



wwPDB EM Validation Summary Report ⓘ

May 19, 2024 – 01:24 am BST

PDB ID : 6TB4
EMDB ID : EMD-10438
Title : Structure of SAGA bound to TBP
Authors : Papai, G.; Frechard, A.; Kolesnikova, O.; Crucifix, C.; Schultz, P.; Ben-Shem, A.
Deposited on : 2019-10-31
Resolution : 3.80 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ 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.dev92
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

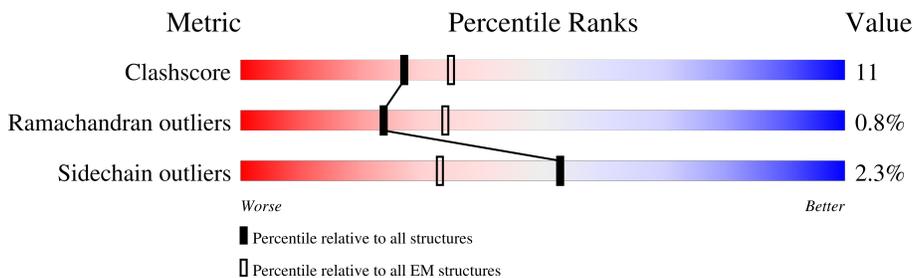
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 Å.

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



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

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 $\leq 5\%$. 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	M	243	
2	A	448	
3	C	698	
4	F	517	
5	D	341	
6	E	1191	
7	J	217	
8	K	609	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	G	722	
10	H	485	
11	I	153	
12	L	3825	
13	B	76	

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 40740 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 TATA-box Binding Protein (TBP).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	M	180	1415	921	242	246	6	0	0

- Molecule 2 is a protein called Transcriptional coactivator HFI1/ADA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	173	1300	816	228	250	6	0	0

- Molecule 3 is a protein called SAGA-associated factor 73 (Sgf73).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	65	518	331	94	90	3	0	0

- Molecule 4 is a protein called Spt20.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	F	210	1682	1071	292	315	4	0	0

- Molecule 5 is a protein called Subunit of the SAGA and SAGA-like transcriptional regulatory complexes, interacts with Spt15p to act.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	D	209	1616	1016	298	295	7	0	0

- Molecule 6 is a protein called Subunit of the SAGA transcriptional regulatory complex, involved in proper assembly of the complex.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	E	154	Total	C	N	O	S	0	0
			1232	784	208	233	7		

- Molecule 7 is a protein called Transcription initiation factor TFIID subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	J	96	Total	C	N	O	S	0	0
			768	489	120	156	3		

- Molecule 8 is a protein called Subunit (61/68 kDa) of TFIID and SAGA complexes.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	K	154	Total	C	N	O	S	0	0
			1192	747	216	226	3		

- Molecule 9 is a protein called Subunit (90 kDa) of TFIID and SAGA complexes.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	G	522	Total	C	N	O	S	0	0
			4075	2581	719	756	19		

- Molecule 10 is a protein called Subunit (60 kDa) of TFIID and SAGA complexes.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	H	421	Total	C	N	O	S	0	0
			3263	2084	556	617	6		

- Molecule 11 is a protein called Subunit (17 kDa) of TFIID and SAGA complexes, involved in RNA polymerase II transcription initiation.

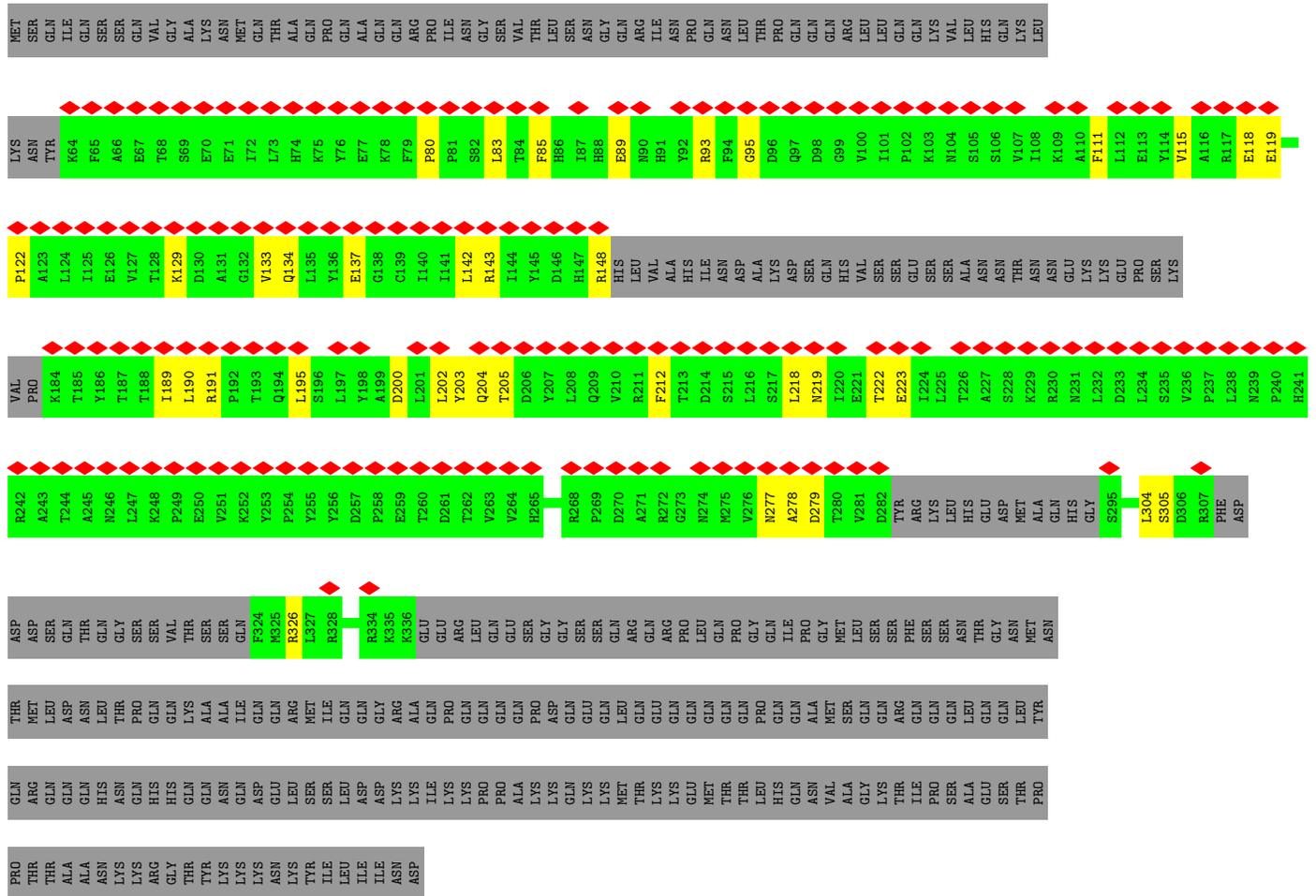
Mol	Chain	Residues	Atoms					AltConf	Trace
11	I	123	Total	C	N	O	S	0	0
			981	632	169	178	2		

- Molecule 12 is a protein called Transcription-associated protein.

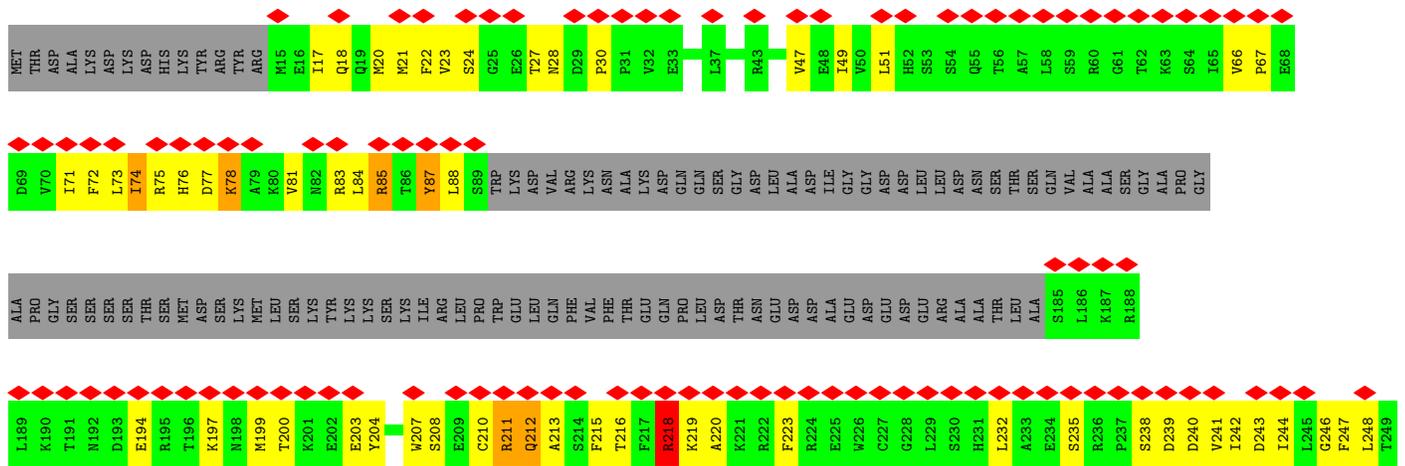
Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	2968	Total	C	N	O	S	0	0
			22318	14296	3864	4071	87		

