Full wwPDB EM Validation Report ⓘ

May 27, 2024 – 12:29 PM JST

PDB ID	:	7EGF
EMDB ID	:	EMD-31115
Title	:	TFIID lobe A subcomplex
Authors	:	Chen, X.; Wu, Z.; Li, J.; Zhao, D.; Xu, Y.
Deposited on	:	2021-03-24
Resolution	:	3.16 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:

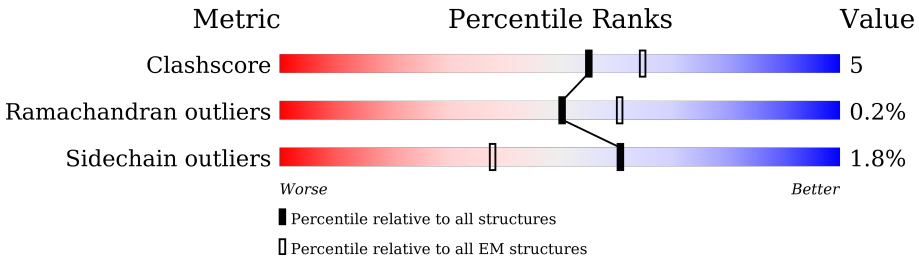
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.16 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.



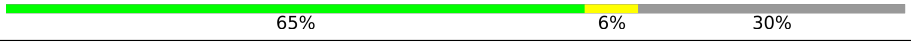
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 <=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	P	339	48% · 48%
2	c	929	13% 86%
3	d	1085	14% 85%
4	e	800	66% · 33%
5	f	677	21% 79%
6	i	264	46% 54%
7	j	218	42% · 56%
8	k	211	46% 54%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	l	161	
10	m	124	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 13303 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	P	177	Total	C	N	O	S	0	0
			1412	918	249	238	7		

- Molecule 2 is a protein called Transcription initiation factor TFIID subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	c	127	Total	C	N	O	S	0	0
			1011	638	174	193	6		

- Molecule 3 is a protein called Transcription initiation factor TFIID subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	d	158	Total	C	N	O	S	0	0
			1307	814	238	252	3		

- Molecule 4 is a protein called Transcription initiation factor TFIID subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	e	539	Total	C	N	O	S	0	0
			4327	2746	748	814	19		

- Molecule 5 is a protein called Transcription initiation factor TFIID subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	f	143	Total	C	N	O	S	0	0
			1135	727	185	216	7		

- Molecule 6 is a protein called Transcription initiation factor TFIID subunit 9.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	i	121	Total	C	N	O	S	0	0
			967	615	167	178	7		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	j	95	Total	C	N	O	S	0	0
			759	488	124	143	4		

- Molecule 8 is a protein called Transcription initiation factor TFIID subunit 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	k	98	Total	C	N	O	S	0	0
			785	499	142	139	5		

- Molecule 9 is a protein called Transcription initiation factor TFIID subunit 12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	l	107	Total	C	N	O	S	0	0
			876	547	158	166	5		

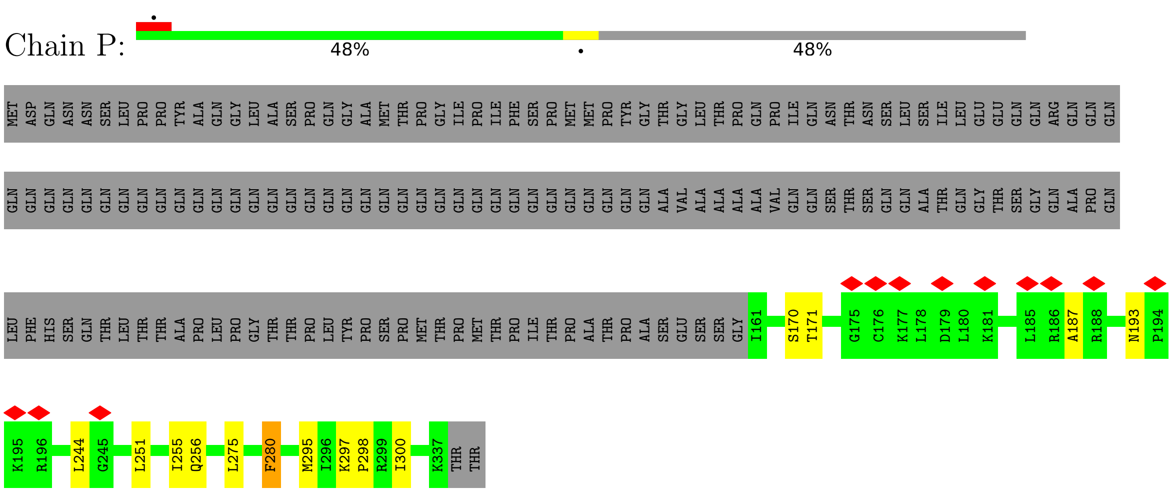
- Molecule 10 is a protein called Transcription initiation factor TFIID subunit 13.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	m	87	Total	C	N	O	S	0	0
			724	456	131	131	6		

