wwPDB EM Validation Summary Report

Feb 12, 2024 – 07:07 PM EST

PDB ID	:	7TVE
EMDB ID	:	EMD-26140
Title	:	ATP and DNA bound SMC5/6 core complex
Authors	:	Yu, Y.; Patel, D.J.
Deposited on	:	2022-02-04
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:

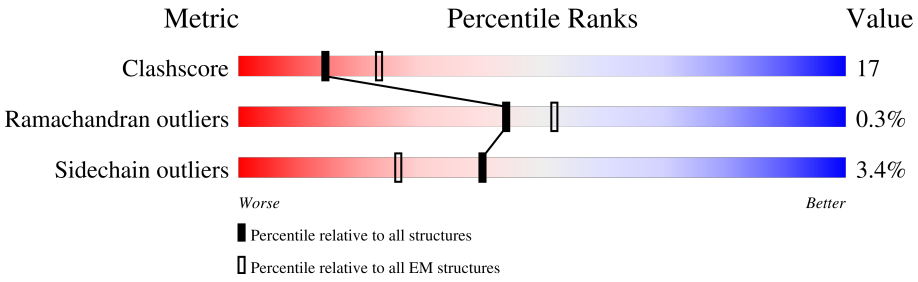
EMDB validation analysis	:	0.0.1.dev70
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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 <=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	68	9% 31% 60%
2	B	78	10% 9% 26% 65%
3	C	337	58% 33% 8%
4	D	1157	26% 12% 62%
5	E	1094	26% 15% 58%
6	F	305	57% 30% 11%
7	G	403	9% 51% 25% 24%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 15456 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 DNA chain called DNA (68-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	27	Total	C	N	O	P	0	0
			567	270	135	135	27		

- Molecule 2 is a DNA chain called DNA (78-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	27	Total	C	N	O	P	0	0
			540	270	54	189	27		

- Molecule 3 is a protein called Non-structural maintenance of chromosomes element 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	310	Total	C	N	O	S	0	0
			2489	1578	424	475	12		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	7	HIS	GLN	conflict	UNP A0A7I9FFW3
C	337	GLY	-	expression tag	UNP A0A7I9FFW3

- Molecule 4 is a protein called Structural maintenance of chromosomes protein 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	445	Total	C	N	O	S	0	0
			3512	2196	618	681	17		

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1048	GLN	GLU	engineered mutation	UNP Q12749
D	1115	GLY	-	expression tag	UNP Q12749

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1116	SER	-	expression tag	UNP Q12749
D	1117	LEU	-	expression tag	UNP Q12749
D	1118	GLU	-	expression tag	UNP Q12749
D	1119	VAL	-	expression tag	UNP Q12749
D	1120	LEU	-	expression tag	UNP Q12749
D	1121	PHE	-	expression tag	UNP Q12749
D	1122	GLN	-	expression tag	UNP Q12749
D	1123	GLY	-	expression tag	UNP Q12749
D	1124	PRO	-	expression tag	UNP Q12749
D	1125	GLY	-	expression tag	UNP Q12749
D	1126	GLY	-	expression tag	UNP Q12749
D	1127	SER	-	expression tag	UNP Q12749
D	1128	SER	-	expression tag	UNP Q12749
D	1129	ALA	-	expression tag	UNP Q12749
D	1130	TRP	-	expression tag	UNP Q12749
D	1131	SER	-	expression tag	UNP Q12749
D	1132	HIS	-	expression tag	UNP Q12749
D	1133	PRO	-	expression tag	UNP Q12749
D	1134	GLN	-	expression tag	UNP Q12749
D	1135	PHE	-	expression tag	UNP Q12749
D	1136	GLU	-	expression tag	UNP Q12749
D	1137	LYS	-	expression tag	UNP Q12749
D	1138	GLY	-	expression tag	UNP Q12749
D	1139	GLY	-	expression tag	UNP Q12749
D	1140	GLY	-	expression tag	UNP Q12749
D	1141	SER	-	expression tag	UNP Q12749
D	1142	GLY	-	expression tag	UNP Q12749
D	1143	GLY	-	expression tag	UNP Q12749
D	1144	GLY	-	expression tag	UNP Q12749
D	1145	SER	-	expression tag	UNP Q12749
D	1146	GLY	-	expression tag	UNP Q12749
D	1147	GLY	-	expression tag	UNP Q12749
D	1148	GLY	-	expression tag	UNP Q12749
D	1149	SER	-	expression tag	UNP Q12749
D	1150	TRP	-	expression tag	UNP Q12749
D	1151	SER	-	expression tag	UNP Q12749
D	1152	HIS	-	expression tag	UNP Q12749
D	1153	PRO	-	expression tag	UNP Q12749
D	1154	GLN	-	expression tag	UNP Q12749
D	1155	PHE	-	expression tag	UNP Q12749
D	1156	GLU	-	expression tag	UNP Q12749
D	1157	LYS	-	expression tag	UNP Q12749

- Molecule 5 is a protein called Structural maintenance of chromosomes protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	461	Total	C	N	O	S	0	0
			3661	2304	636	707	14		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1015	GLN	GLU	engineered mutation	UNP Q08204
E	1094	GLY	-	expression tag	UNP Q08204

- Molecule 6 is a protein called Non-structural maintenance of chromosome element 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	270	Total	C	N	O	S	0	0
			2150	1365	362	418	5		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	MET	-	initiating methionine	UNP Q05541
F	304	GLY	-	expression tag	UNP Q05541

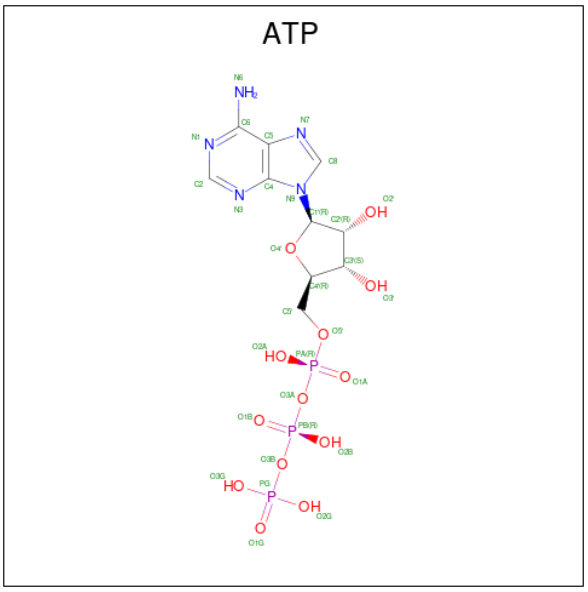
- Molecule 7 is a protein called Non-structural maintenance of chromosome element 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	307	Total	C	N	O	S	0	0
			2506	1579	458	458	11		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	403	GLY	-	expression tag	UNP P43124

- Molecule 8 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).