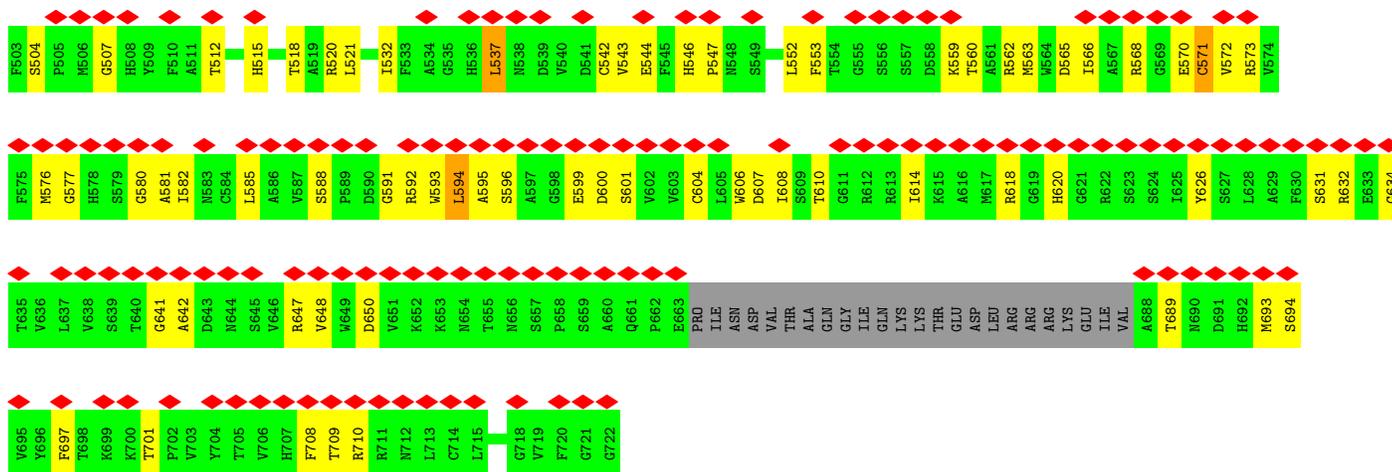
- Molecule 13 is a protein called Transcriptional adapter 3 (Ada3).

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	B	76	380	228	76	76	0	0

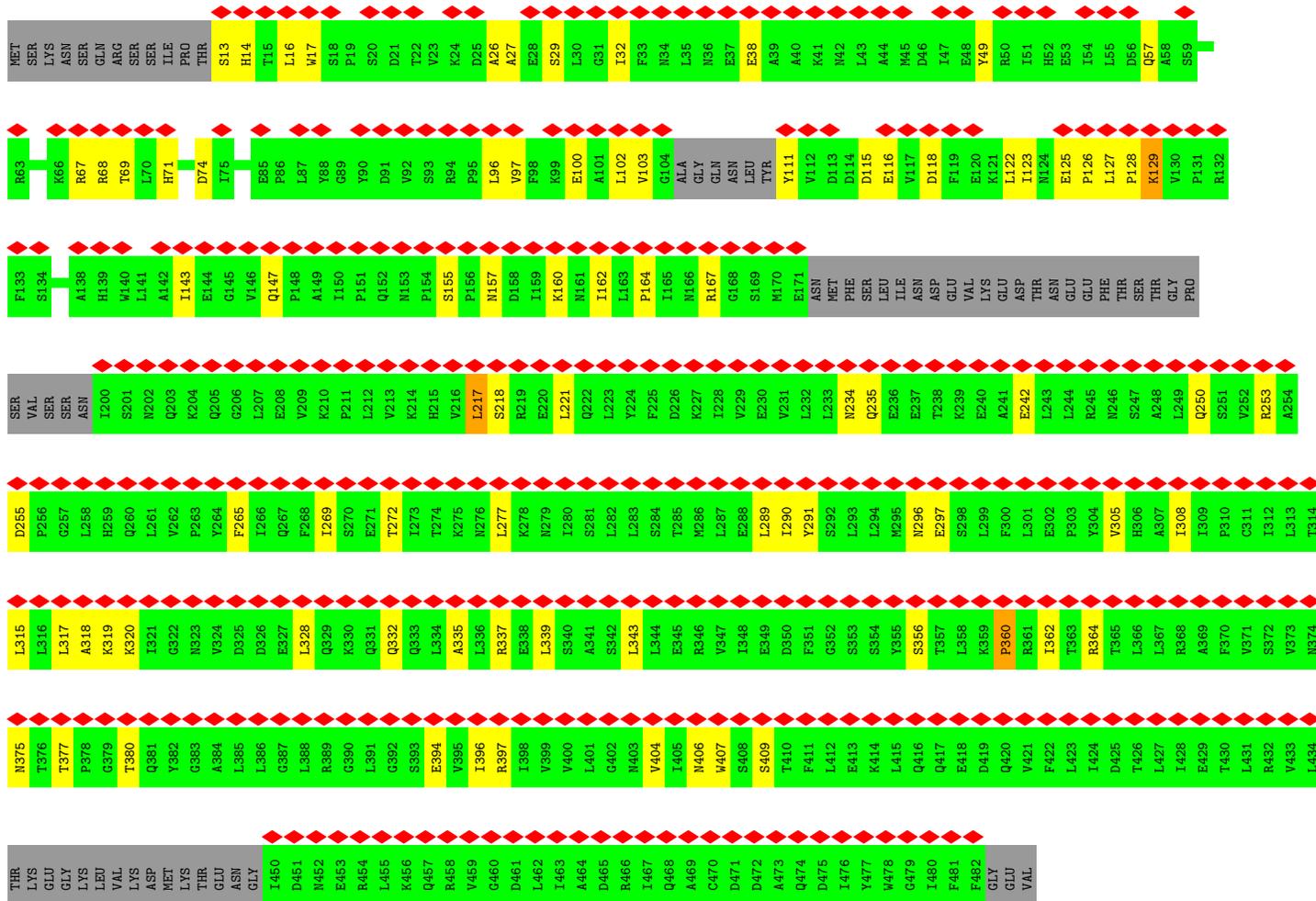
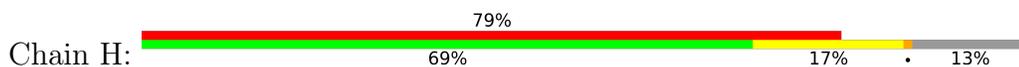


• Molecule 5: Subunit of the SAGA and SAGA-like transcriptional regulatory complexes, interacts with Spt15p to act