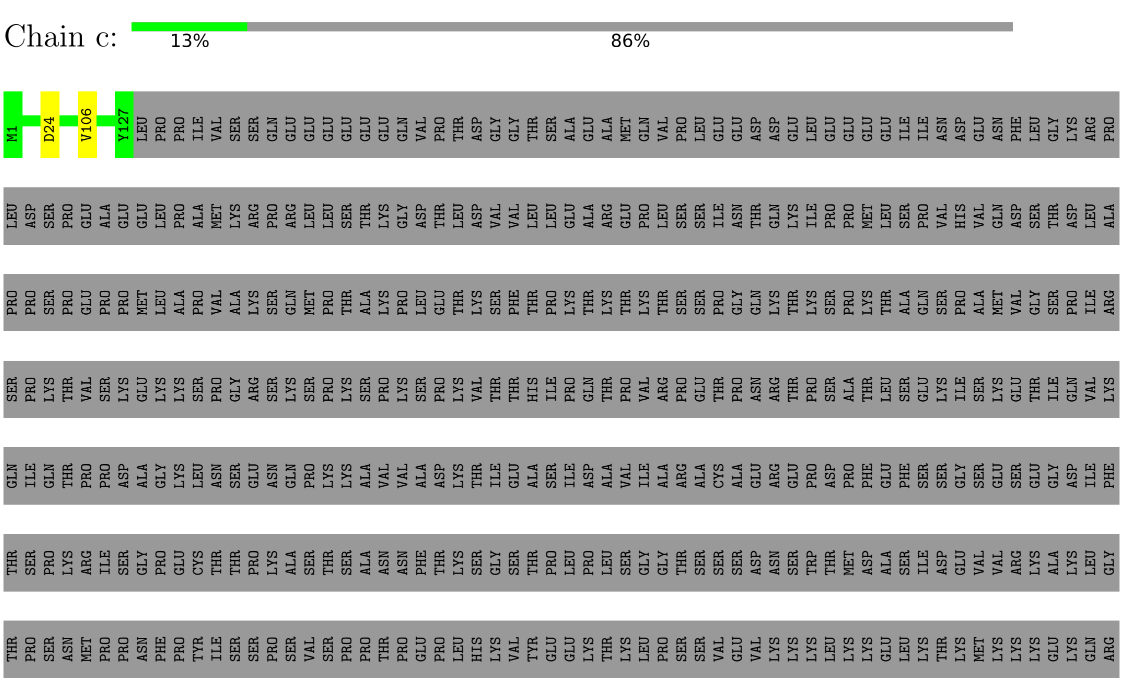
3 Residue-property plots

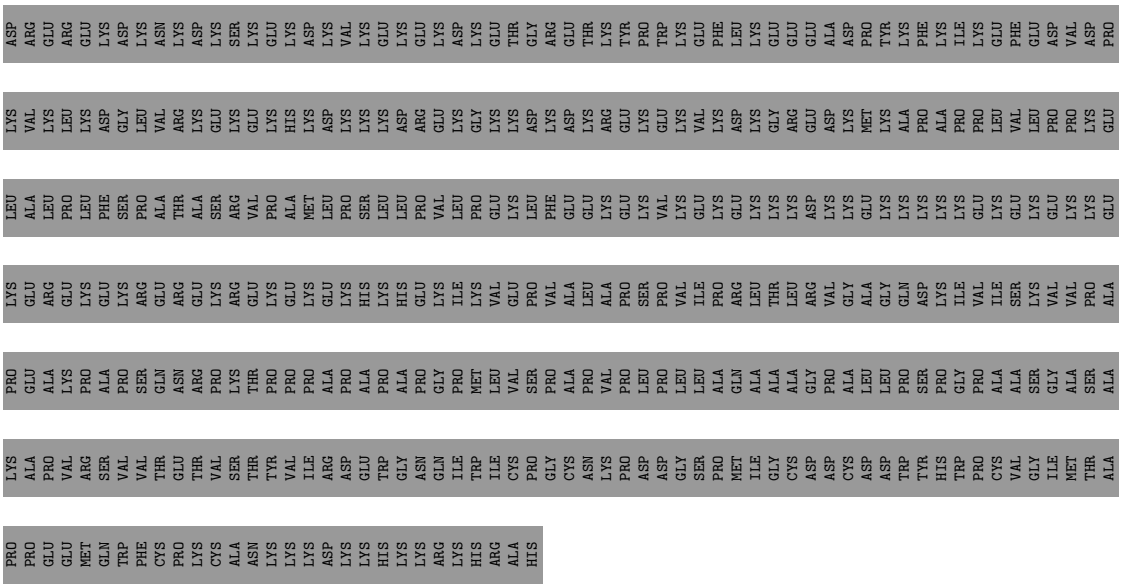
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: TATA-box-binding protein



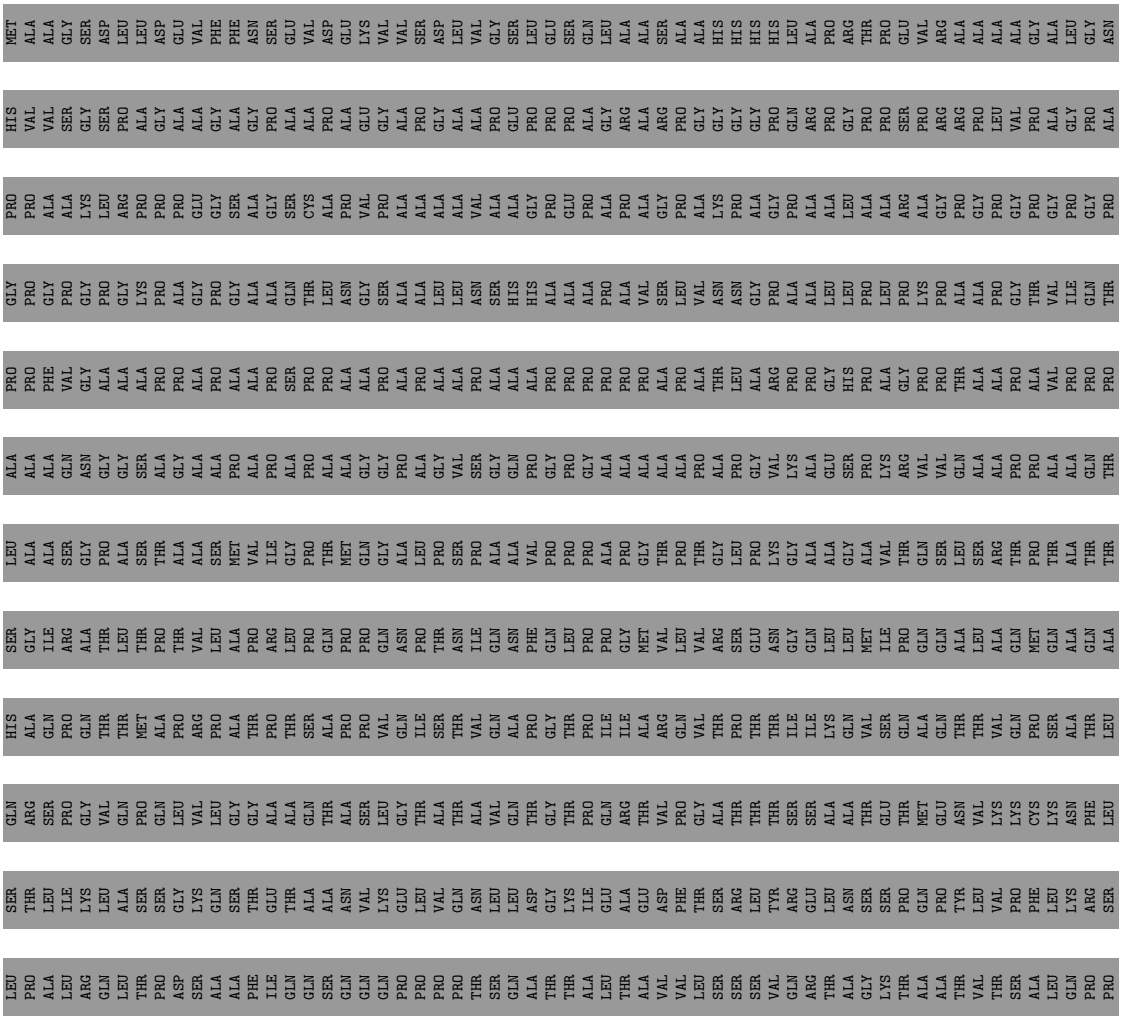
- Molecule 2: Transcription initiation factor TFIID subunit 3

