wwPDB EM Validation Summary Report (i)

Feb 2, 2023 – 02:17 PM JST

PDB ID	:	7F0R
EMDB ID	:	EMD-31403
Title	:	Cryo-EM structure of Pseudomonas aeruginosa SutA transcription activation complex
Authors	:	He, D.W.; You, L.L.; Zhang, Y.
Deposited on	:	2021-06-06
Resolution	:	5.80 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

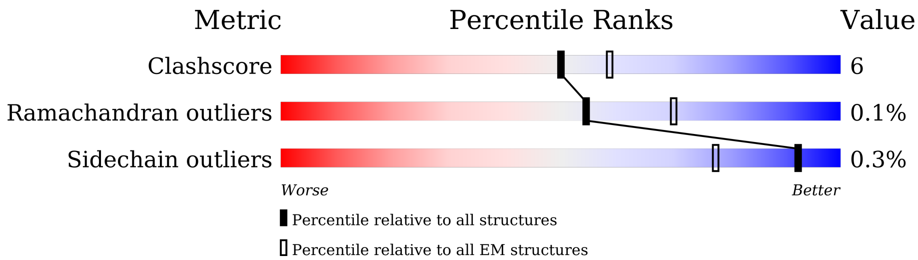
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

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 5.80 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	345	• 58% 5% 38%
1	B	345	6% 55% 6% 39%
2	C	1359	12% 85% 12% •
3	D	1412	12% 81% 14% 5%
4	E	88	6% 74% • 24%
5	G	109	• 30% • 69%
6	F	342	6% 73% 9% 19%
7	I	70	50% • 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	70	47% 10% 43%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 25722 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 alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	215	1554	980	285	286	3	0	0
1	B	209	1500	944	269	284	3	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	initiating methionine	UNP O52760
A	-10	GLY	-	expression tag	UNP O52760
A	-9	HIS	-	expression tag	UNP O52760
A	-8	HIS	-	expression tag	UNP O52760
A	-7	HIS	-	expression tag	UNP O52760
A	-6	HIS	-	expression tag	UNP O52760
A	-5	HIS	-	expression tag	UNP O52760
A	-4	HIS	-	expression tag	UNP O52760
A	-3	HIS	-	expression tag	UNP O52760
A	-2	HIS	-	expression tag	UNP O52760
A	-1	HIS	-	expression tag	UNP O52760
A	0	HIS	-	expression tag	UNP O52760
B	-11	MET	-	initiating methionine	UNP O52760
B	-10	GLY	-	expression tag	UNP O52760
B	-9	HIS	-	expression tag	UNP O52760
B	-8	HIS	-	expression tag	UNP O52760
B	-7	HIS	-	expression tag	UNP O52760
B	-6	HIS	-	expression tag	UNP O52760
B	-5	HIS	-	expression tag	UNP O52760
B	-4	HIS	-	expression tag	UNP O52760
B	-3	HIS	-	expression tag	UNP O52760
B	-2	HIS	-	expression tag	UNP O52760
B	-1	HIS	-	expression tag	UNP O52760
B	0	HIS	-	expression tag	UNP O52760

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	1324	9188	5730	1679	1747	32	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	MET	-	initiating methionine	UNP Q51561
C	0	GLY	-	expression tag	UNP Q51561

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	1339	9324	5890	1693	1704	37	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP Q9HWC9
D	1	LEU	-	expression tag	UNP Q9HWC9
D	1400	GLY	-	expression tag	UNP Q9HWC9
D	1401	SER	-	expression tag	UNP Q9HWC9
D	1402	GLY	-	expression tag	UNP Q9HWC9
D	1403	SER	-	expression tag	UNP Q9HWC9
D	1404	TRP	-	expression tag	UNP Q9HWC9
D	1405	SER	-	expression tag	UNP Q9HWC9
D	1406	HIS	-	expression tag	UNP Q9HWC9
D	1407	PRO	-	expression tag	UNP Q9HWC9
D	1408	GLN	-	expression tag	UNP Q9HWC9
D	1409	PHE	-	expression tag	UNP Q9HWC9
D	1410	GLU	-	expression tag	UNP Q9HWC9
D	1411	LYS	-	expression tag	UNP Q9HWC9