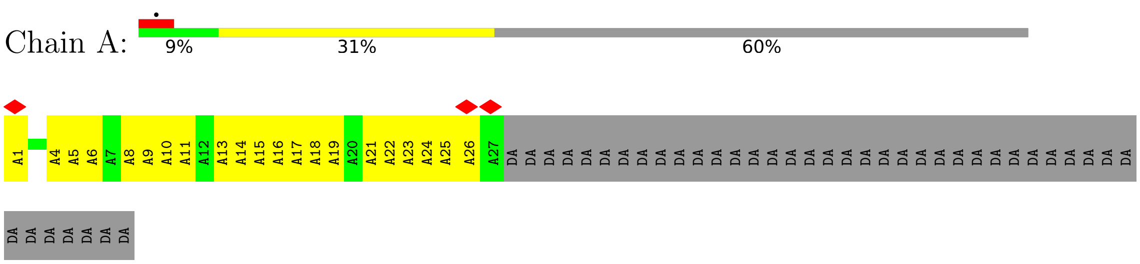


Mol	Chain	Residues	Atoms					AltConf
8	E	1	Total	C	N	O	P	0
			31	10	5	13	3	

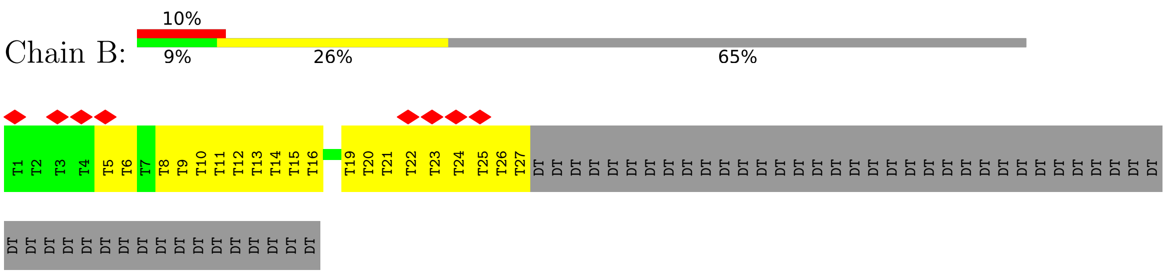
3 Residue-property plots (i)

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.

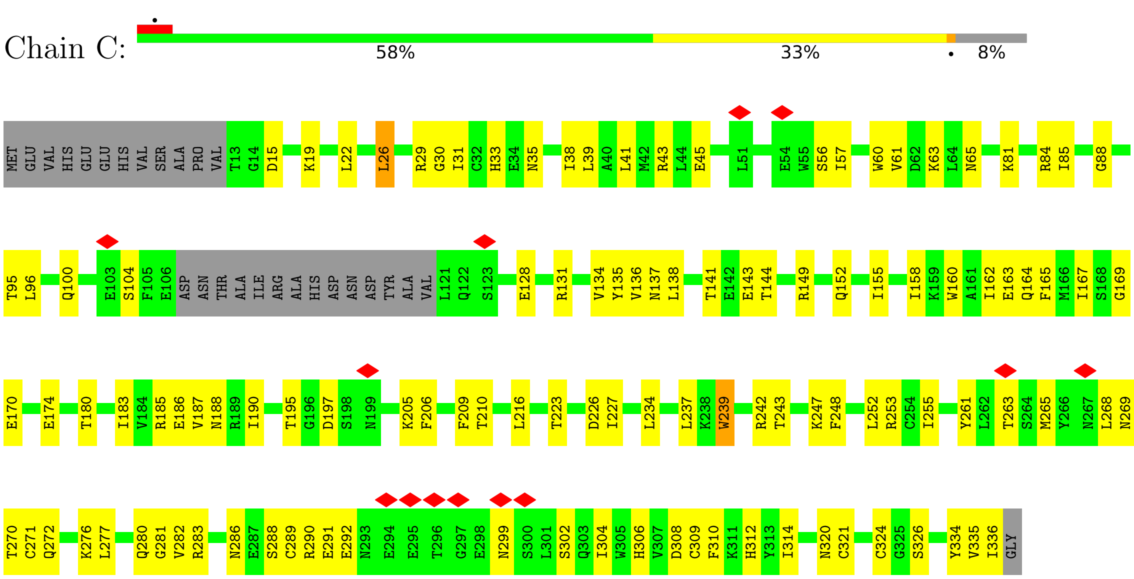
- Molecule 1: DNA (68-MER)



- Molecule 2: DNA (78-MER)