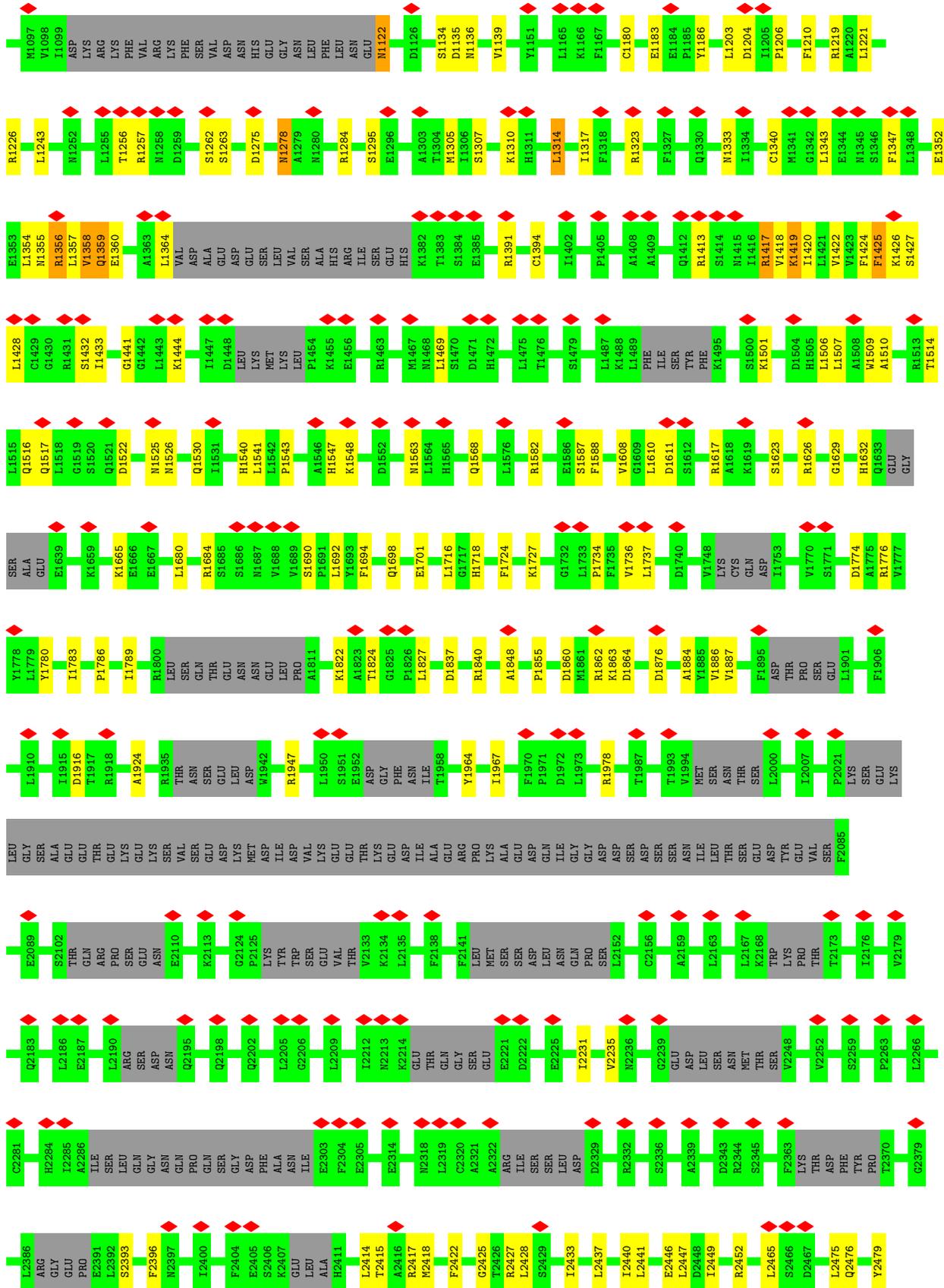


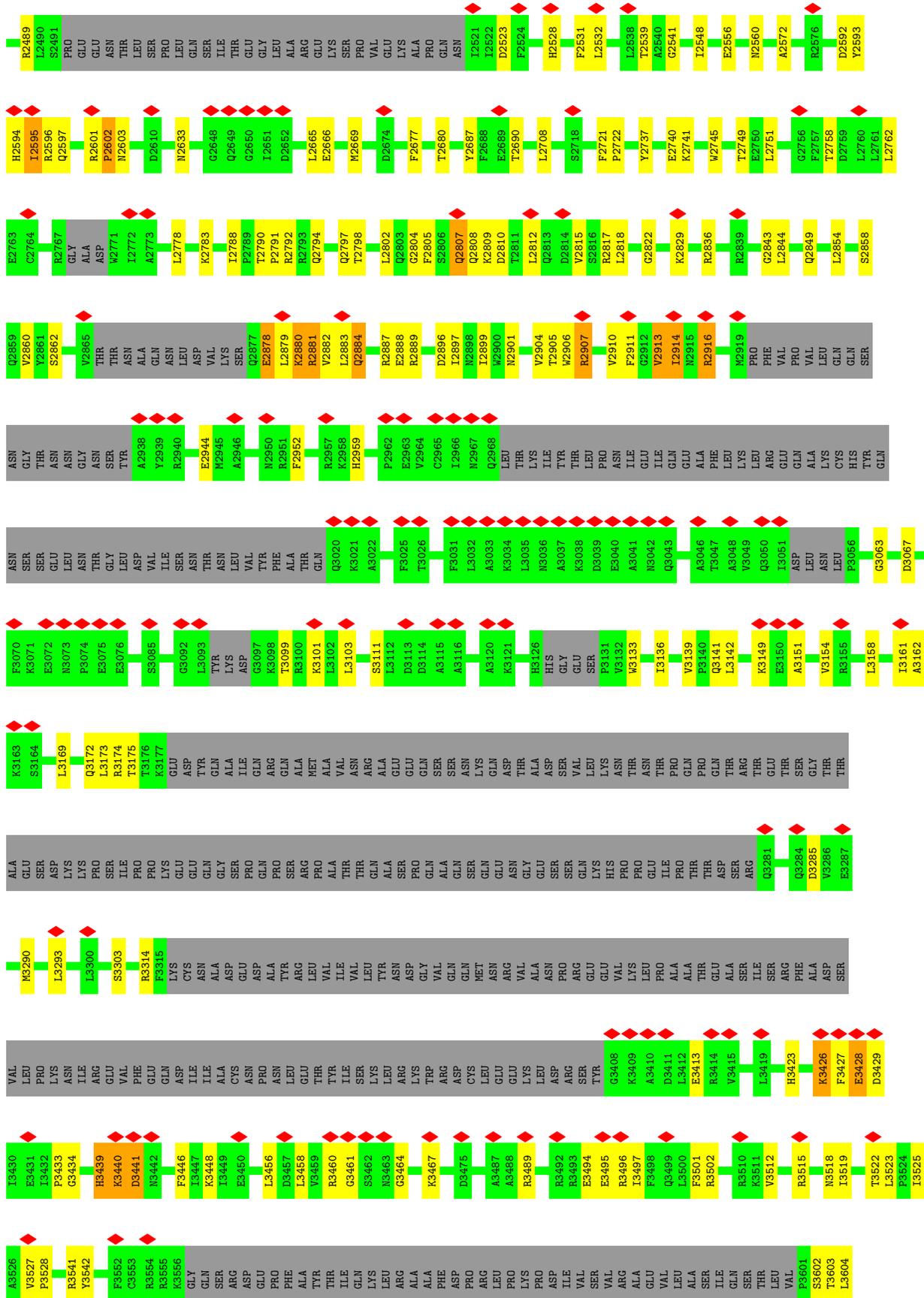


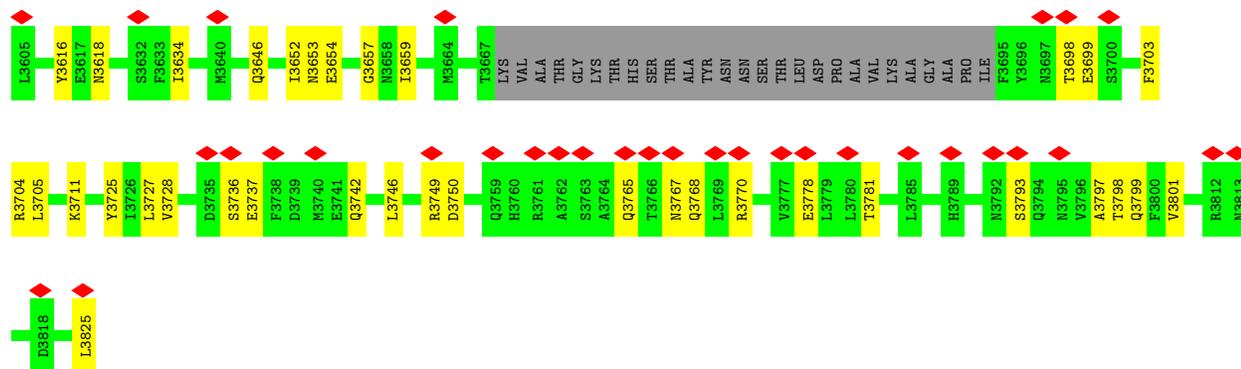
• Molecule 10: Subunit (60 kDa) of TFIID and SAGA complexes



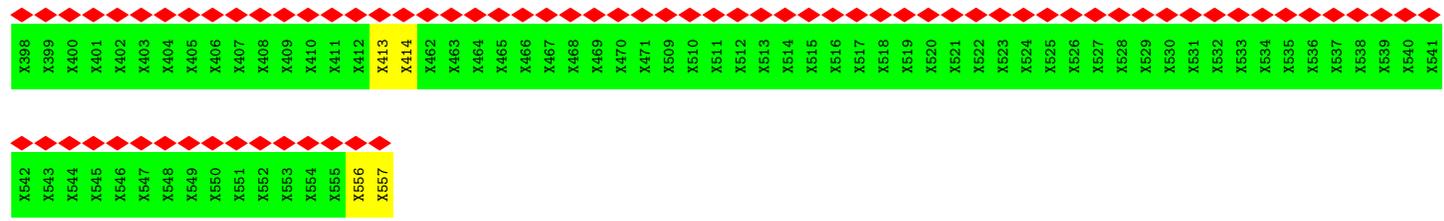
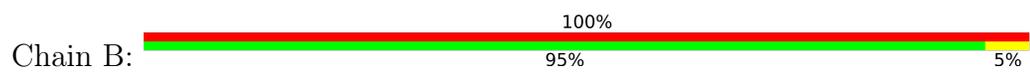
• Molecule 11: Subunit (17 kDa) of TFIID and SAGA complexes, involved in RNA polymerase II transcription initiation







• Molecule 13: Transcriptional adapter 3 (Ada3)