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	67	460	289	87	83	1	0	0

- Molecule 5 is a protein called Transcriptional factor SutA.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	G	34	181	106	38	37	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	GLY	-	expression tag	UNP Q9HTR9
G	-2	ALA	-	expression tag	UNP Q9HTR9
G	-1	MET	-	expression tag	UNP Q9HTR9
G	0	GLY	-	expression tag	UNP Q9HTR9

- Molecule 6 is a protein called RNA polymerase sigma factor RpoS.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	278	1911	1194	358	356	3	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	335	LEU	-	expression tag	UNP P45684
F	336	GLU	-	expression tag	UNP P45684
F	337	HIS	-	expression tag	UNP P45684
F	338	HIS	-	expression tag	UNP P45684
F	339	HIS	-	expression tag	UNP P45684
F	340	HIS	-	expression tag	UNP P45684
F	341	HIS	-	expression tag	UNP P45684
F	342	HIS	-	expression tag	UNP P45684

- Molecule 7 is a DNA chain called DNA (70-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	I	38	777	369	141	229	38	0	0

- Molecule 8 is a DNA chain called DNA (70-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	H	40	824	390	156	238	40	0	0

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	AltConf
9	D	2	Total Zn 2 2	0

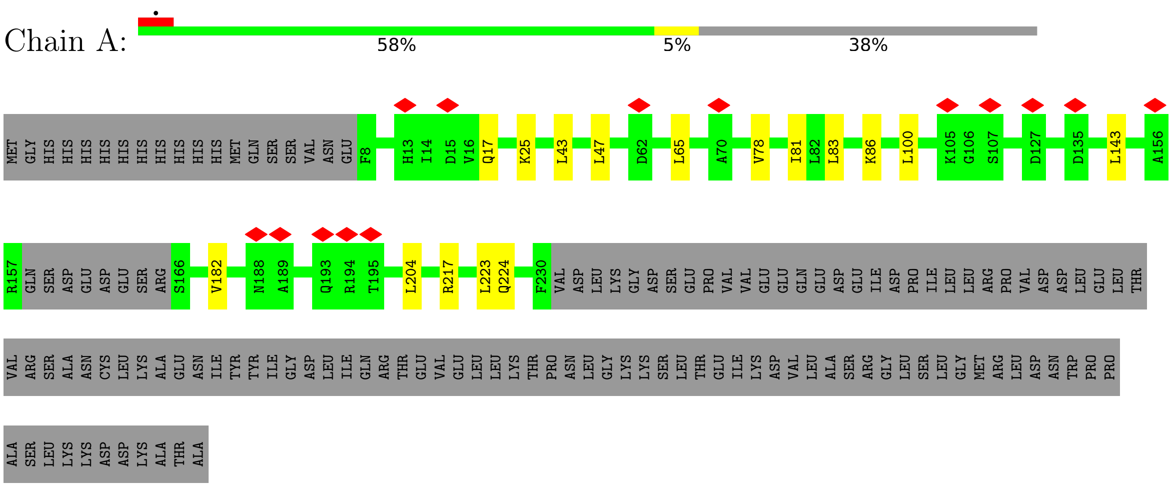
- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	AltConf
10	D	1	Total Mg 1 1	0