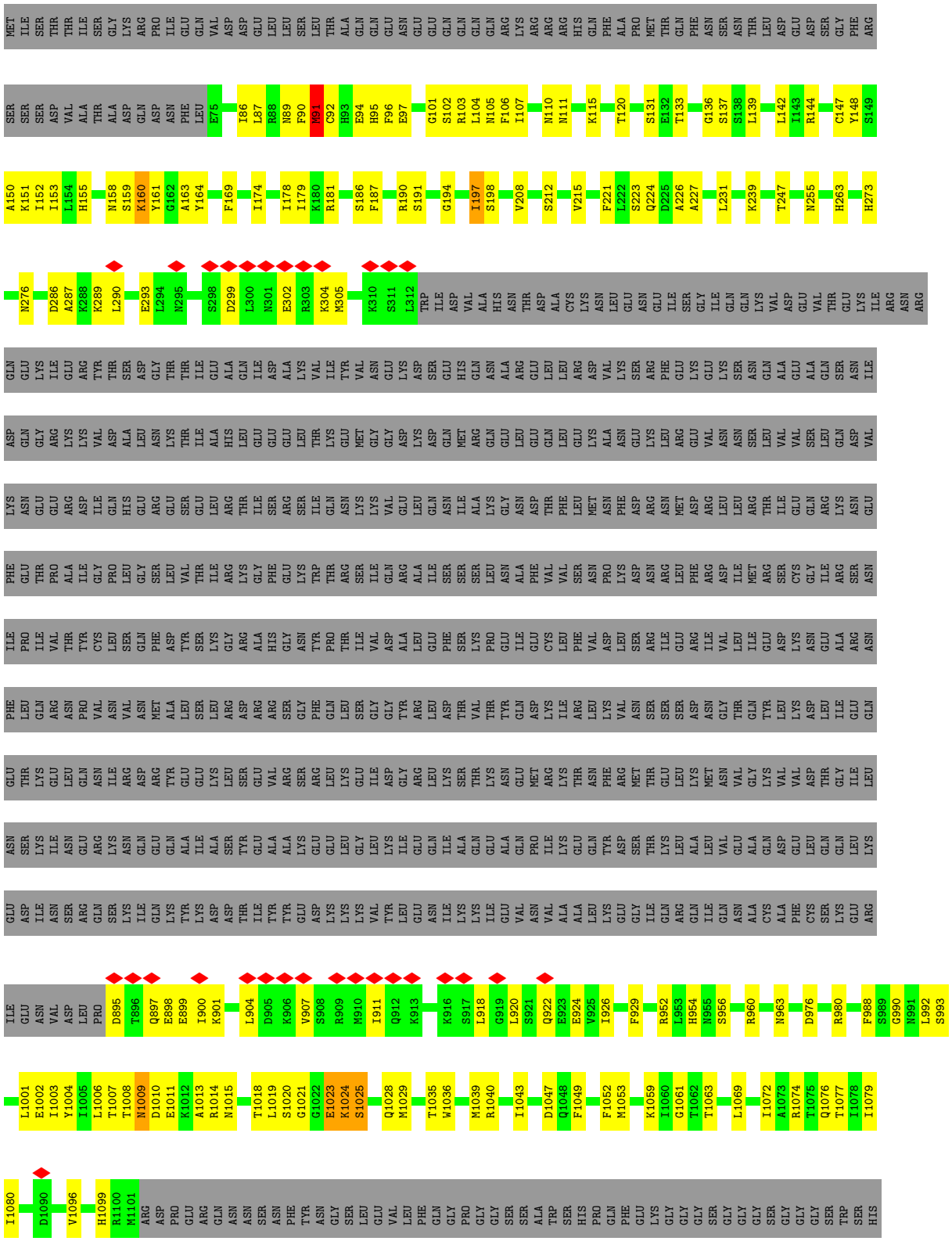


- Molecule 3: Non-structural maintenance of chromosomes element 1



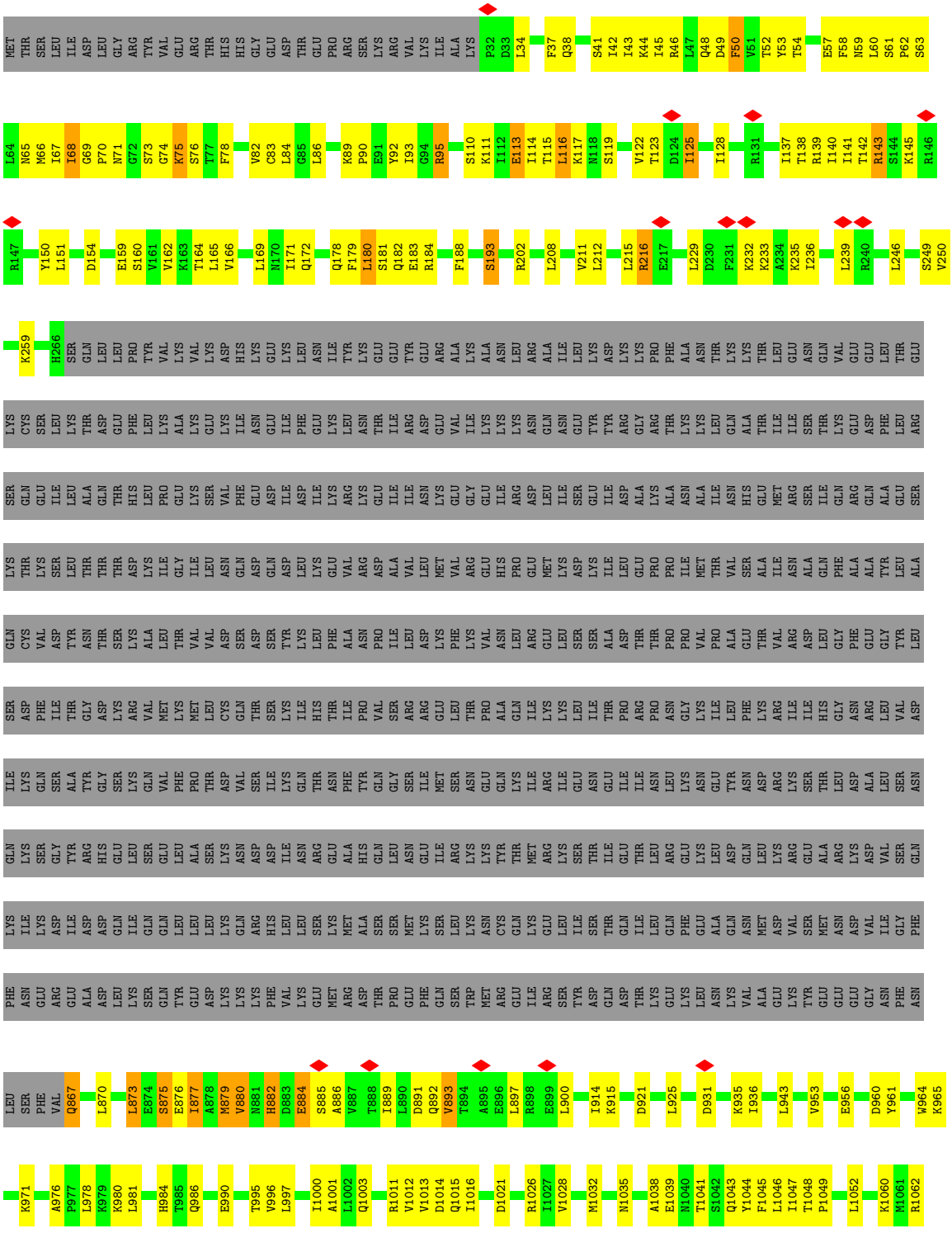
• Molecule 4: Structural maintenance of chromosomes protein 6

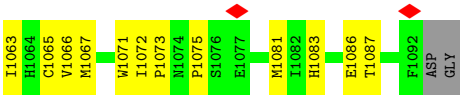
Chain D:



• Molecule 5: Structural maintenance of chromosomes protein 5

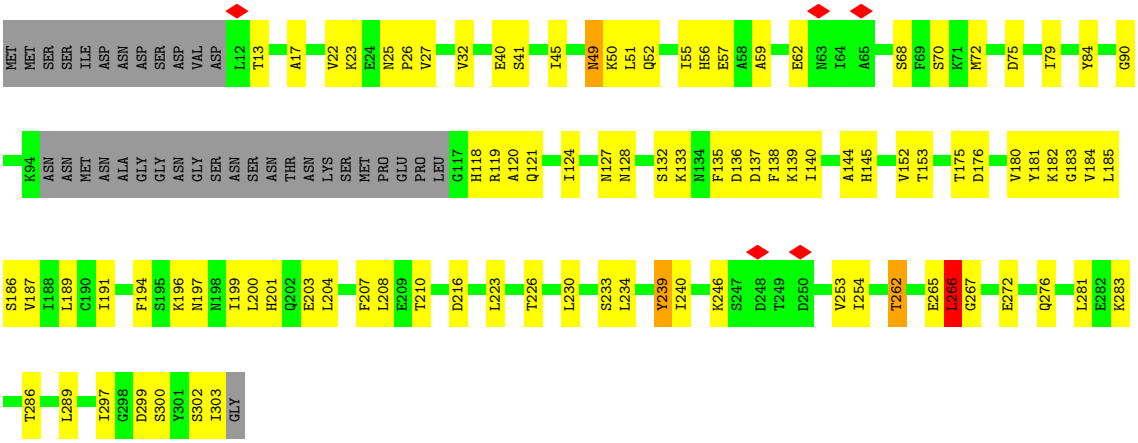
Chain E:





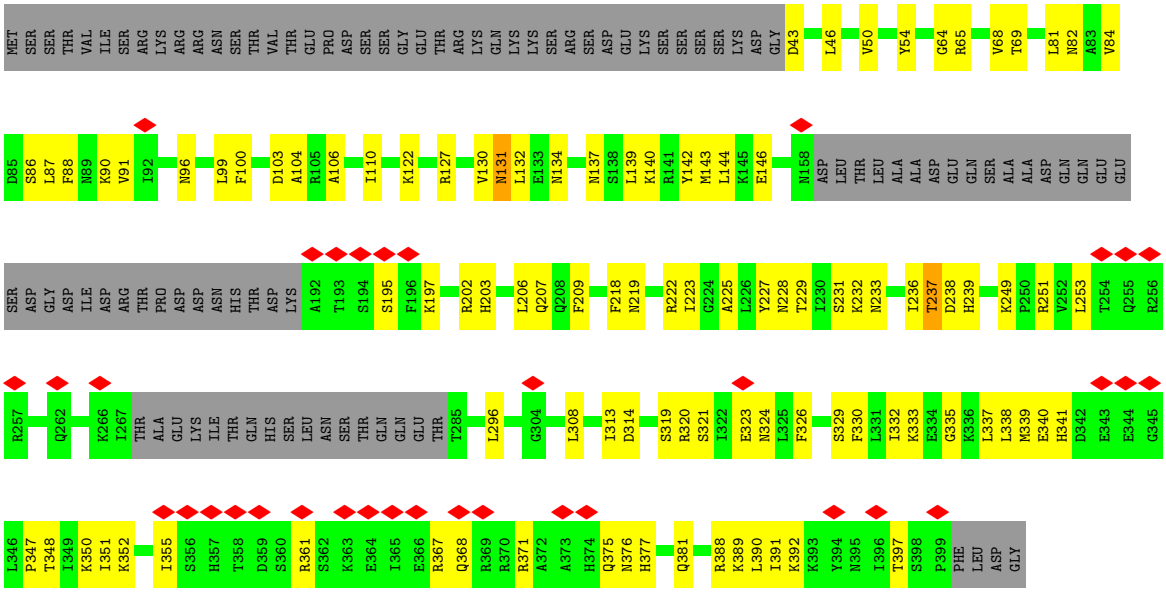
• Molecule 6: Non-structural maintenance of chromosome element 3

Chain F:



• Molecule 7: Non-structural maintenance of chromosome element 4

Chain G:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	201249	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 ($e^-/Å^2$)	53.55	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.155	Depositor
Minimum map value	-0.086	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.012	Depositor
Map size (Å)	320.7, 320.7, 320.7	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.069, 1.069, 1.069	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ATP

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	0.18	0/647	0.59	0/995
2	B	0.18	0/593	0.62	0/914
3	C	0.32	0/2531	0.52	0/3425
4	D	0.28	0/3554	0.46	0/4759
5	E	0.28	0/3716	0.47	0/5009
6	F	0.34	0/2178	0.53	0/2929
7	G	0.30	0/2551	0.49	0/3419
All	All	0.29	0/15770	0.50	0/21450

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	A	567	0	298	29	0
2	B	540	0	325	38	0
3	C	2489	0	2459	88	0
4	D	3512	0	3565	113	0
5	E	3661	0	3695	174	0
6	F	2150	0	2198	70	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	2506	0	2535	104	0
8	E	31	0	12	0	0
All	All	15456	0	15087	527	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 17.

The worst 5 of 527 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:A:8:DA:H2	2:B:21:DT:N3	1.48	1.12
1:A:18:DA:H2	2:B:11:DT:O2	1.52	0.91
5:E:142:THR:HG21	5:E:145:LYS:HB2	1.55	0.87
5:E:142:THR:CG2	5:E:145:LYS:HB2	2.04	0.86
5:E:193:SER:HB3	5:E:965:LYS:HB2	1.58	0.85

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
3	C	306/337 (91%)	269 (88%)	36 (12%)	1 (0%)	41 74
4	D	441/1157 (38%)	388 (88%)	50 (11%)	3 (1%)	22 60
5	E	457/1094 (42%)	390 (85%)	67 (15%)	0	100 100
6	F	266/305 (87%)	226 (85%)	39 (15%)	1 (0%)	34 70
7	G	301/403 (75%)	255 (85%)	46 (15%)	0	100 100
All	All	1771/3296 (54%)	1528 (86%)	238 (13%)	5 (0%)	44 74

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	F	266	LEU
4	D	91	MET
4	D	1021	GLY
3	C	252	LEU
4	D	197	ILE

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	
3	C	273/297 (92%)	271 (99%)	2 (1%)	84	91
4	D	387/1032 (38%)	374 (97%)	13 (3%)	37	64
5	E	412/1003 (41%)	385 (93%)	27 (7%)	16	48
6	F	243/273 (89%)	235 (97%)	8 (3%)	38	65
7	G	280/367 (76%)	276 (99%)	4 (1%)	67	81
All	All	1595/2972 (54%)	1541 (97%)	54 (3%)	40	64

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	184	ARG
5	E	877	ILE
6	F	266	LEU
5	E	193	SER
5	E	870	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
5	E	882	HIS
7	G	260	ASN
5	E	1035	ASN
7	G	353	GLN
7	G	134	ASN

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)

1 ligand is 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	$\#\|Z\| > 2$
8	ATP	E	1101	-	26,33,33	0.93	1 (3%)	31,52,52	1.61	5 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	ATP	E	1101	-	-	7/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	E	1101	ATP	C5-C4	2.48	1.47	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	E	1101	ATP	PA-O3A-PB	-3.63	120.36	132.83
8	E	1101	ATP	PB-O3B-PG	-3.61	120.43	132.83
8	E	1101	ATP	C3'-C2'-C1'	3.41	106.12	100.98
8	E	1101	ATP	N3-C2-N1	-3.16	123.73	128.68
8	E	1101	ATP	C4-C5-N7	-2.70	106.59	109.40