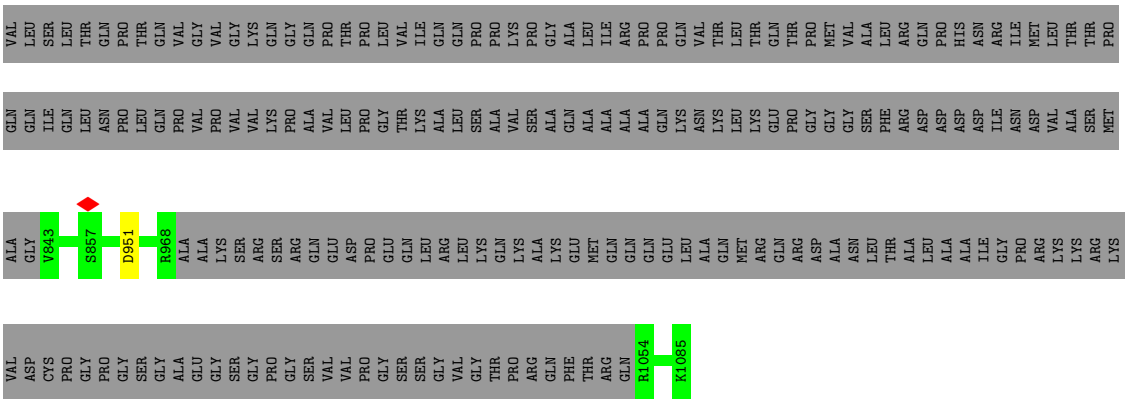




• Molecule 3: Transcription initiation factor TFIID subunit 4

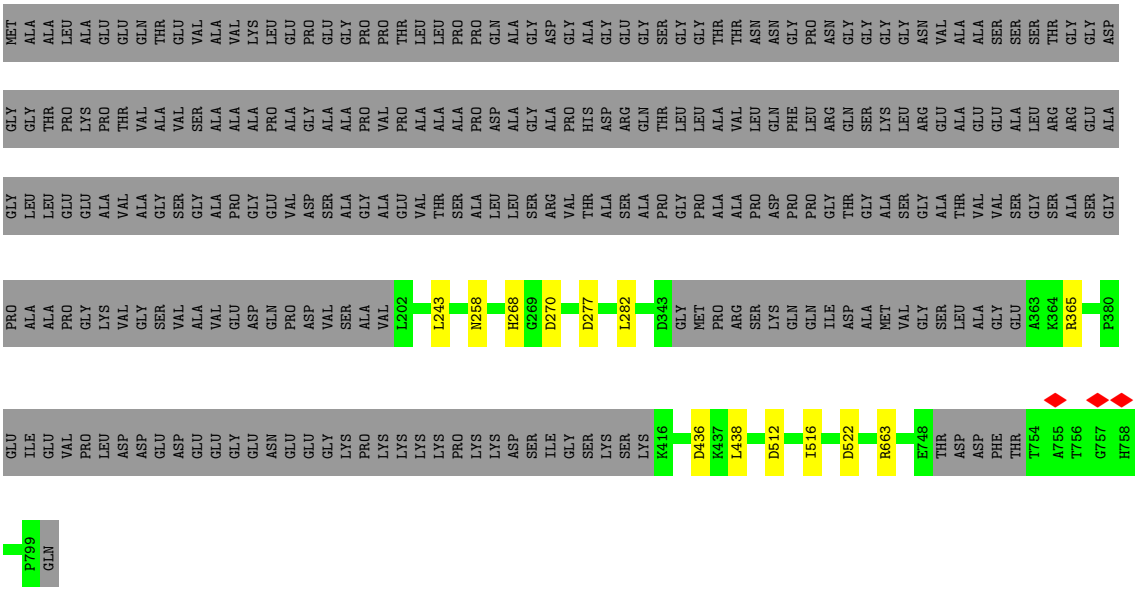
Chain d: 14% 85%





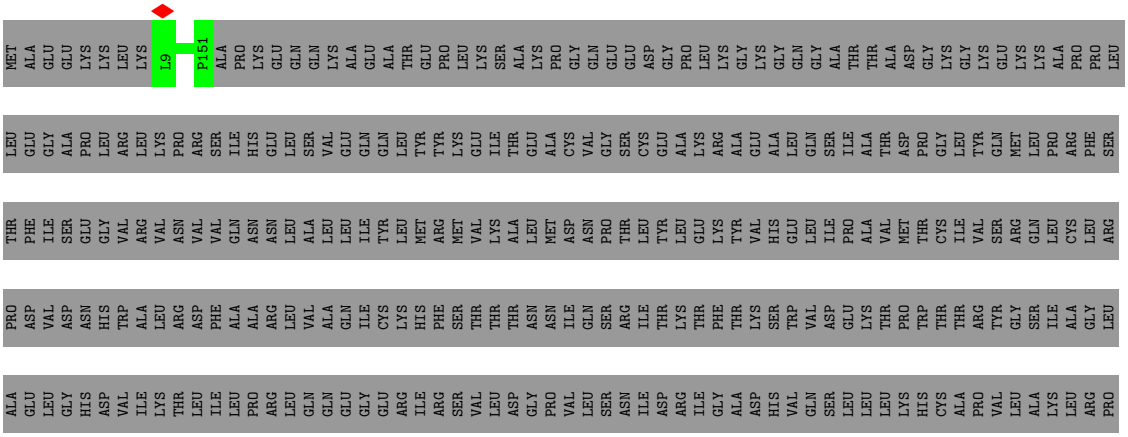
• Molecule 4: Transcription initiation factor TFIID subunit 5

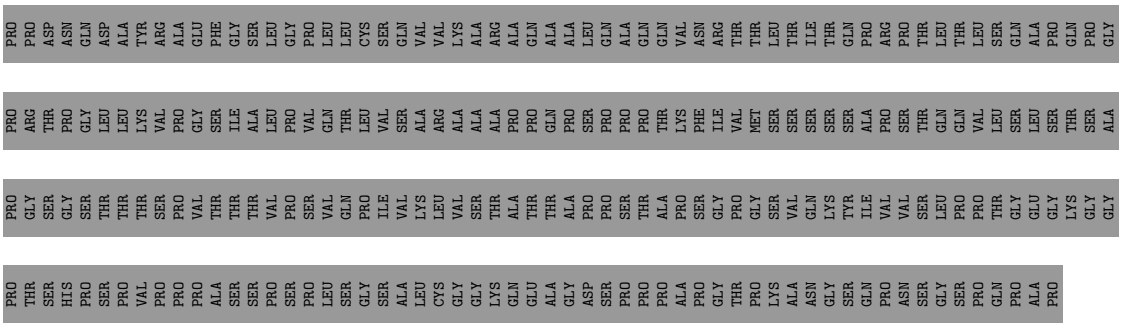
Chain e:



• Molecule 5: Transcription initiation factor TFIID subunit 6

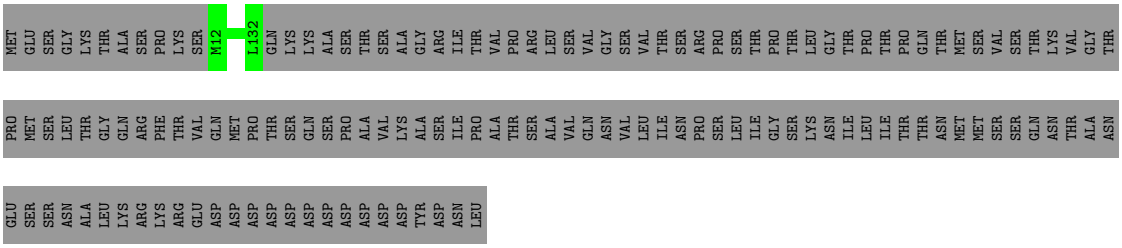
Chain f:





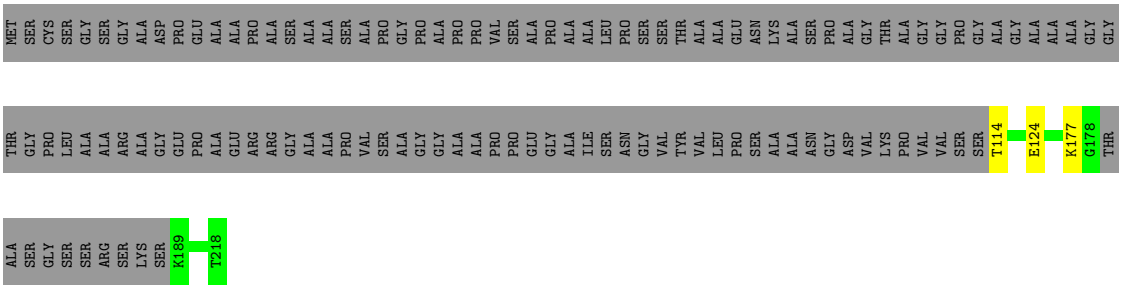
• Molecule 6: Transcription initiation factor TFIID subunit 9

Chain i:



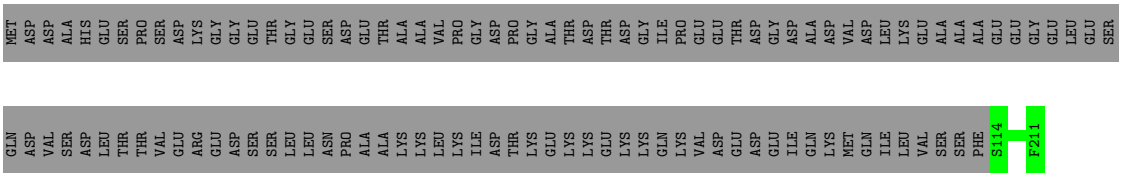
• Molecule 7: Transcription initiation factor TFIID subunit 10

Chain j:



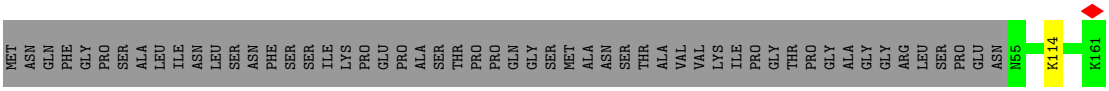
• Molecule 8: Transcription initiation factor TFIID subunit 11

Chain k:

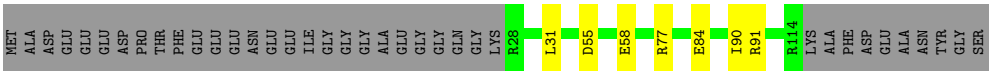


• Molecule 9: Transcription initiation factor TFIID subunit 12

Chain l:



• Molecule 10: Transcription initiation factor TFIID subunit 13



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	15078	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 ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	4.942	Depositor
Minimum map value	-2.786	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.065	Depositor
Recommended contour level	0.106	Depositor
Map size (Å)	337.59998, 337.59998, 337.59998	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.055, 1.055, 1.055	Depositor

5 Model quality

5.1 Standard geometry

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	P	0.46	0/1438	0.57	0/1935
2	c	0.39	0/1035	0.54	0/1406
3	d	0.38	0/1321	0.53	0/1772
4	e	0.43	0/4433	0.59	0/6004
5	f	0.29	0/1161	0.54	0/1574
6	i	0.29	0/989	0.46	0/1343
7	j	0.53	0/775	0.63	0/1049
8	k	0.30	0/799	0.47	0/1070
9	l	0.43	0/888	0.55	0/1194
10	m	0.59	0/733	0.64	0/977
All	All	0.42	0/13572	0.56	0/18324

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	P	1412	0	1506	9	0
2	c	1011	0	975	0	0
3	d	1307	0	1318	0	0
4	e	4327	0	4197	0	0
5	f	1135	0	1131	0	0
6	i	967	0	977	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	j	759	0	759	0	0
8	k	785	0	818	0	0
9	l	876	0	893	0	0
10	m	724	0	744	0	0
All	All	13303	0	13318	9	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 5.

All (9) 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:P:280:PHE:CD1	1:P:280:PHE:N	2.79	0.50
1:P:298:PRO:HB2	1:P:300:ILE:HG12	1.95	0.48
1:P:170:SER:HA	1:P:255:ILE:HA	1.95	0.47
1:P:171:THR:OG1	1:P:256:GLN:HG2	2.17	0.45
1:P:170:SER:HB3	1:P:255:ILE:HG12	1.99	0.45
1:P:295:MET:HG2	1:P:298:PRO:HD2	1.97	0.45
1:P:187:ALA:HB2	1:P:244:LEU:HD11	2.03	0.40
1:P:275:LEU:HD23	1:P:275:LEU:HA	1.93	0.40
1:P:297:LYS:HB3	1:P:298:PRO:HD3	2.03	0.40

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	P	175/339 (52%)	165 (94%)	9 (5%)	1 (1%)	25 62
2	c	125/929 (14%)	116 (93%)	9 (7%)	0	100 100
3	d	154/1085 (14%)	150 (97%)	4 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	e	531/800 (66%)	482 (91%)	48 (9%)	1 (0%)	47 78
5	f	141/677 (21%)	128 (91%)	13 (9%)	0	100 100
6	i	119/264 (45%)	115 (97%)	4 (3%)	0	100 100
7	j	91/218 (42%)	85 (93%)	4 (4%)	2 (2%)	6 32
8	k	96/211 (46%)	91 (95%)	5 (5%)	0	100 100
9	l	105/161 (65%)	100 (95%)	5 (5%)	0	100 100
10	m	85/124 (68%)	82 (96%)	3 (4%)	0	100 100
All	All	1622/4808 (34%)	1514 (93%)	104 (6%)	4 (0%)	50 78

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	e	522	ASP
7	j	124	GLU
7	j	177	LYS
1	P	251	LEU

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	P	153/293 (52%)	151 (99%)	2 (1%)	69 86
2	c	113/833 (14%)	111 (98%)	2 (2%)	59 81
3	d	146/815 (18%)	145 (99%)	1 (1%)	84 93
4	e	475/657 (72%)	463 (98%)	12 (2%)	47 75
5	f	127/574 (22%)	127 (100%)	0	100 100
6	i	107/235 (46%)	107 (100%)	0	100 100
7	j	83/154 (54%)	82 (99%)	1 (1%)	71 87
8	k	87/182 (48%)	87 (100%)	0	100 100
9	l	98/141 (70%)	97 (99%)	1 (1%)	76 89

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
10	m	80/106 (76%)	73 (91%)	7 (9%)	10 34
All	All	1469/3990 (37%)	1443 (98%)	26 (2%)	61 81