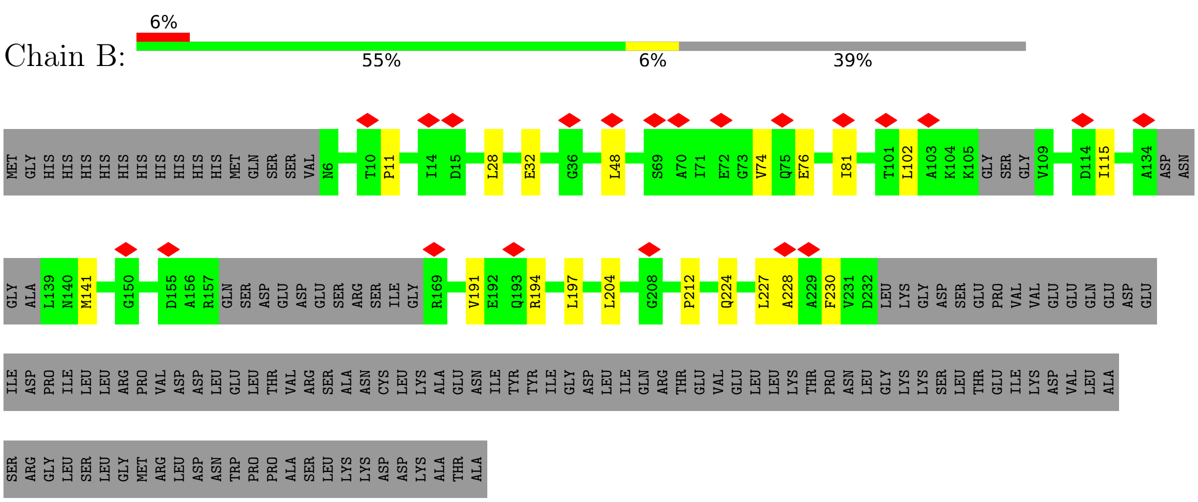
3 Residue-property plots

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.

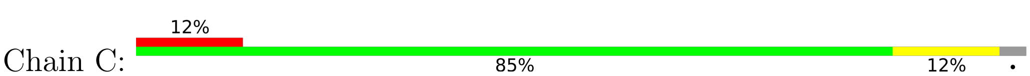
- Molecule 1: DNA-directed RNA polymerase subunit alpha