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

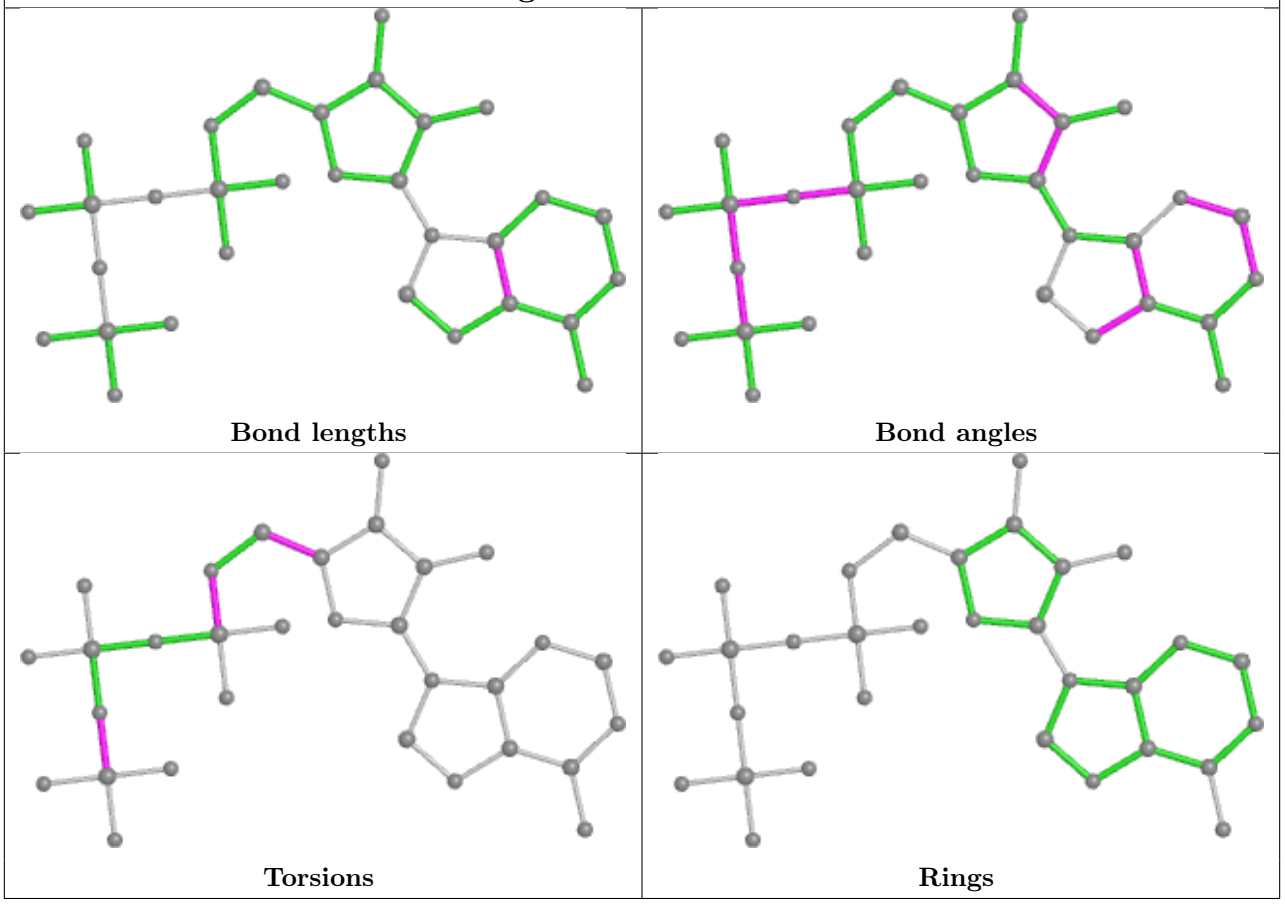
Mol	Chain	Res	Type	Atoms
8	E	1101	ATP	PB-O3B-PG-O2G
8	E	1101	ATP	C5'-O5'-PA-O1A
8	E	1101	ATP	O4'-C4'-C5'-O5'
8	E	1101	ATP	C3'-C4'-C5'-O5'
8	E	1101	ATP	C5'-O5'-PA-O3A

There are no ring outliers.

No monomer is involved in short contacts.

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

Ligand ATP E 1101



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

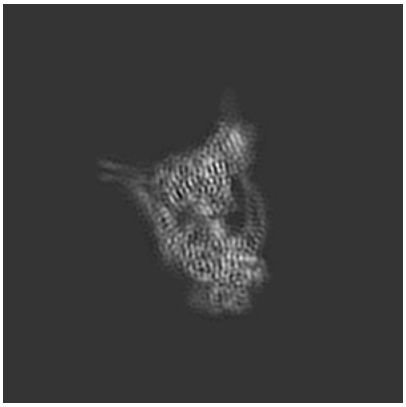
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-26140. These allow visual inspection of the internal detail of the map and identification of artifacts.

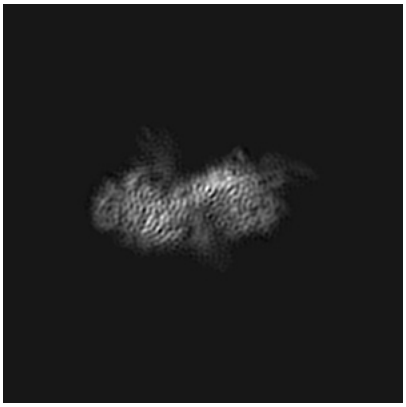
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

6.1.1 Primary map



X



Y

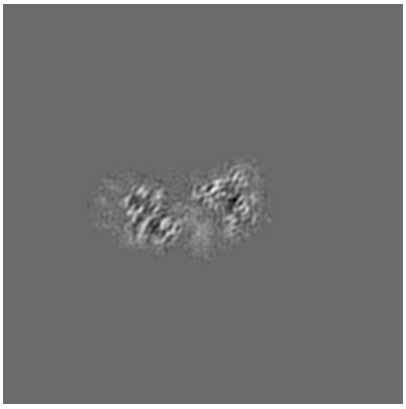
Z

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

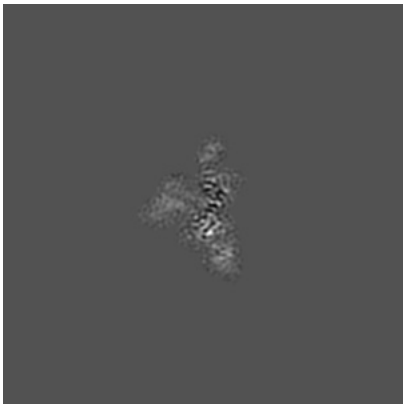
6.2 Central slices

6.2.1 Primary map

X Index: 150



Y Index: 150



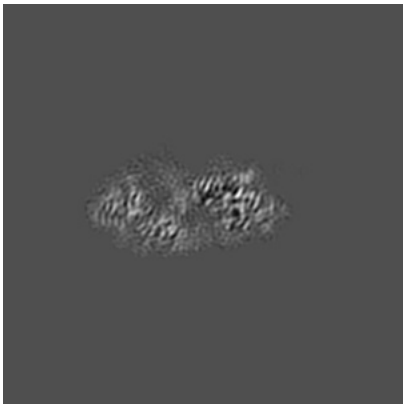
Z Index: 150

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: 150



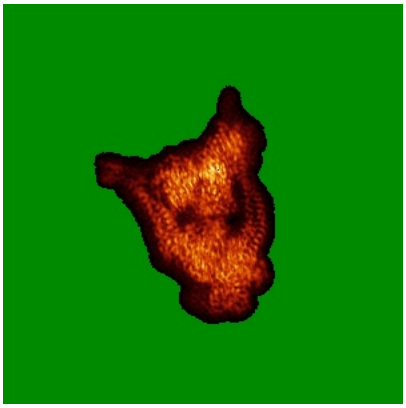
Y Index: 160

Z Index: 172

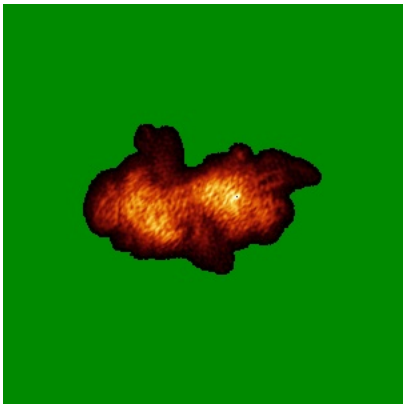
The images above show the largest variance slices of the map in three orthogonal directions.