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	354104	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	52.8	Depositor
Minimum defocus (nm)	0.8	Depositor
Maximum defocus (nm)	4.5	Depositor
Magnification	105000	Depositor
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor
Maximum map value	3.125	Depositor
Minimum map value	-1.547	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.059	Depositor
Recommended contour level	0.7	Depositor
Map size (Å)	558.08, 558.08, 558.08	wwPDB
Map dimensions	512, 512, 512	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.09, 1.09, 1.09	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	M	0.88	2/1442 (0.1%)	0.78	2/1942 (0.1%)
2	A	0.44	0/1319	0.60	0/1794
3	C	0.35	0/528	0.61	0/710
4	F	0.34	0/1718	0.58	0/2335
5	D	0.51	0/1641	0.65	0/2213
6	E	0.45	0/1246	0.62	0/1667
7	J	0.47	0/779	0.60	0/1051
8	K	0.42	0/1213	0.66	0/1647
9	G	0.52	0/4177	0.60	0/5661
10	H	0.38	0/3315	0.60	0/4500
11	I	0.44	0/1006	0.63	0/1374
12	L	0.33	0/22712	0.54	0/30825
All	All	0.41	2/41096 (0.0%)	0.58	2/55719 (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	M	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	M	194	ILE	C-N	27.38	1.97	1.34
1	M	183	SER	C-N	6.38	1.48	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	M	62	GLY	N-CA-C	6.94	130.46	113.10
1	M	194	ILE	CA-C-N	-6.13	103.71	117.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	M	183	SER	Mainchain

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	M	1415	0	1492	88	0
2	A	1300	0	1254	30	0
3	C	518	0	528	9	0
4	F	1682	0	1622	30	0
5	D	1616	0	1558	122	0
6	E	1232	0	1276	40	0
7	J	768	0	754	11	0
8	K	1192	0	1214	89	0
9	G	4075	0	3934	102	0
10	H	3263	0	3258	69	0
11	I	981	0	982	18	0
12	L	22318	0	20960	376	0
13	B	380	0	85	2	0
All	All	40740	0	38917	865	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 865 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)
8:K:469:LEU:HD13	12:L:2788:ILE:CG2	1.31	1.51
8:K:469:LEU:CD1	12:L:2788:ILE:HG21	1.57	1.34
12:L:828:ARG:H	12:L:829:PRO:HD3	1.06	1.20
1:M:194:ILE:HG22	1:M:195:TYR:N	1.57	1.19
1:M:194:ILE:C	1:M:195:TYR:N	1.97	1.17

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	M	178/243 (73%)	147 (83%)	24 (14%)	7 (4%)	3	28
2	A	169/448 (38%)	151 (89%)	18 (11%)	0	100	100
3	C	61/698 (9%)	56 (92%)	5 (8%)	0	100	100
4	F	202/517 (39%)	177 (88%)	24 (12%)	1 (0%)	29	66
5	D	203/341 (60%)	177 (87%)	19 (9%)	7 (3%)	3	31
6	E	150/1191 (13%)	135 (90%)	13 (9%)	2 (1%)	12	48
7	J	92/217 (42%)	82 (89%)	9 (10%)	1 (1%)	14	51
8	K	150/609 (25%)	125 (83%)	23 (15%)	2 (1%)	12	48
9	G	512/722 (71%)	448 (88%)	63 (12%)	1 (0%)	47	79
10	H	413/485 (85%)	365 (88%)	48 (12%)	0	100	100
11	I	119/153 (78%)	98 (82%)	21 (18%)	0	100	100
12	L	2874/3825 (75%)	2623 (91%)	233 (8%)	18 (1%)	25	62
All	All	5123/9449 (54%)	4584 (90%)	500 (10%)	39 (1%)	24	57

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	77	GLY
1	M	99	PHE
1	M	110	LYS
5	D	200	THR
5	D	327	ARG

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	M	152/208 (73%)	138 (91%)	14 (9%)	9	35
2	A	134/394 (34%)	132 (98%)	2 (2%)	65	81
3	C	55/627 (9%)	55 (100%)	0	100	100
4	F	179/471 (38%)	179 (100%)	0	100	100
5	D	166/306 (54%)	154 (93%)	12 (7%)	14	45
6	E	142/1101 (13%)	136 (96%)	6 (4%)	30	58
7	J	85/183 (46%)	84 (99%)	1 (1%)	71	84
8	K	133/524 (25%)	130 (98%)	3 (2%)	50	72
9	G	439/635 (69%)	430 (98%)	9 (2%)	53	74
10	H	352/438 (80%)	348 (99%)	4 (1%)	73	85
11	I	104/130 (80%)	103 (99%)	1 (1%)	76	86
12	L	2156/3450 (62%)	2114 (98%)	42 (2%)	57	76
All	All	4097/8467 (48%)	4003 (98%)	94 (2%)	53	72

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	1122	ASN
12	L	2489	ARG
12	L	1314	LEU
12	L	1419	LYS
12	L	2878	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
10	H	64	HIS
10	H	157	ASN
12	L	2884	GLN
12	L	2633	ASN
12	L	2698	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
13	B	2
1	M	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	414:UNK	C	462:UNK	N	28.86
1	B	471:UNK	C	509:UNK	N	23.98
1	M	194:ILE	C	195:TYR	N	1.97

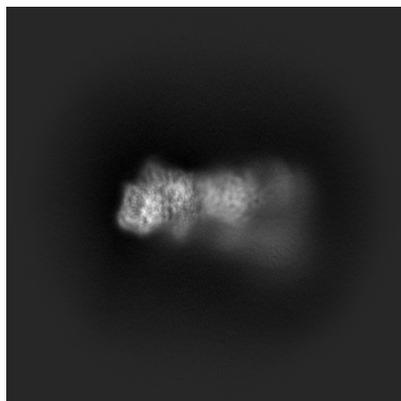
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-10438. These allow visual inspection of the internal detail of the map and identification of artifacts.

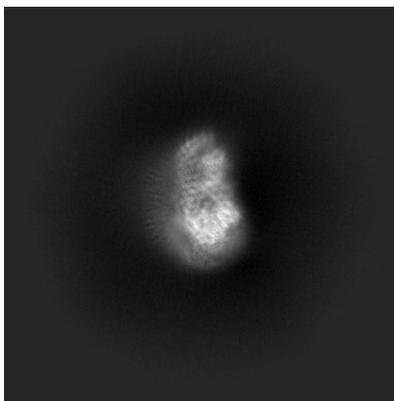
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

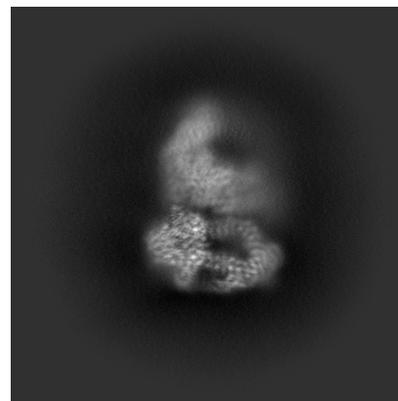
6.1.1 Primary map



X

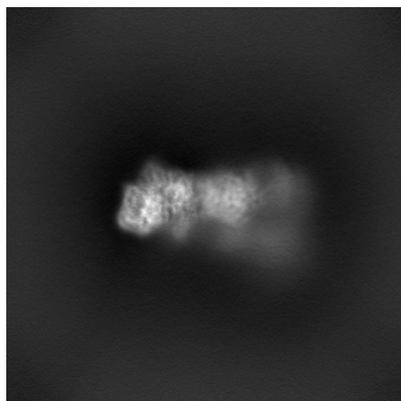


Y

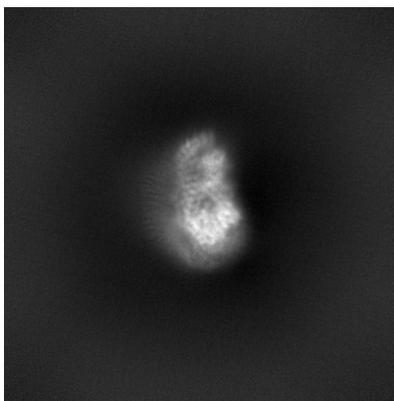


Z

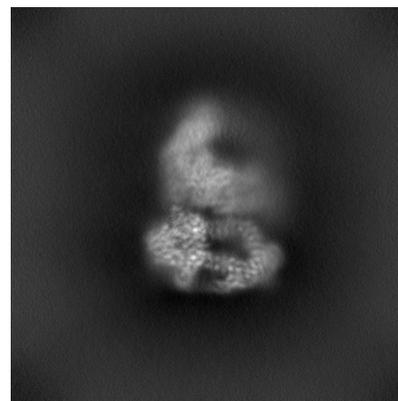
6.1.2 Raw map



X



Y

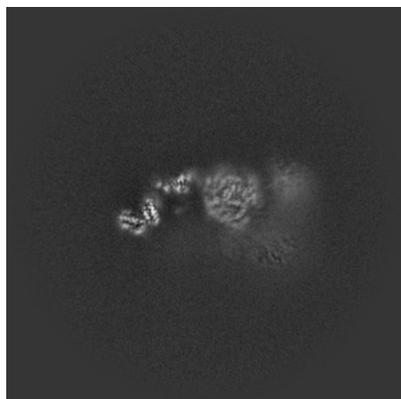


Z

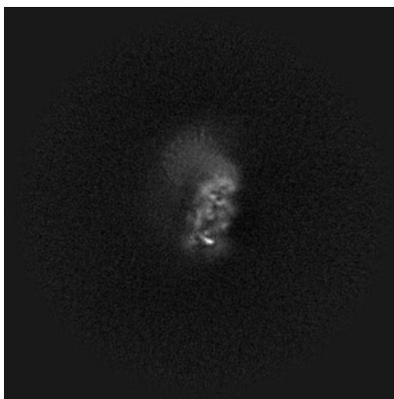
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