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	193	ASN
1	P	280	PHE
2	c	24	ASP
2	c	106	VAL
3	d	951	ASP
4	e	243	LEU
4	e	258	ASN
4	e	268	HIS
4	e	270	ASP
4	e	277	ASP
4	e	282	LEU
4	e	365	ARG
4	e	436	ASP
4	e	438	LEU
4	e	512	ASP
4	e	516	ILE
4	e	663	ARG
7	j	114	THR
9	l	114	LYS
10	m	31	LEU
10	m	55	ASP
10	m	58	GLU
10	m	77	ARG
10	m	84	GLU
10	m	90	ILE
10	m	91	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	P	189	ASN
1	P	193	ASN
3	d	912	ASN
3	d	943	GLN
3	d	1069	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	e	246	HIS
4	e	294	ASN
4	e	336	HIS
4	e	616	HIS
6	i	81	GLN
7	j	160	GLN
7	j	173	HIS
8	k	186	HIS
9	l	73	ASN
9	l	105	HIS
9	l	119	HIS
10	m	70	HIS
10	m	107	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

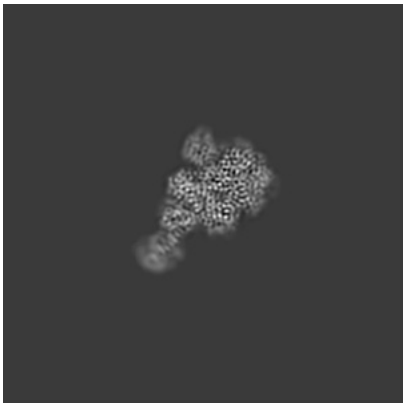
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-31115. These allow visual inspection of the internal detail of the map and identification of artifacts.

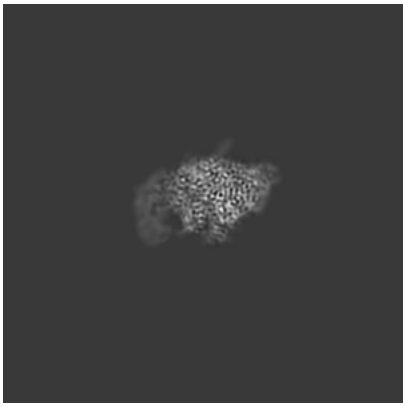
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

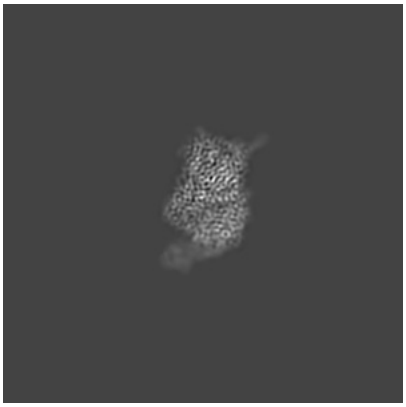
6.1.1 Primary map



X



Y



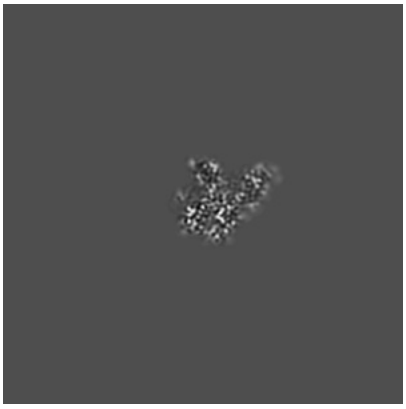
Z

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

6.2 Central slices

6.2.1 Primary map

X Index: 160



Y Index: 160

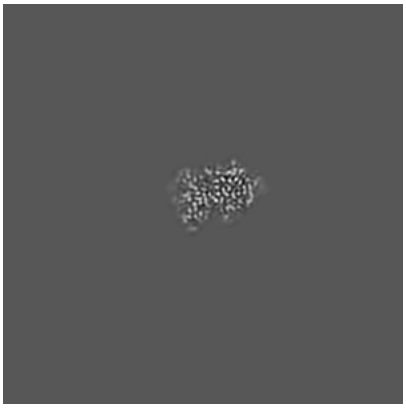
Z Index: 160

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: 178



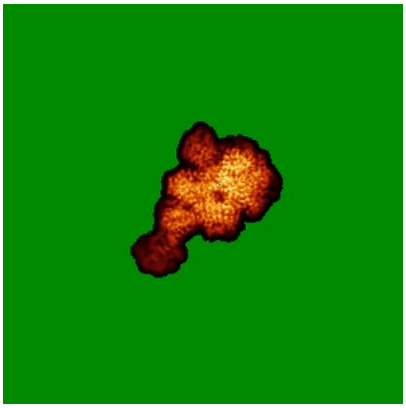
Y Index: 178

Z Index: 175

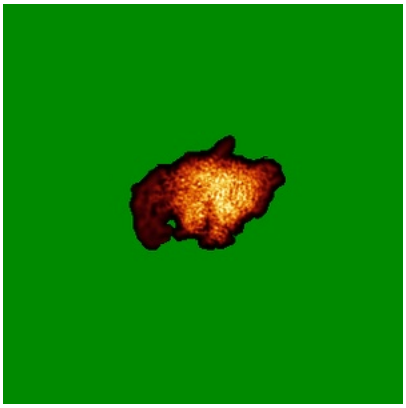
The images above show the largest variance slices of the map in three orthogonal directions.