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

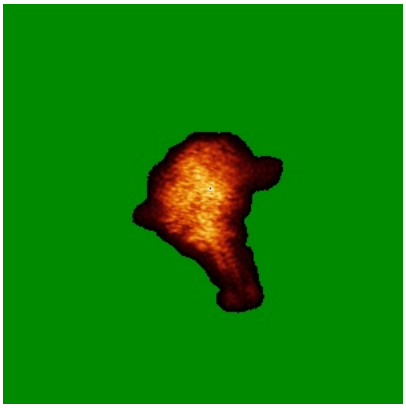
6.4.1 Primary map



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

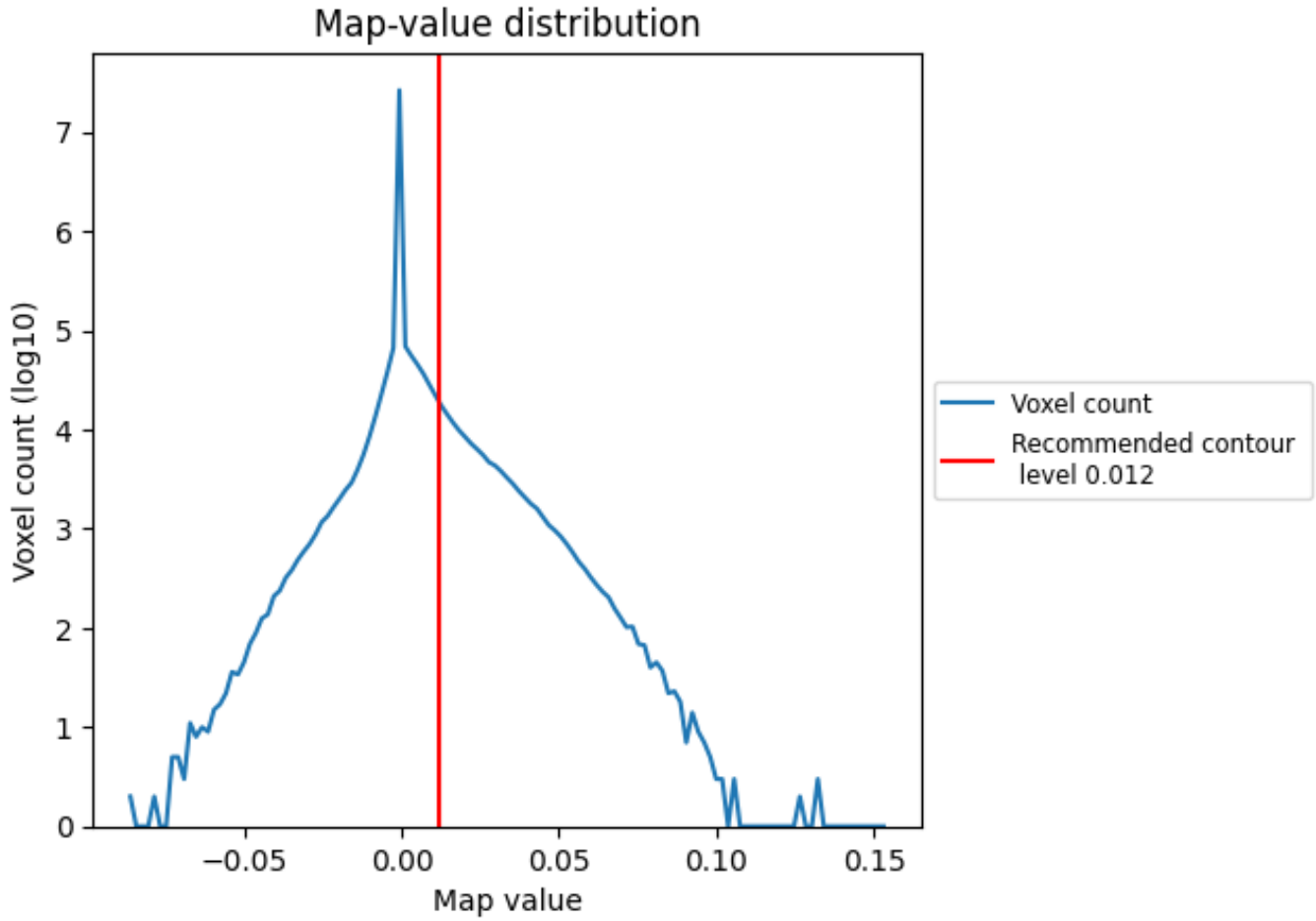
6.6 Mask visualisation (i)

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

7 Map analysis

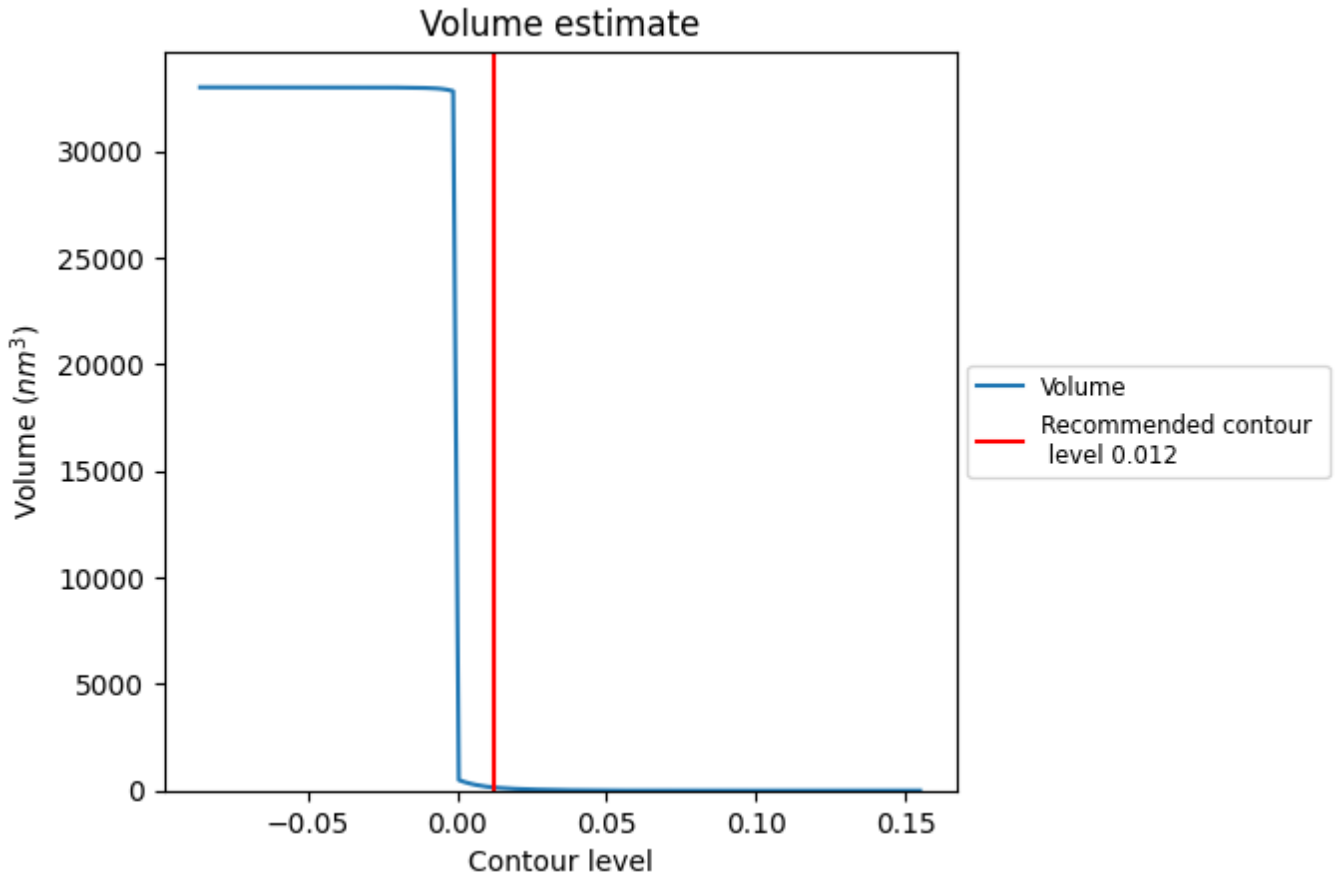
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