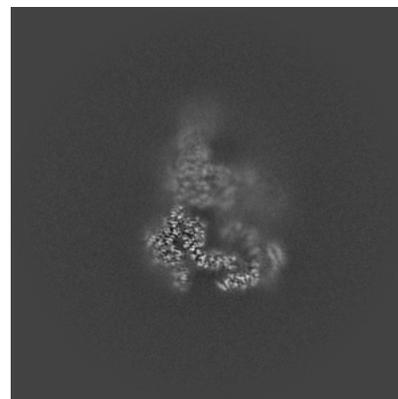
6.2.1 Primary map



X Index: 256

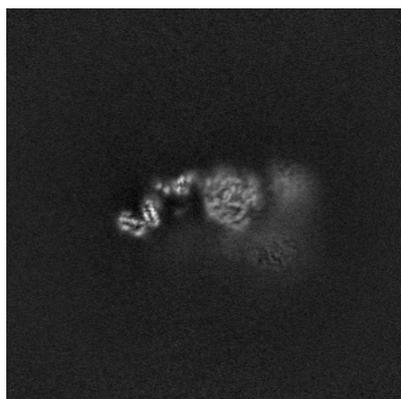


Y Index: 256



Z Index: 256

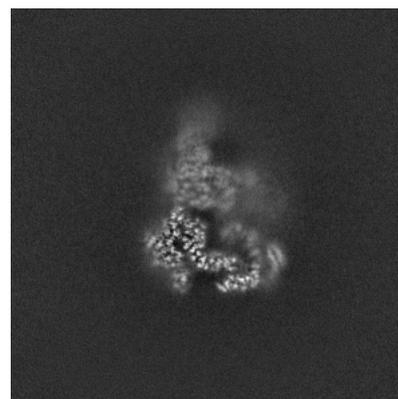
6.2.2 Raw map



X Index: 256



Y Index: 256

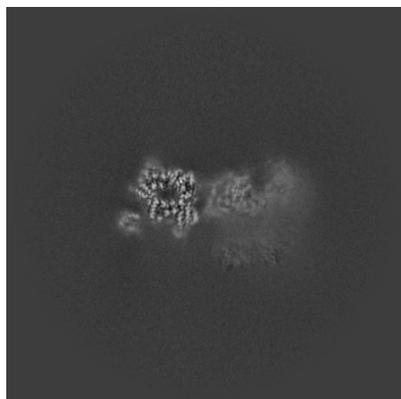


Z Index: 256

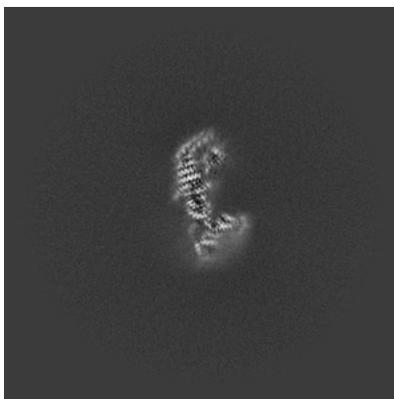
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

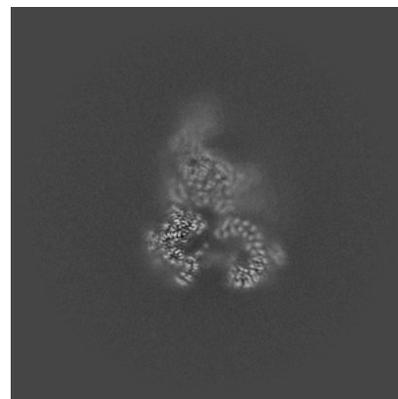
6.3.1 Primary map



X Index: 236

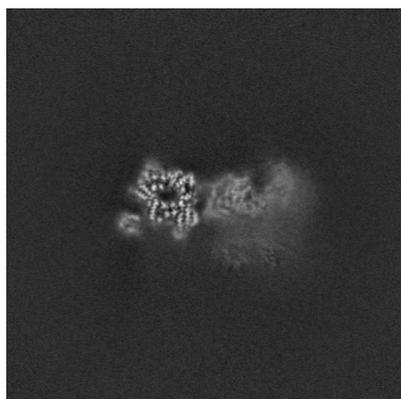


Y Index: 182



Z Index: 268

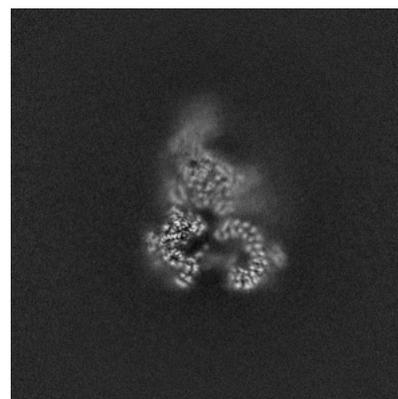
6.3.2 Raw map



X Index: 236



Y Index: 182

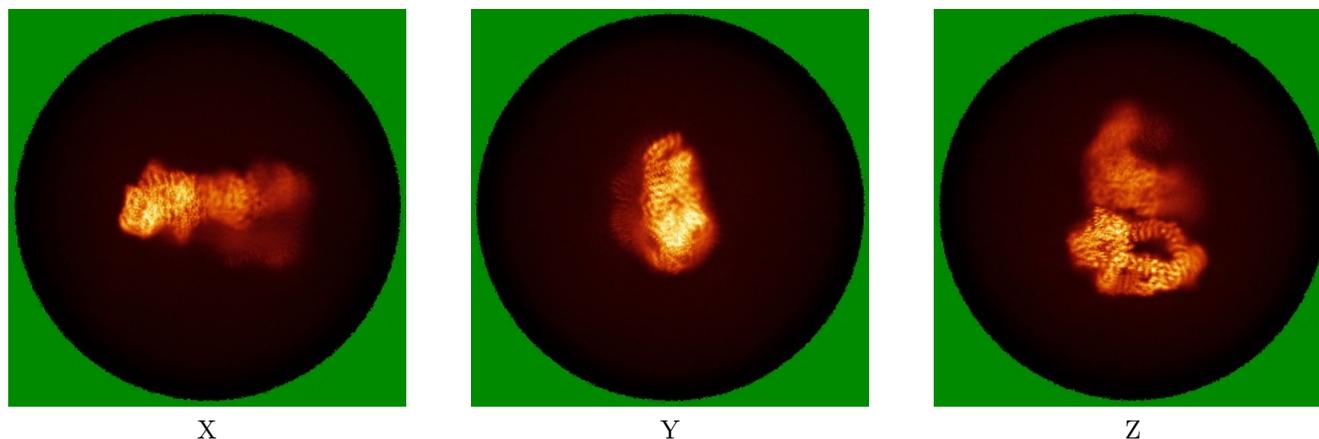


Z Index: 268

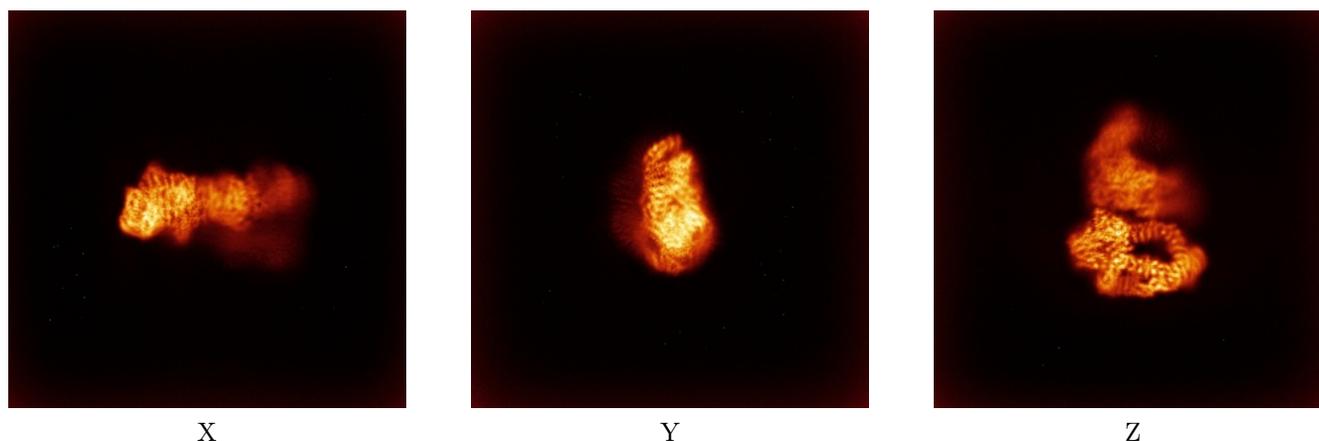
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

6.4.1 Primary map



6.4.2 Raw map



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



The images above show the 3D surface view of the map at the recommended contour level 0.7. 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



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.