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

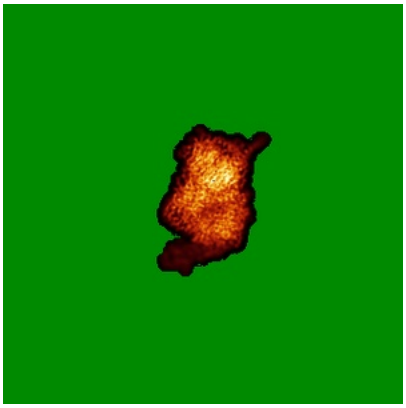
6.4.1 Primary map



X



Y

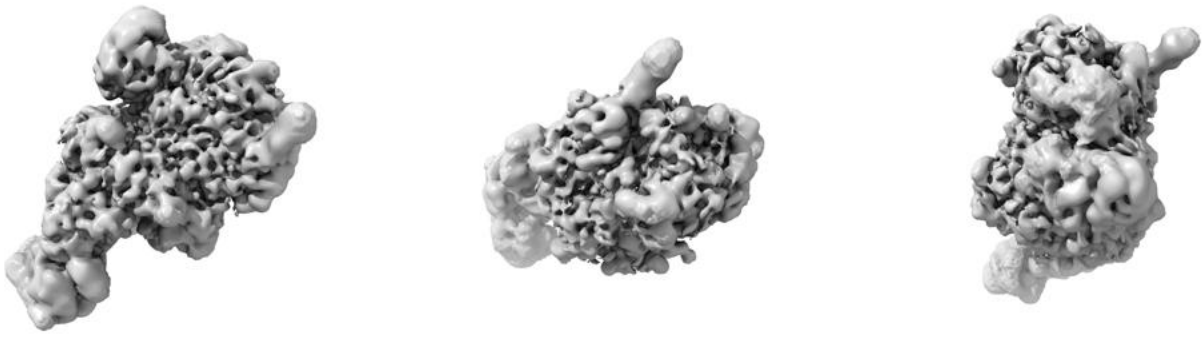


Z

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

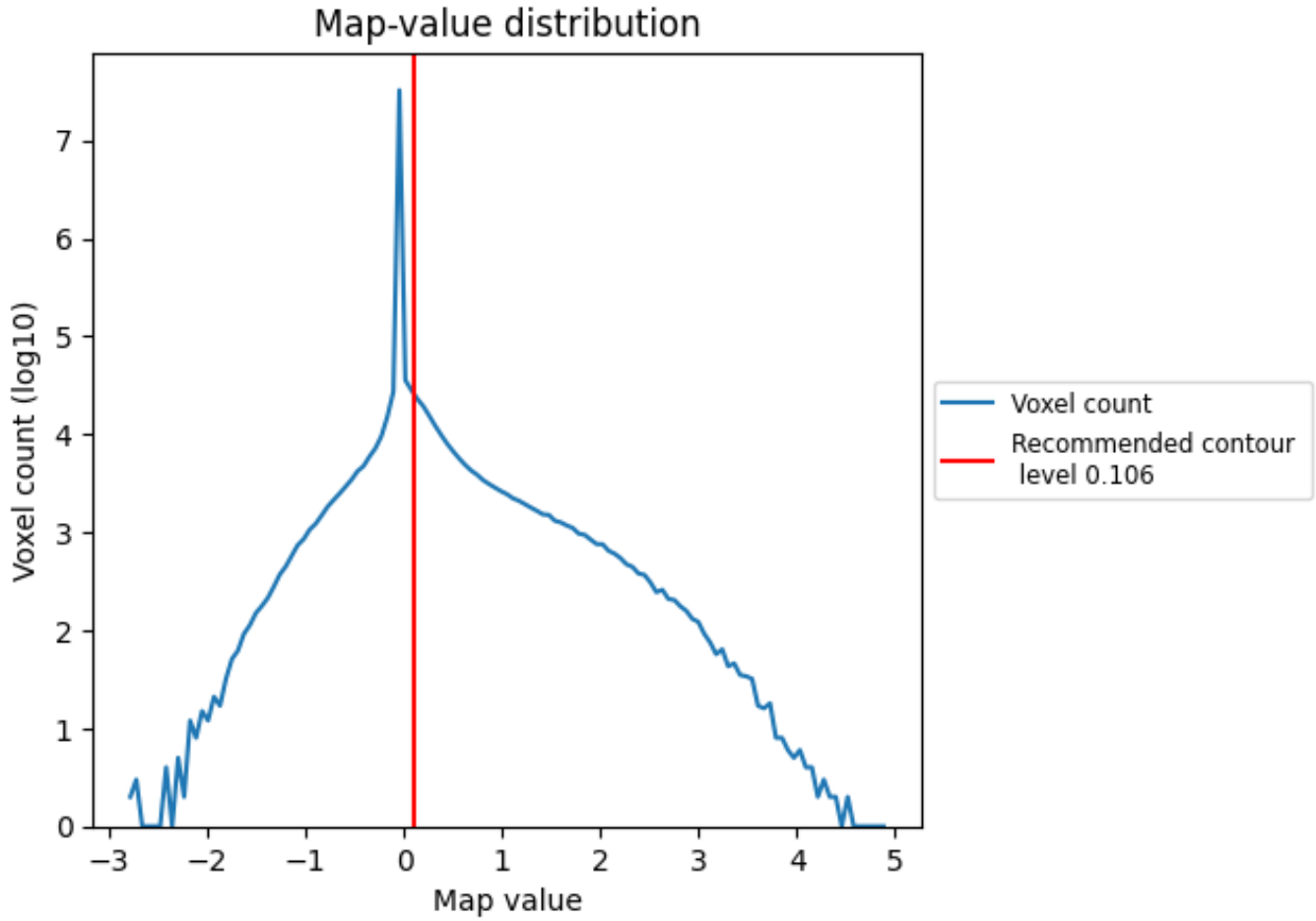
6.6 Mask visualisation (i)

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

7 Map analysis

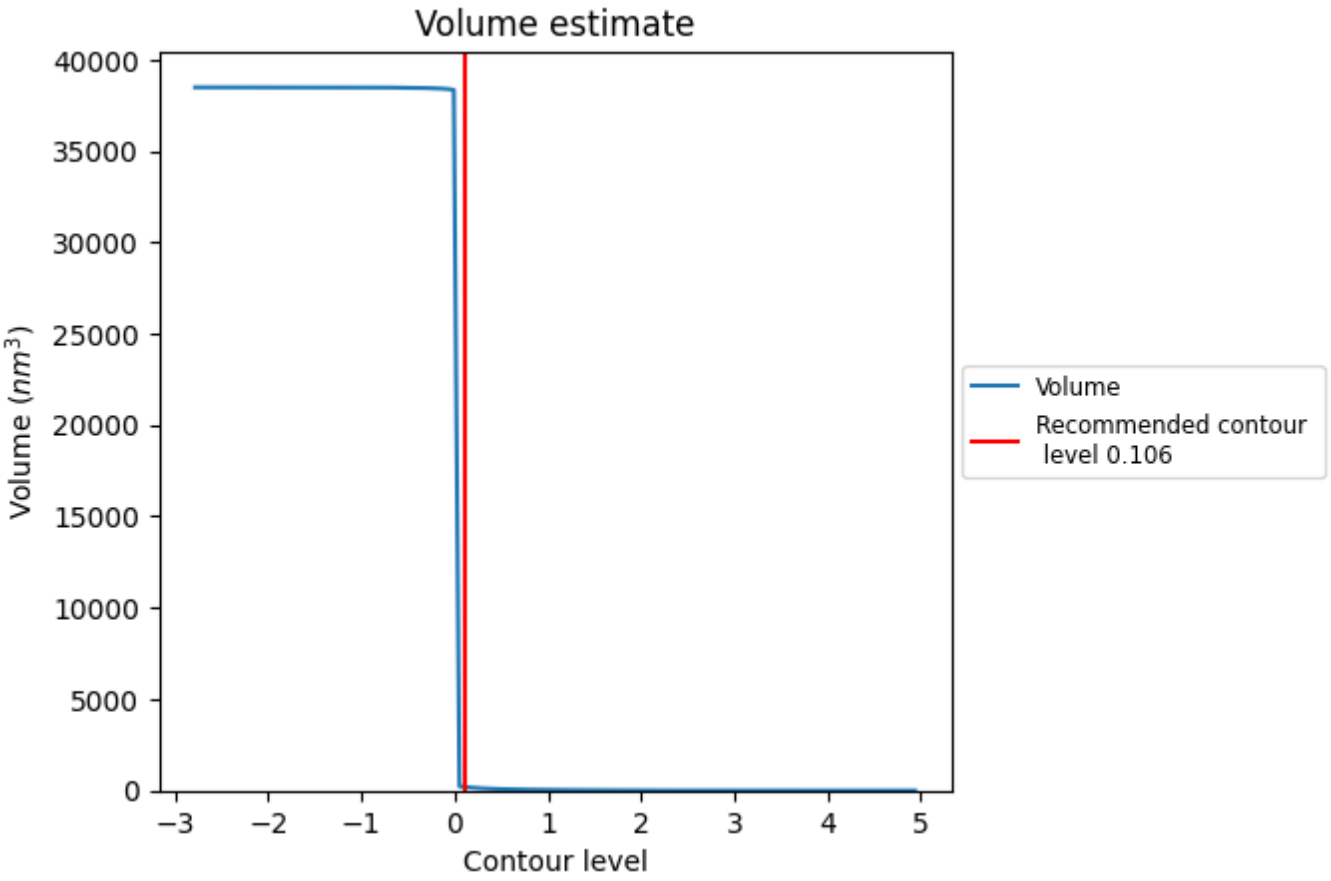
This section contains the results of statistical analysis of the map.