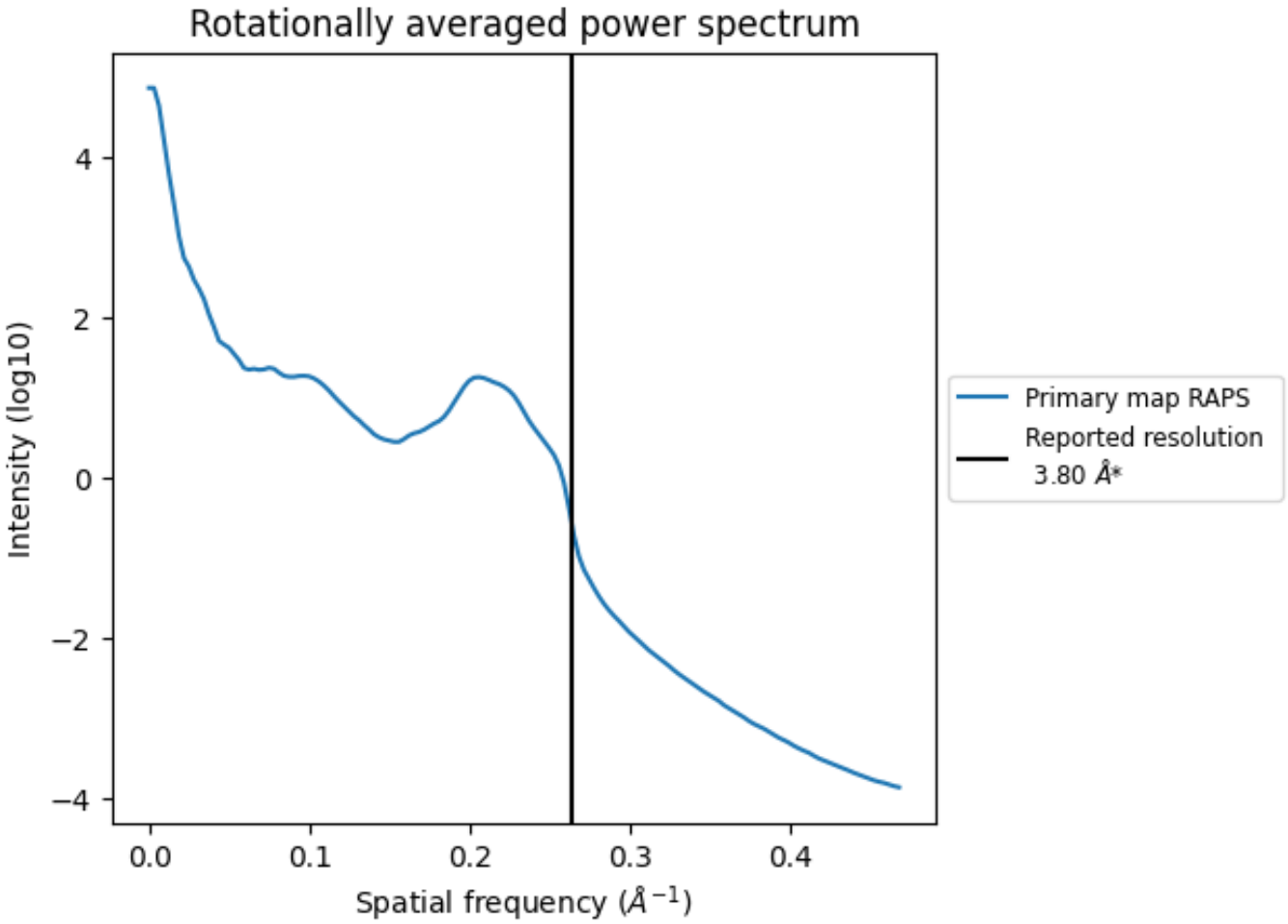
7.2 Volume estimate (i)



The volume at the recommended contour level is 150 nm^3; this corresponds to an approximate mass of 135 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}

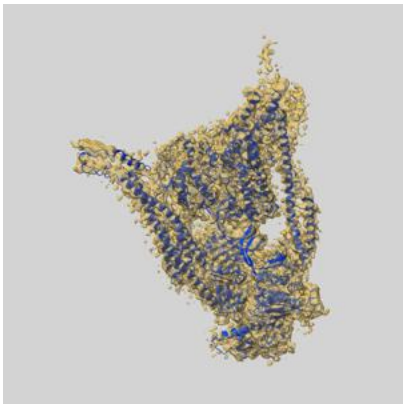
8 Fourier-Shell correlation

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

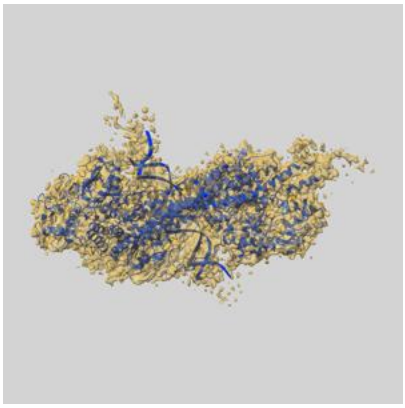
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-26140 and PDB model 7TVE. Per-residue inclusion information can be found in section 3 on page 7.