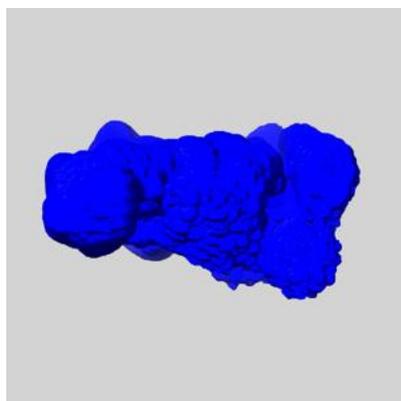
6.6 Mask visualisation [i](#)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

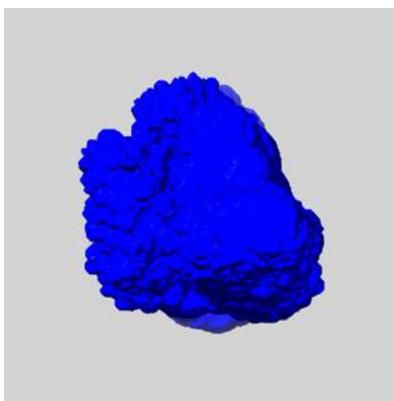
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

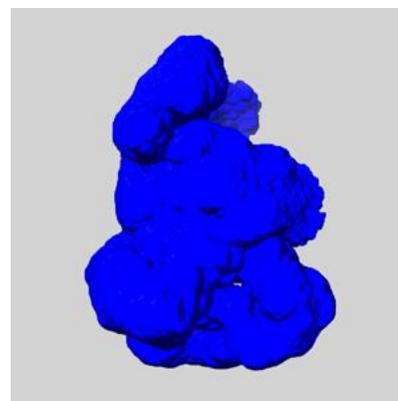
6.6.1 emd_10438_msk_1.map [i](#)



X



Y

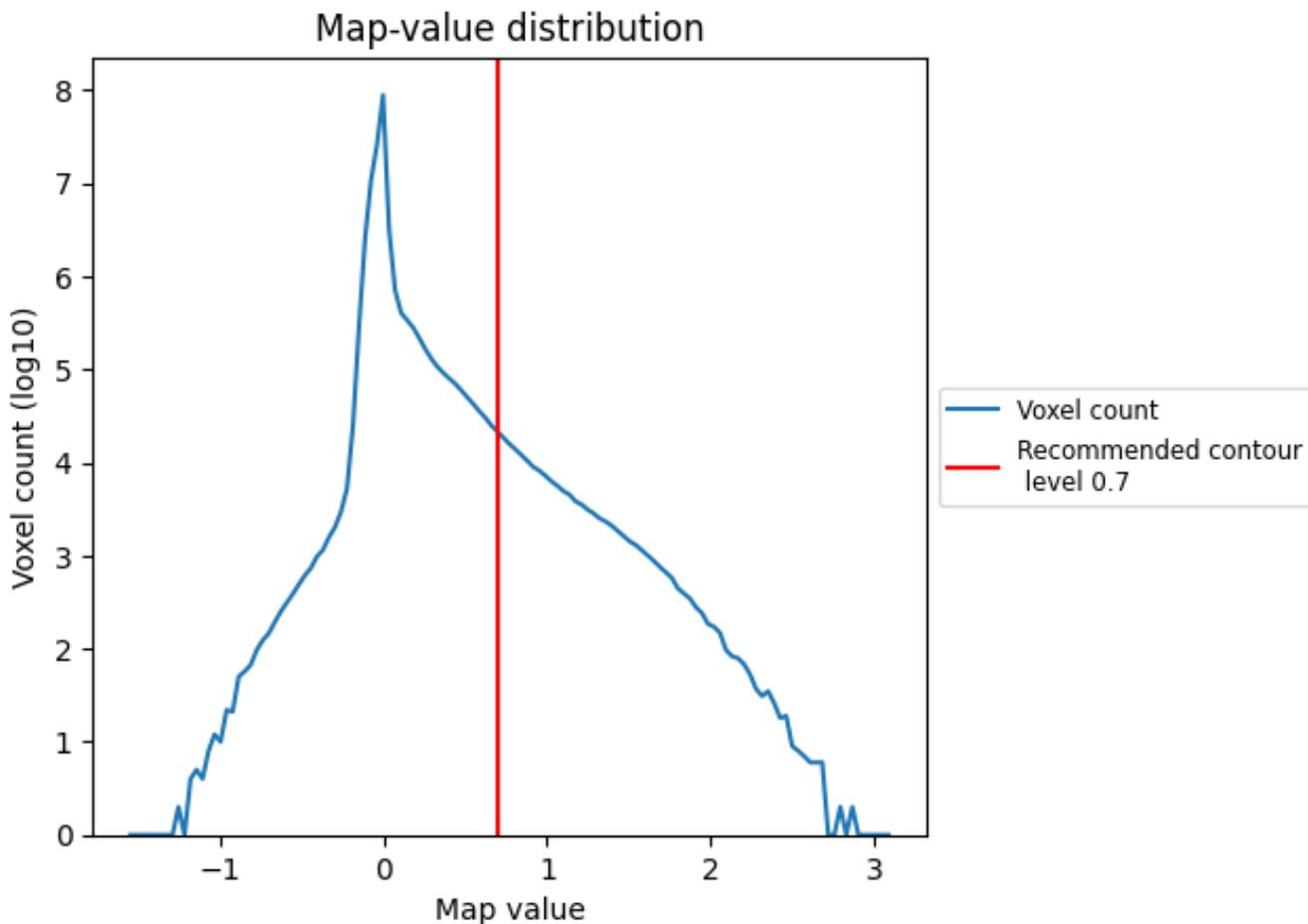


Z

7 Map analysis [i](#)

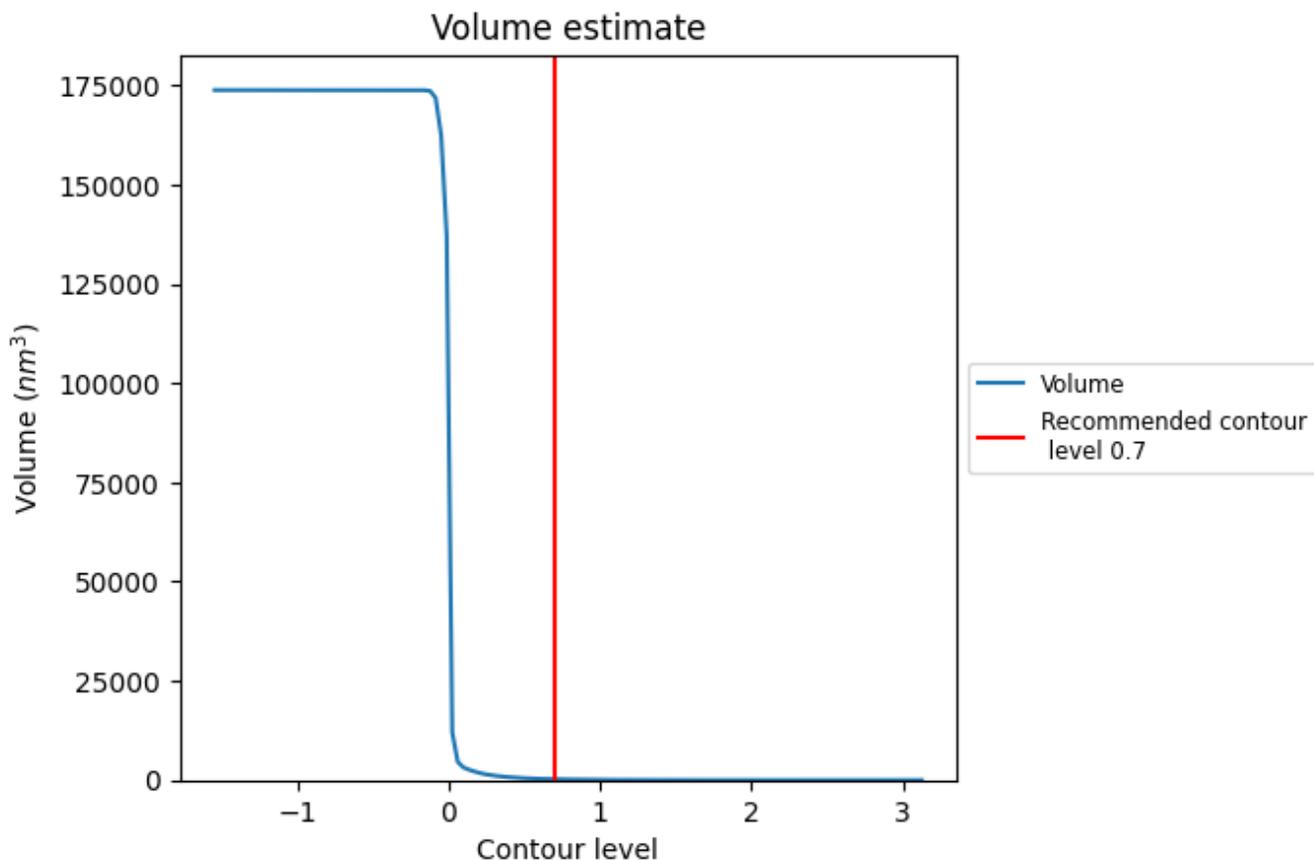
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