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.

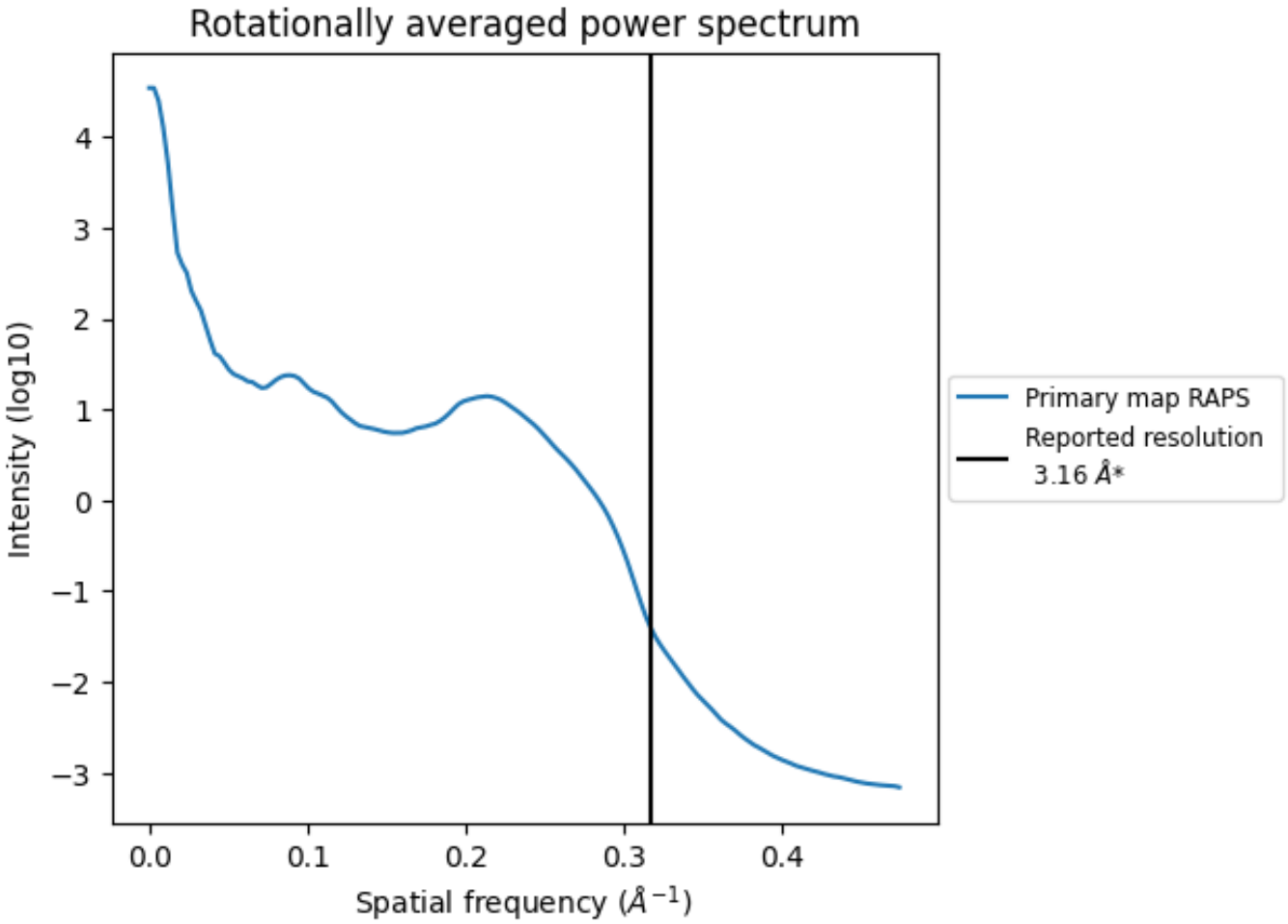
7.2 Volume estimate



The volume at the recommended contour level is 198 nm^3; this corresponds to an approximate mass of 179 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 0.316 Å^{-1}

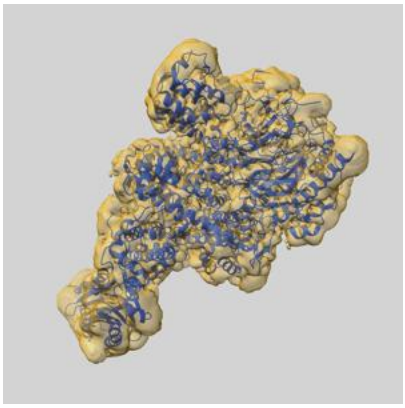
8 Fourier-Shell correlation

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

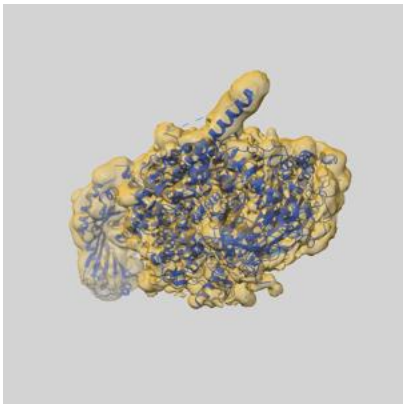
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-31115 and PDB model 7EGF. Per-residue inclusion information can be found in section 3 on page 6.

