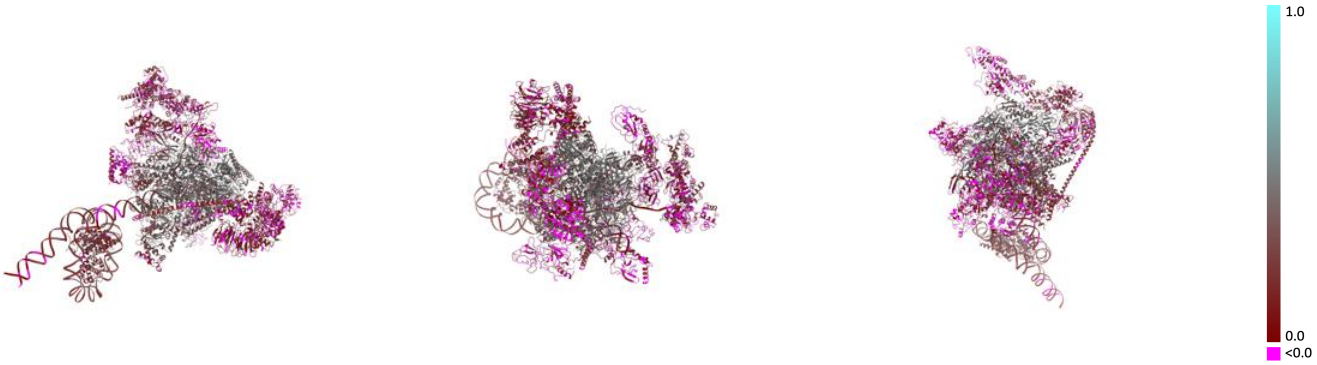
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.02 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.

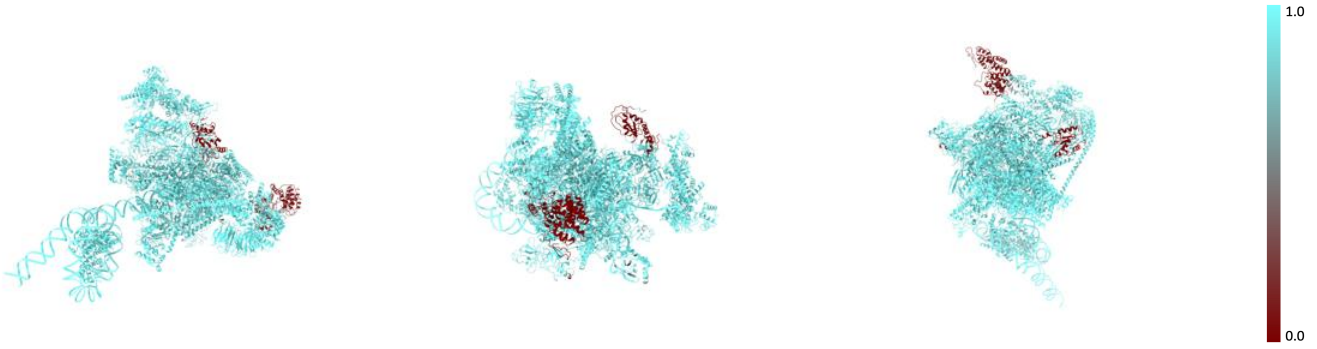
9.2 Q-score mapped to coordinate model (i)



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.