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

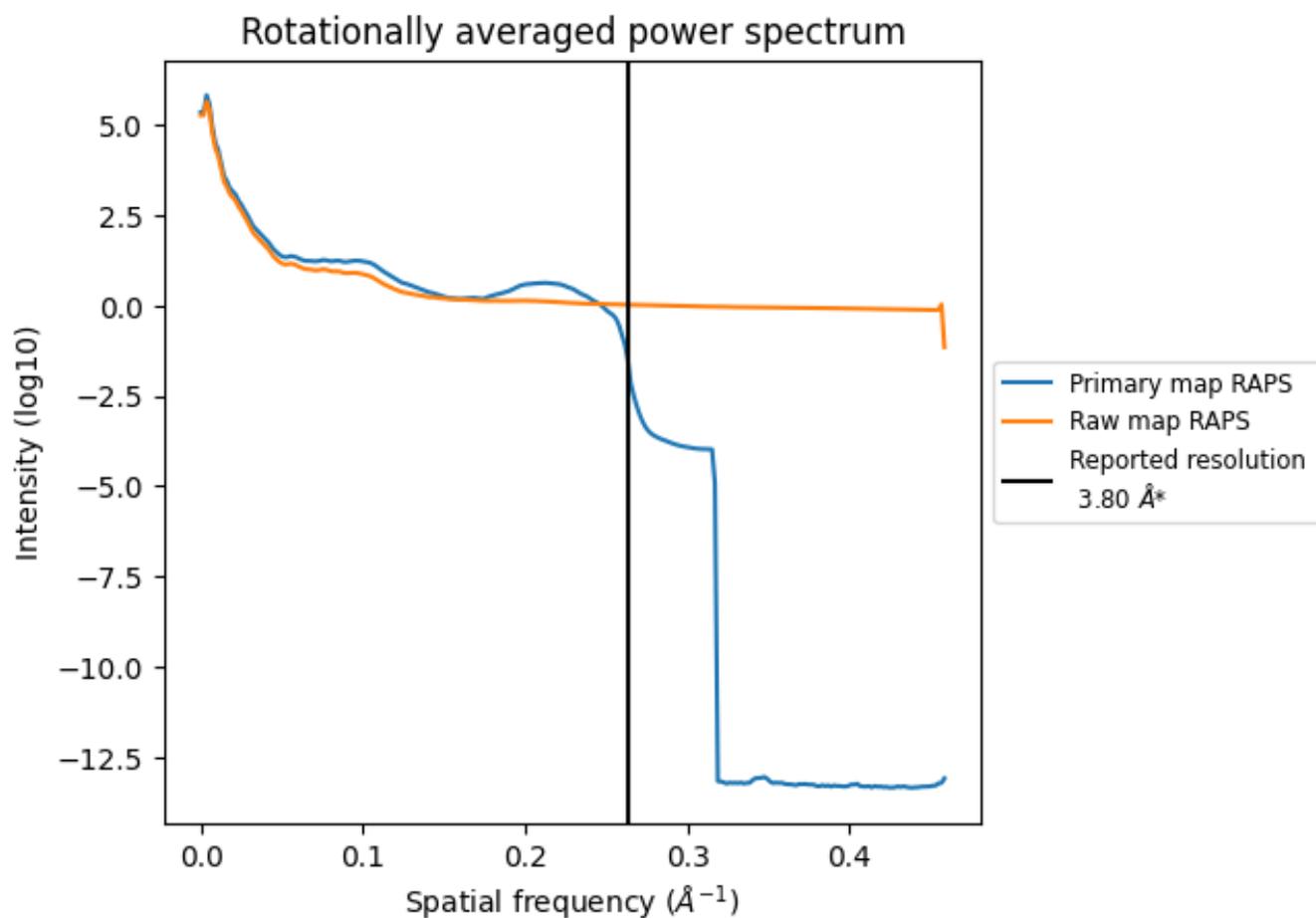
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 221 nm³; this corresponds to an approximate mass of 199 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](#)

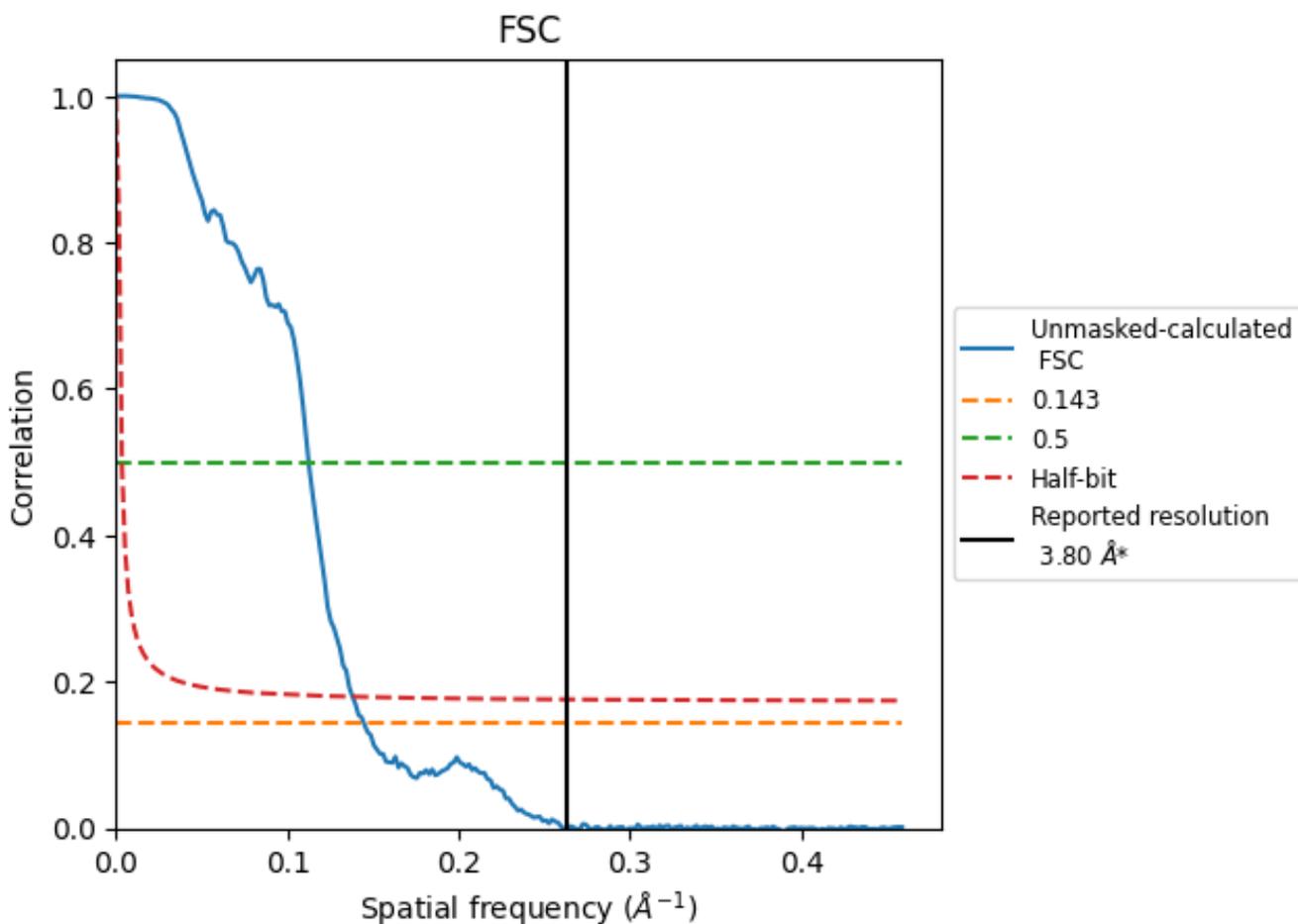


*Reported resolution corresponds to spatial frequency of 0.263 Å⁻¹

8 Fourier-Shell correlation [i](#)

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 [i](#)



*Reported resolution corresponds to spatial frequency of 0.263 \AA^{-1}

8.2 Resolution estimates [i](#)

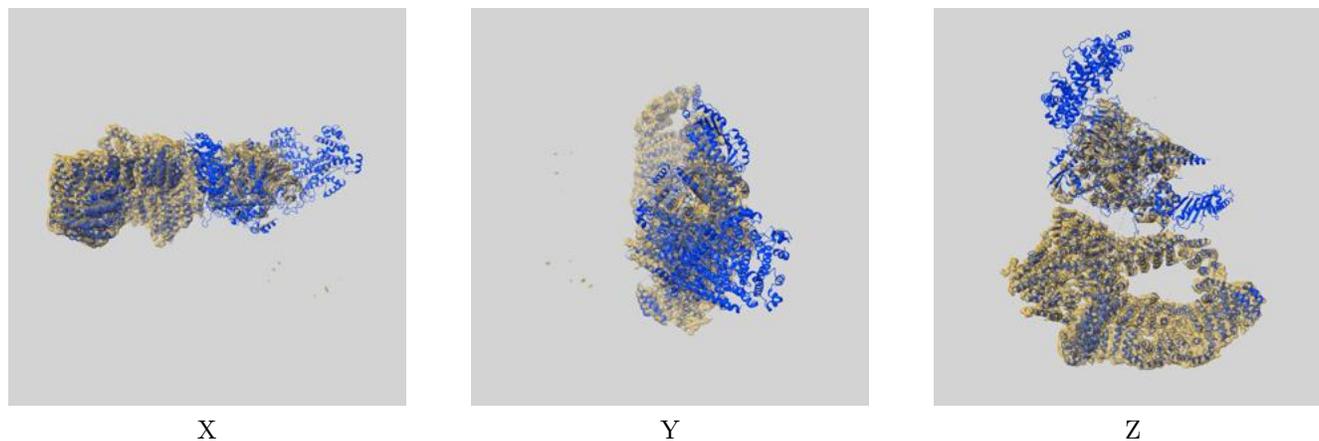
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.80	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	6.91	8.89	7.24