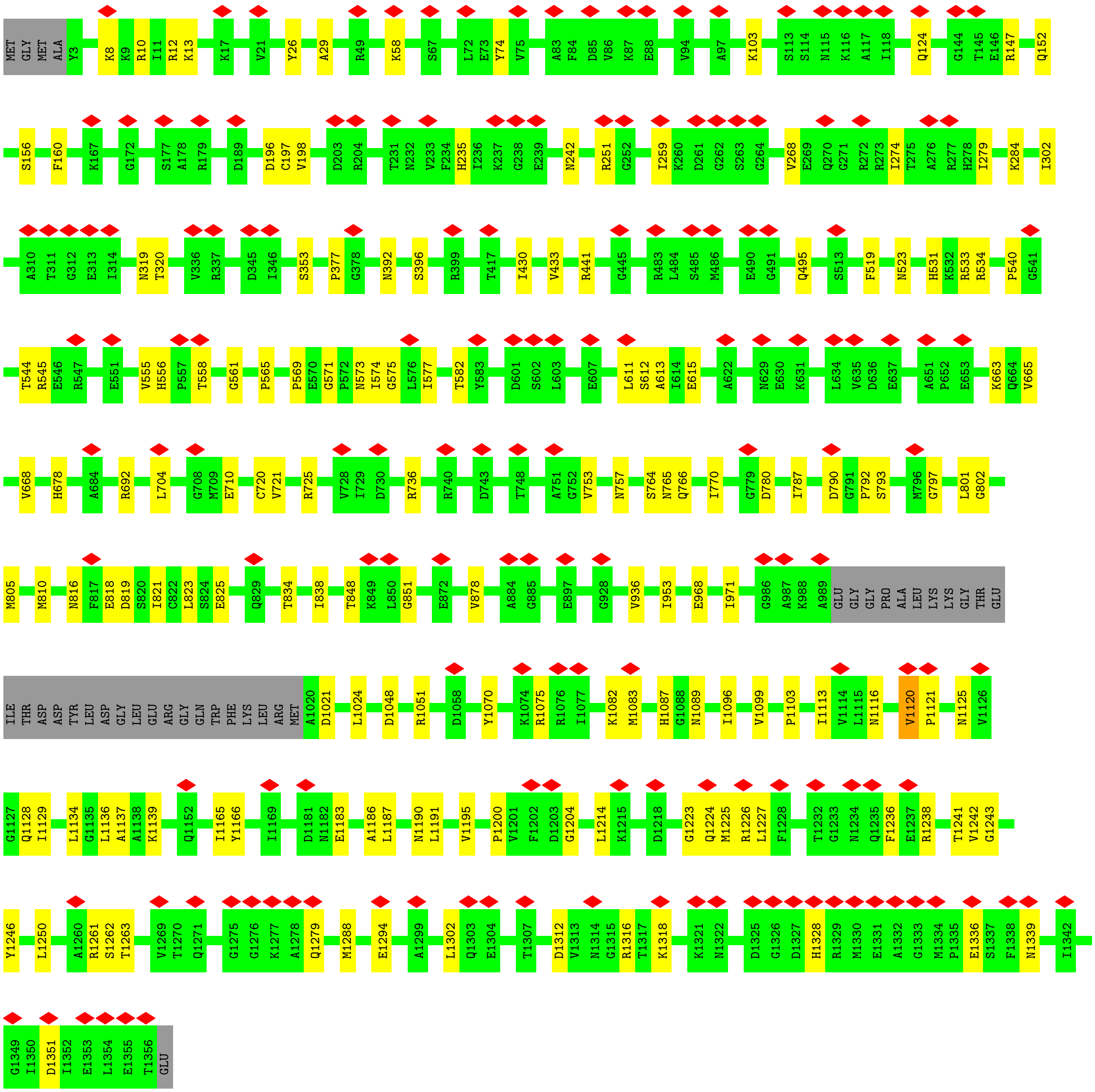


- Molecule 1: DNA-directed RNA polymerase subunit alpha



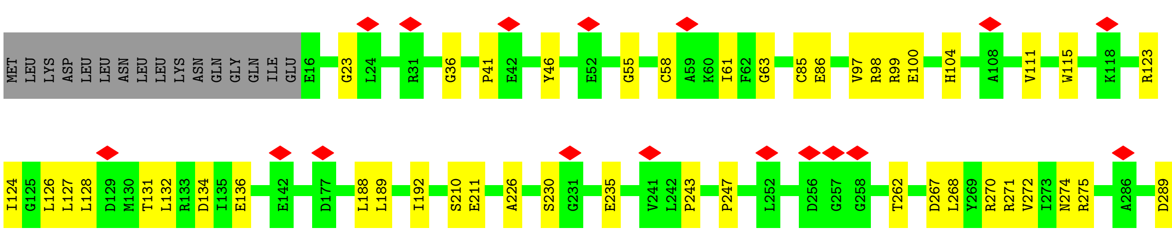
- Molecule 2: DNA-directed RNA polymerase subunit beta