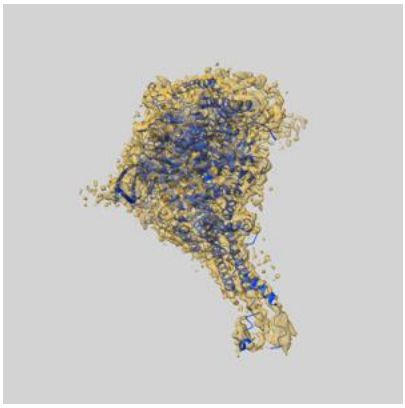
9.1 Map-model overlay



X



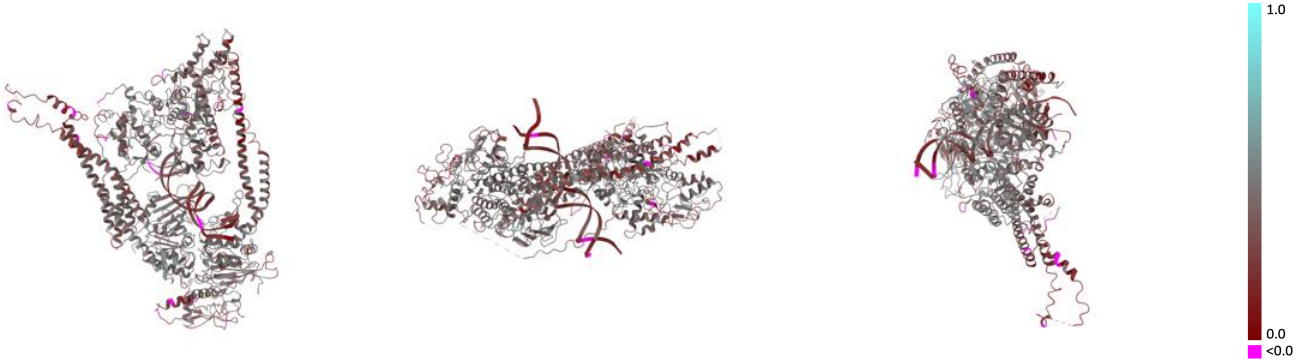
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.012 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.

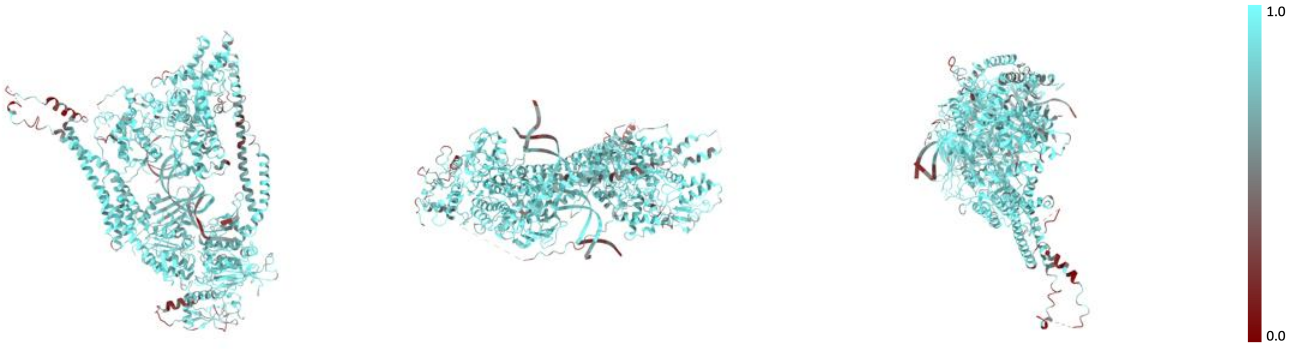
9.2 Q-score mapped to coordinate model (i)



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.

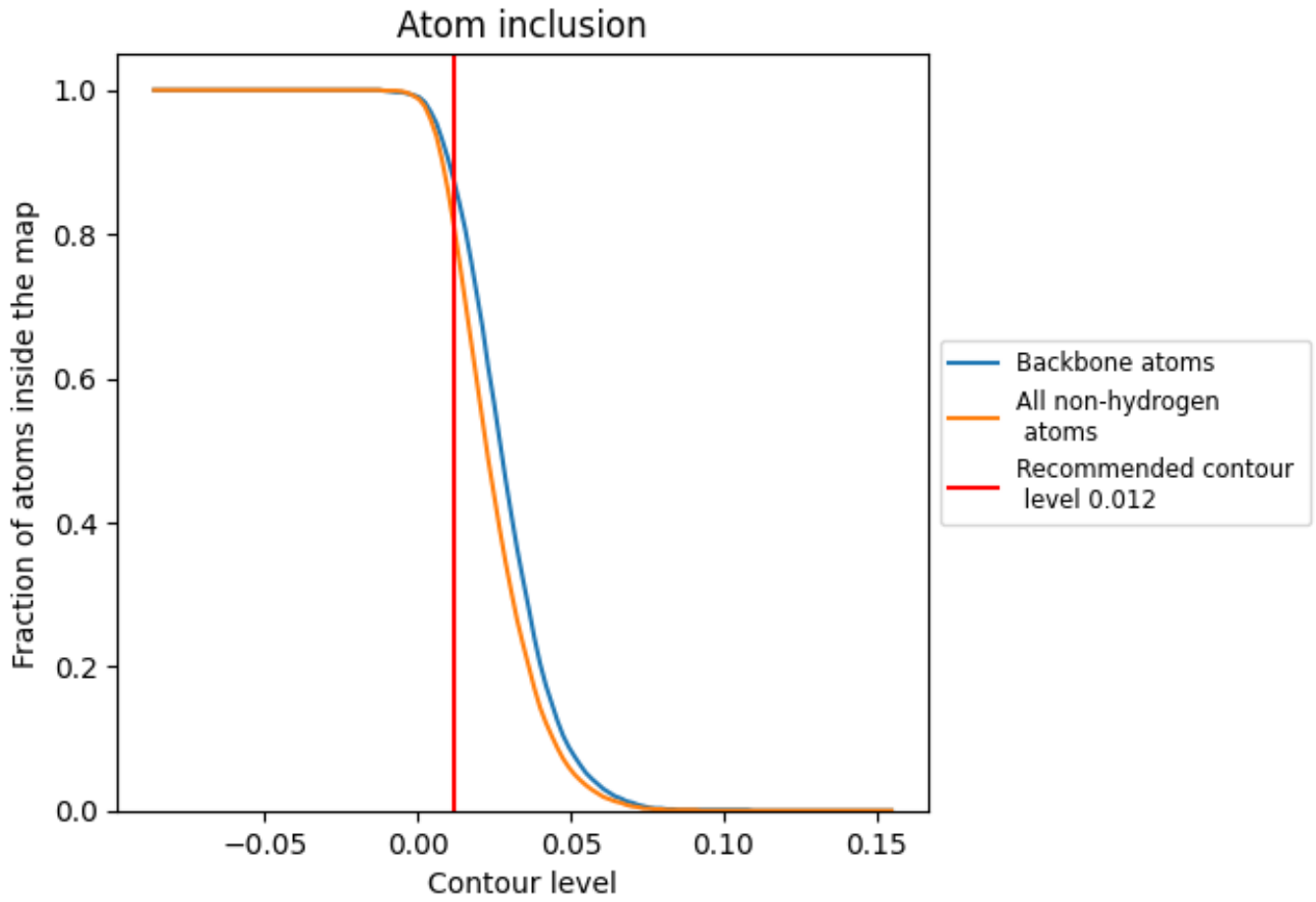
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.012).

9.4 Atom inclusion (i)



At the recommended contour level, 87% of all backbone atoms, 81% 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 (0.012) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8090	0.3780
A	0.6970	0.2210
B	0.6090	0.1740
C	0.8510	0.3940
D	0.8270	0.3890
E	0.8060	0.3850
F	0.8640	0.4280
G	0.7680	0.3740