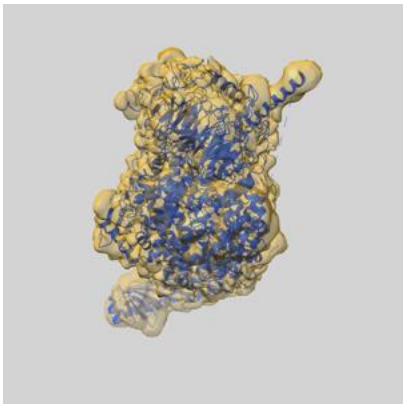
9.1 Map-model overlay



X



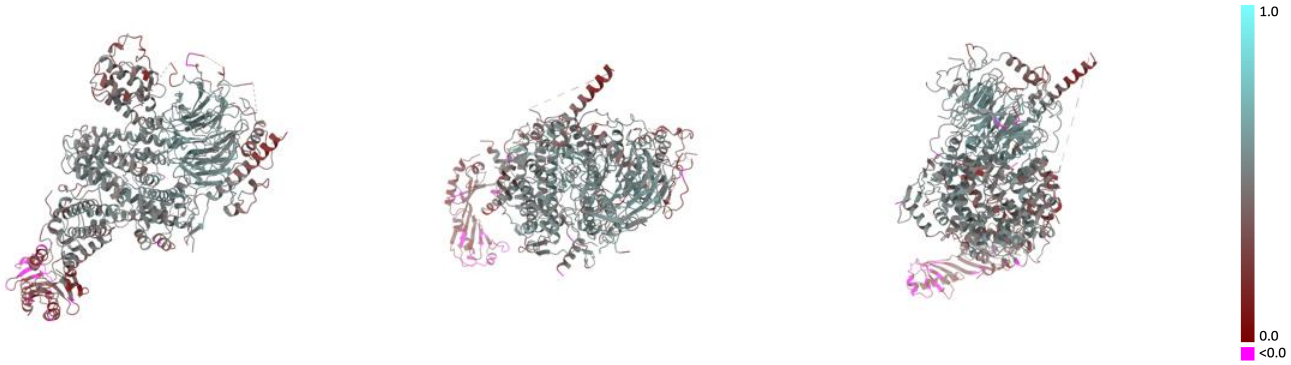
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.106 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.

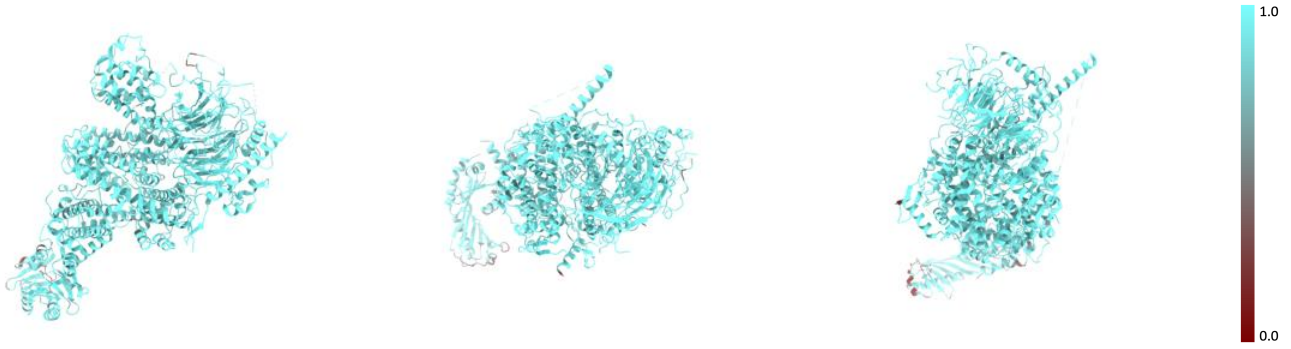
9.2 Q-score mapped to coordinate model (i)



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.

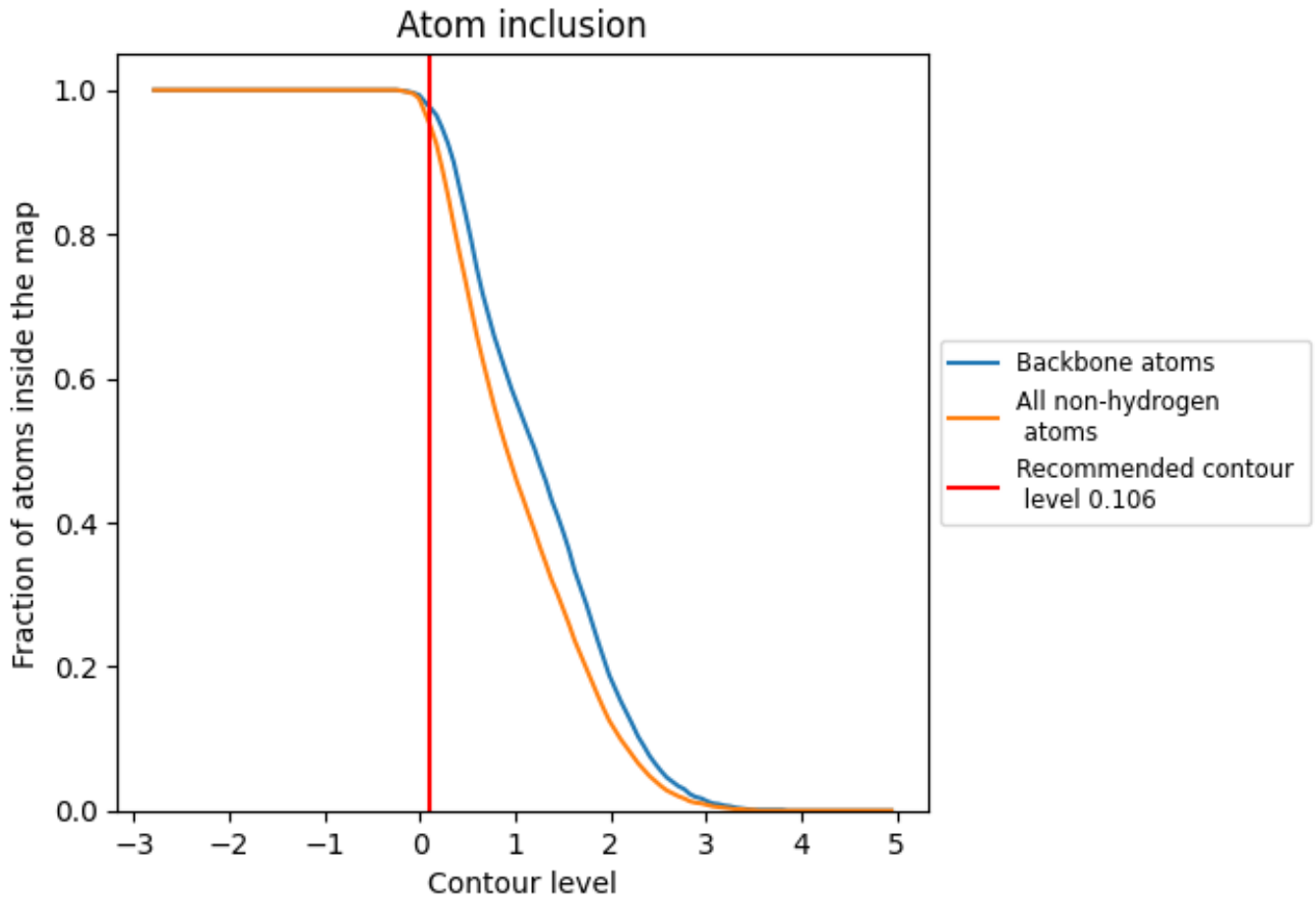
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.106).

9.4 Atom inclusion (i)



At the recommended contour level, 98% of all backbone atoms, 95% 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 (0.106) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9510	0.4430
P	0.8690	0.1710
c	0.9590	0.4920
d	0.9510	0.4330
e	0.9640	0.4810
f	0.9630	0.4980
i	0.9790	0.5090
j	0.9640	0.4890
k	0.9620	0.4670
l	0.9510	0.4560
m	0.9390	0.4330