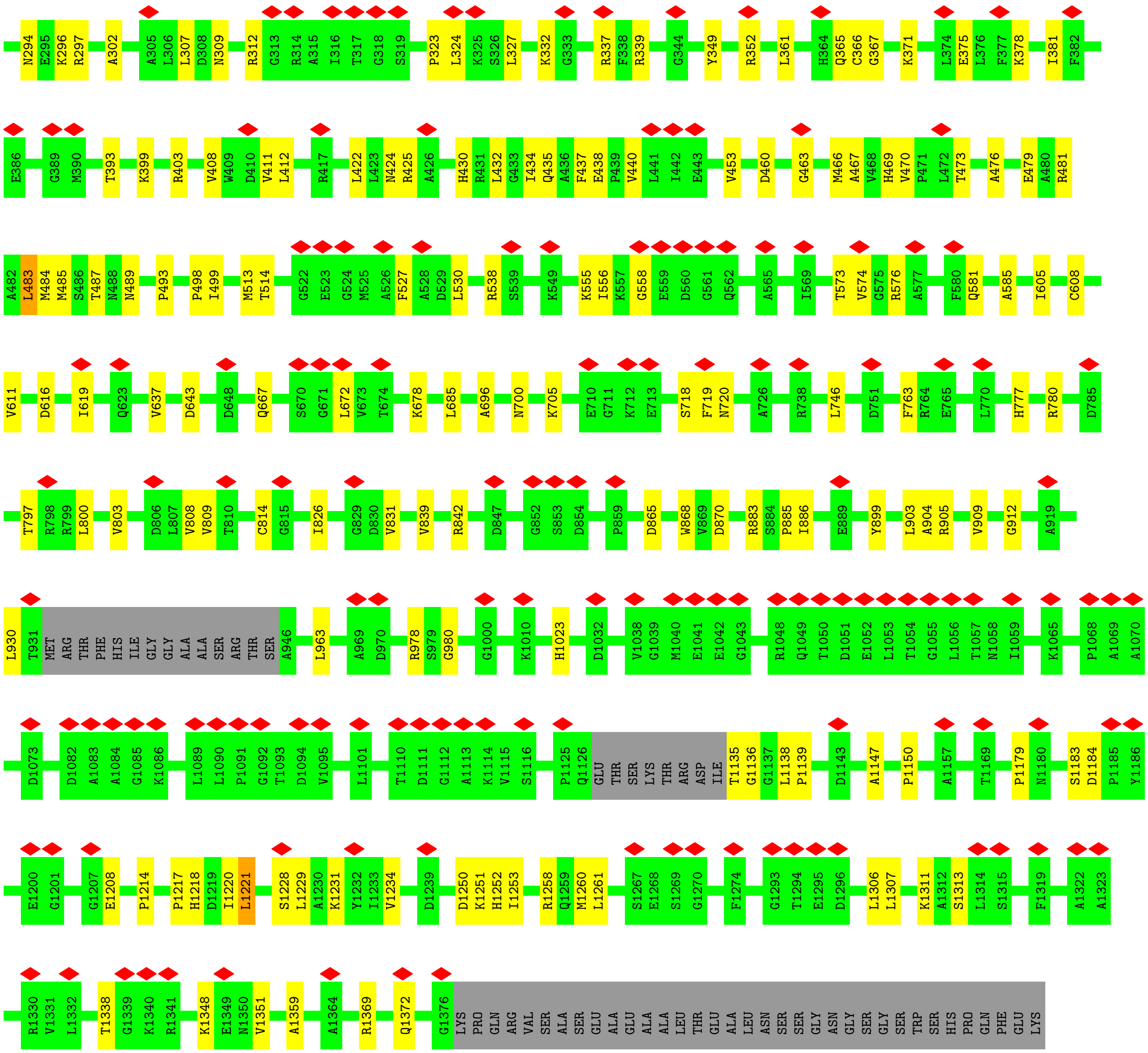




• Molecule 3: DNA-directed RNA polymerase subunit beta'

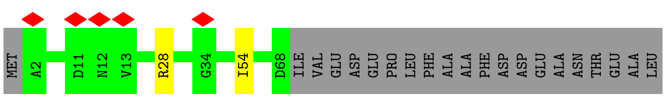
Chain D:





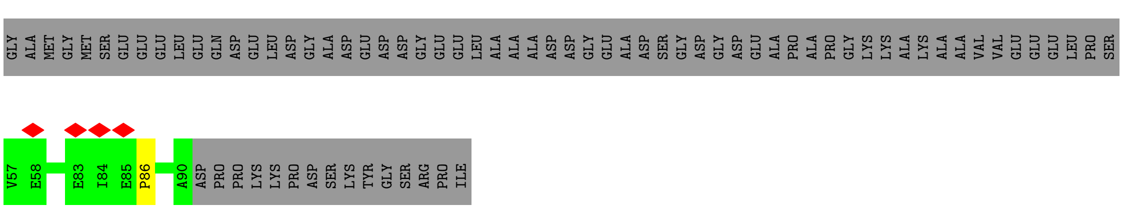
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:

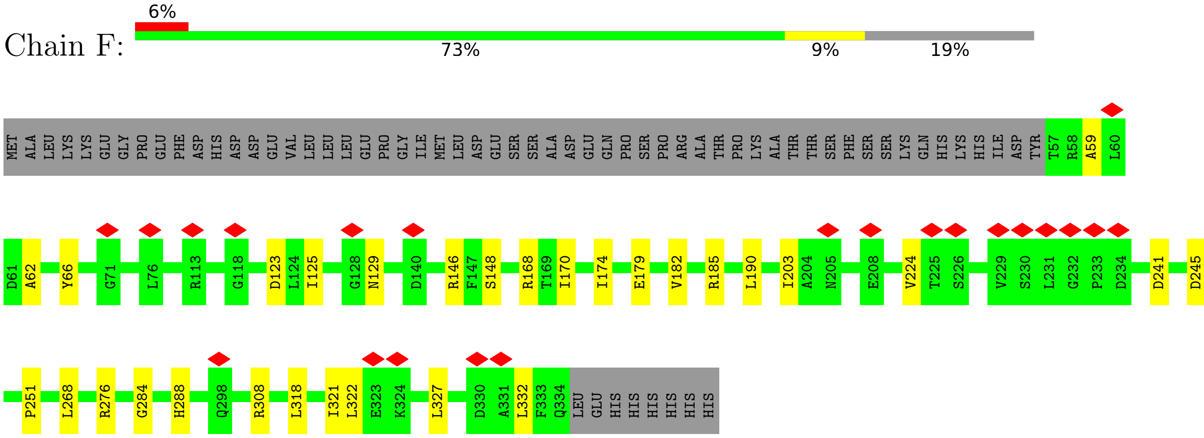


- Molecule 5: Transcriptional factor SutA

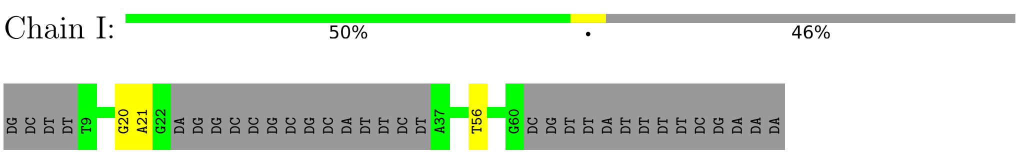
Chain G:



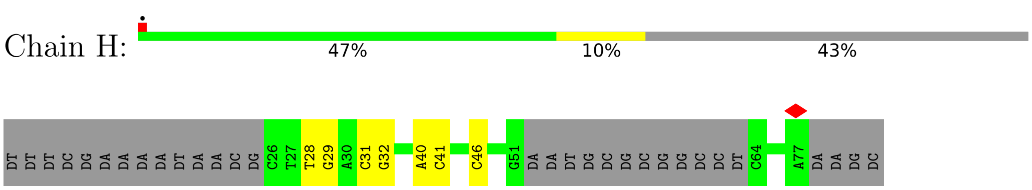
• Molecule 6: RNA polymerase sigma factor RpoS



• Molecule 7: DNA (70-MER)



• Molecule 8: DNA (70-MER)