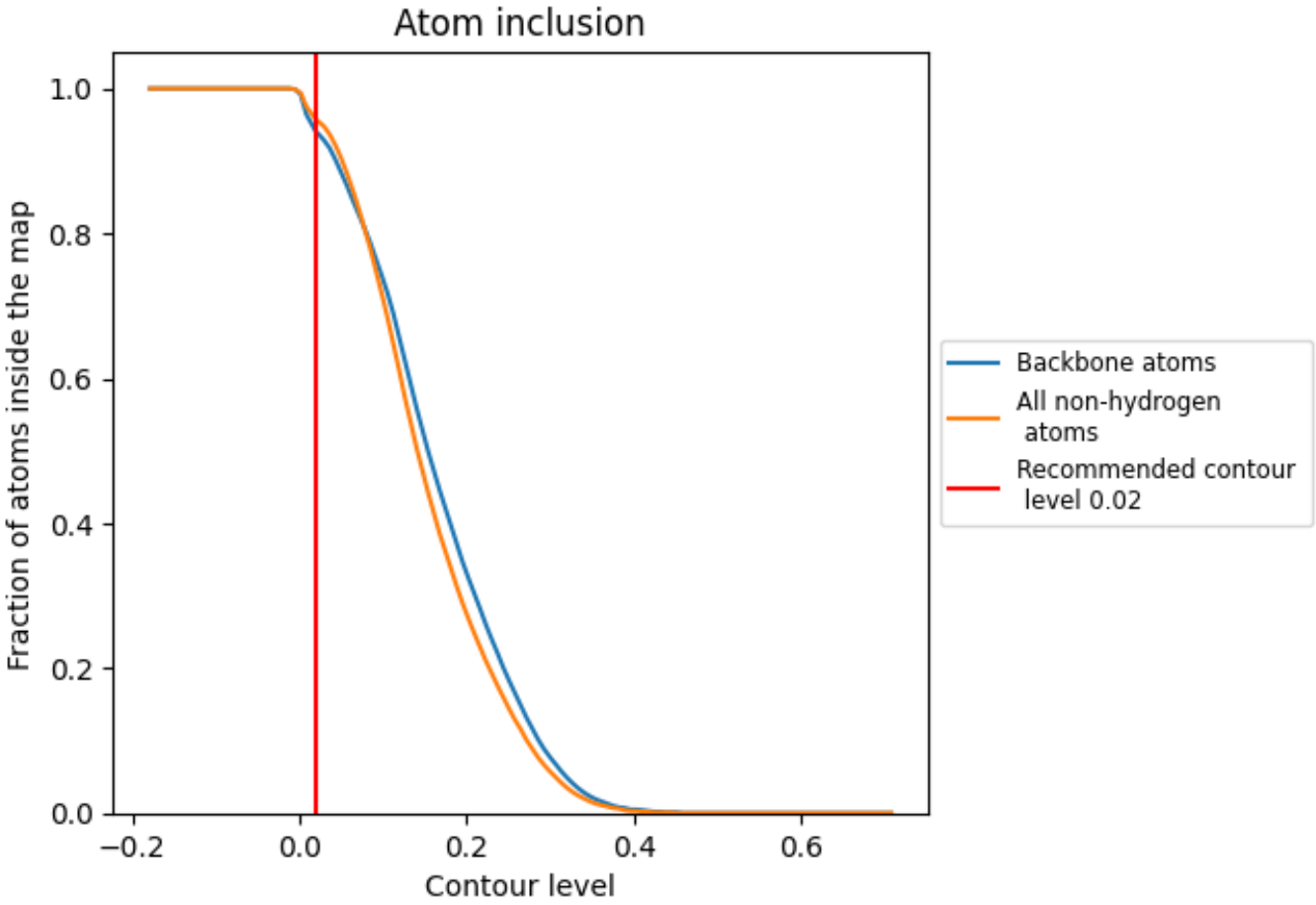
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.02).

9.4 Atom inclusion (i)



At the recommended contour level, 94% of all backbone atoms, 96% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9570	0.2500
A	0.9920	0.3920
B	1.0000	0.4050
C	1.0000	0.4320
D	1.0000	0.1870
E	1.0000	0.3770
F	1.0000	0.4380
G	1.0000	0.2100
H	0.9970	0.4250
I	1.0000	0.3720
J	1.0000	0.4090
K	1.0000	0.4370
L	1.0000	0.3950
M	0.8420	0.0920
N	0.9940	0.1820
O	1.0000	-0.0080
P	1.0000	0.1930
Q	0.7640	0.0960
R	0.9150	0.0610
S	1.0000	0.1990
T	0.9910	0.1850
U	0.9810	0.0890
V	0.7960	0.0690
W	1.0000	0.1390
X	1.0000	0.1580
Y	1.0000	0.0350
Z	0.9920	0.0850
a	1.0000	0.2060
b	1.0000	0.2210
c	1.0000	0.2240
d	0.9990	0.2280
e	1.0000	0.2110
f	1.0000	0.2160