*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 6.91 differs from the reported value 3.8 by more than 10 %

9 Map-model fit [i](#)

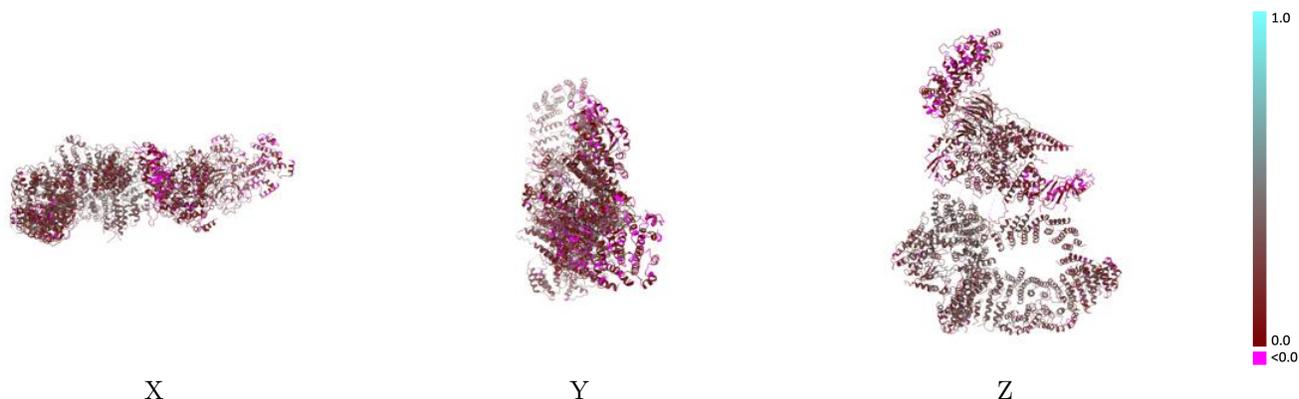
This section contains information regarding the fit between EMDB map EMD-10438 and PDB model 6TB4. Per-residue inclusion information can be found in section 3 on page 7.

9.1 Map-model overlay [i](#)



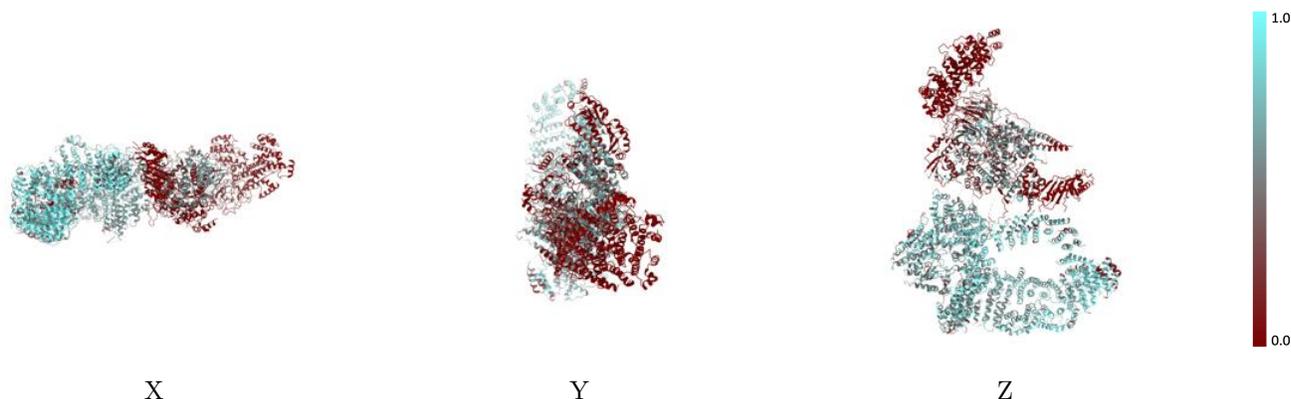
The images above show the 3D surface view of the map at the recommended contour level 0.7 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



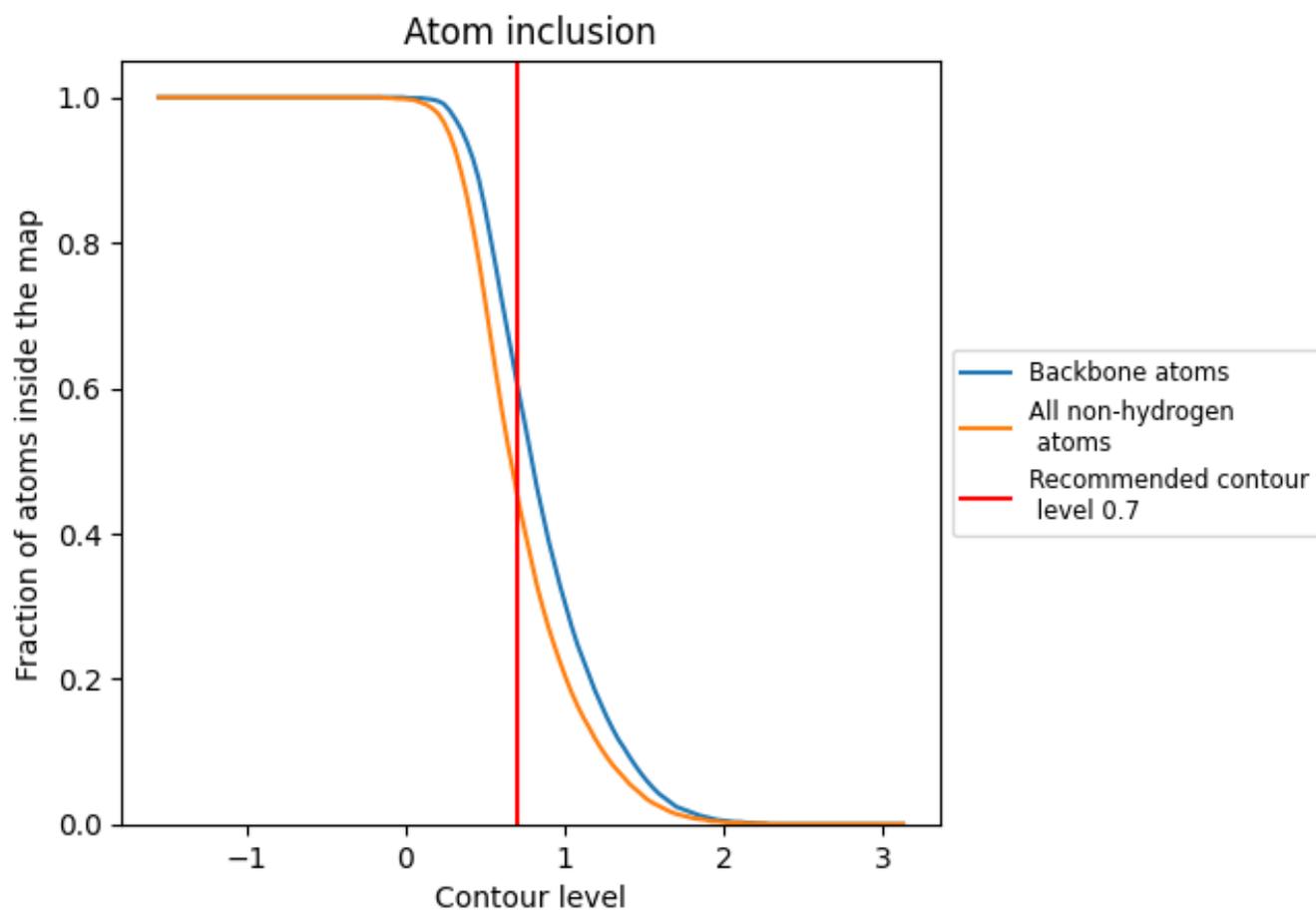
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.7).

9.4 Atom inclusion [i](#)



At the recommended contour level, 61% of all backbone atoms, 46% 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 (0.7) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.4560	 0.2310
A	 0.2350	 0.1830
B	 0.0030	 0.1280
C	 0.0900	 0.1570
D	 0.2950	 0.1690
E	 0.4000	 0.1980
F	 0.1990	 0.1540
G	 0.1960	 0.1800
H	 0.0850	 0.1350
I	 0.2930	 0.1750
J	 0.4020	 0.2060
K	 0.3290	 0.2160
L	 0.6650	 0.2880
M	 0.0020	 0.0490