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	112959	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	60.8	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.042	Depositor
Minimum map value	-0.017	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0065	Depositor
Map size (Å)	300.0, 300.0, 300.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0, 1.0, 1.0	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	0.25	0/1571	0.52	0/2134
1	B	0.25	0/1511	0.53	0/2055
2	C	0.26	0/9308	0.50	0/12643
3	D	0.26	0/9455	0.51	3/12839 (0.0%)
4	E	0.25	0/463	0.46	0/628
5	G	0.23	0/180	0.35	0/247
6	F	0.24	0/1931	0.47	0/2628
7	I	0.57	0/869	0.92	0/1336
8	H	0.51	0/924	0.88	0/1422
All	All	0.28	0/26212	0.55	3/35932 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	483	LEU	CA-CB-CG	5.72	128.45	115.30
3	D	1261	LEU	CA-CB-CG	5.68	128.37	115.30
3	D	685	LEU	CA-CB-CG	5.43	127.79	115.30

There are no chirality outliers.

There are no planarity outliers.

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	A	1554	0	1540	11	0
1	B	1500	0	1458	14	0
2	C	9188	0	8315	102	0
3	D	9324	0	8700	151	0
4	E	460	0	423	2	0
5	G	181	0	102	1	0
6	F	1911	0	1643	24	0
7	I	777	0	429	2	0
8	H	824	0	450	4	0
9	D	2	0	0	0	0
10	D	1	0	0	0	0
All	All	25722	0	23060	281	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 6.

The worst 5 of 281 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)
3:D:1218:HIS:ND1	3:D:1306:LEU:O	1.62	1.29
3:D:930:LEU:HD11	3:D:1138:LEU:CD2	1.88	1.04
3:D:1218:HIS:CE1	3:D:1306:LEU:O	2.12	1.02
3:D:1220:ILE:CG2	3:D:1229:LEU:HD12	1.96	0.95
3:D:1147:ALA:O	3:D:1218:HIS:CD2	2.24	0.90

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	A	211/345 (61%)	206 (98%)	5 (2%)	0	100	100
1	B	201/345 (58%)	194 (96%)	7 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	1320/1359 (97%)	1264 (96%)	56 (4%)	0	100	100
3	D	1333/1412 (94%)	1282 (96%)	49 (4%)	2 (0%)	47	81
4	E	65/88 (74%)	63 (97%)	2 (3%)	0	100	100
5	G	32/109 (29%)	31 (97%)	1 (3%)	0	100	100
6	F	276/342 (81%)	267 (97%)	9 (3%)	0	100	100
All	All	3438/4000 (86%)	3307 (96%)	129 (4%)	2 (0%)	54	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	1023	HIS
3	D	1221	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	A	150/295 (51%)	150 (100%)	0	100	100
1	B	143/295 (48%)	142 (99%)	1 (1%)	84	90
2	C	793/1144 (69%)	789 (100%)	4 (0%)	88	93
3	D	802/1180 (68%)	800 (100%)	2 (0%)	93	96
4	E	37/74 (50%)	37 (100%)	0	100	100
5	G	3/82 (4%)	3 (100%)	0	100	100
6	F	145/307 (47%)	145 (100%)	0	100	100
All	All	2073/3377 (61%)	2066 (100%)	7 (0%)	92	94

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1120	VAL
2	C	1318	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	885	PRO
3	D	678	LYS
2	C	736	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	277	ASN
3	D	700	ASN
3	D	465	GLN
3	D	777	HIS
2	C	767	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)

Of 3 ligands modelled in this entry, 3 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)

There are no chain breaks in this entry.

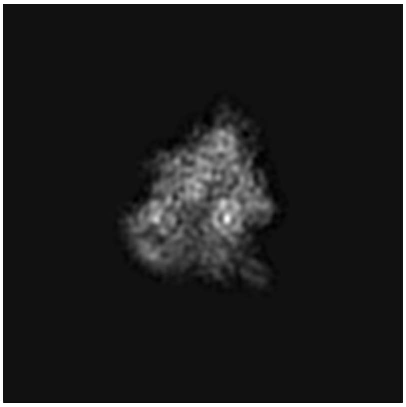
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-31403. These allow visual inspection of the internal detail of the map and identification of artifacts.

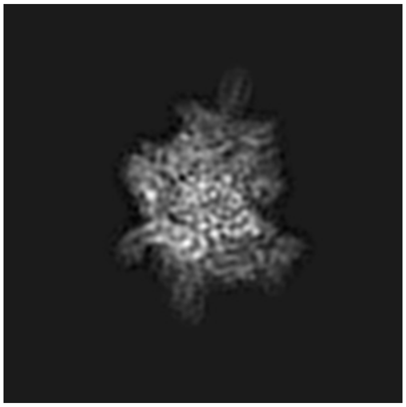
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

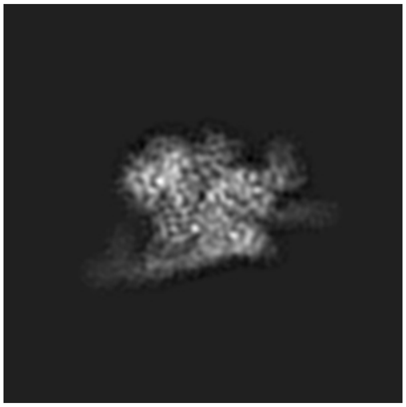
6.1.1 Primary map



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: 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

6.3.1 Primary map

X Index: 128

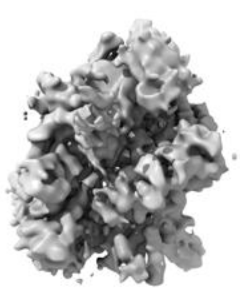
Y Index: 168

Z Index: 148

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

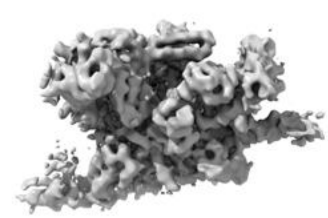
6.4 Orthogonal surface views

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0065. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

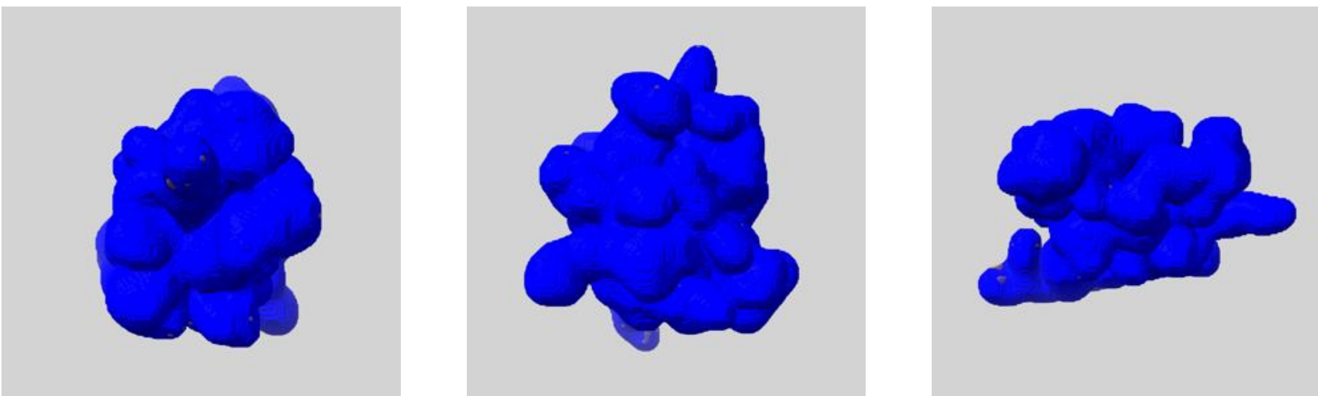
6.5 Mask visualisation (i)

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

6.5.1 emd_31403_msk_1.map (i)

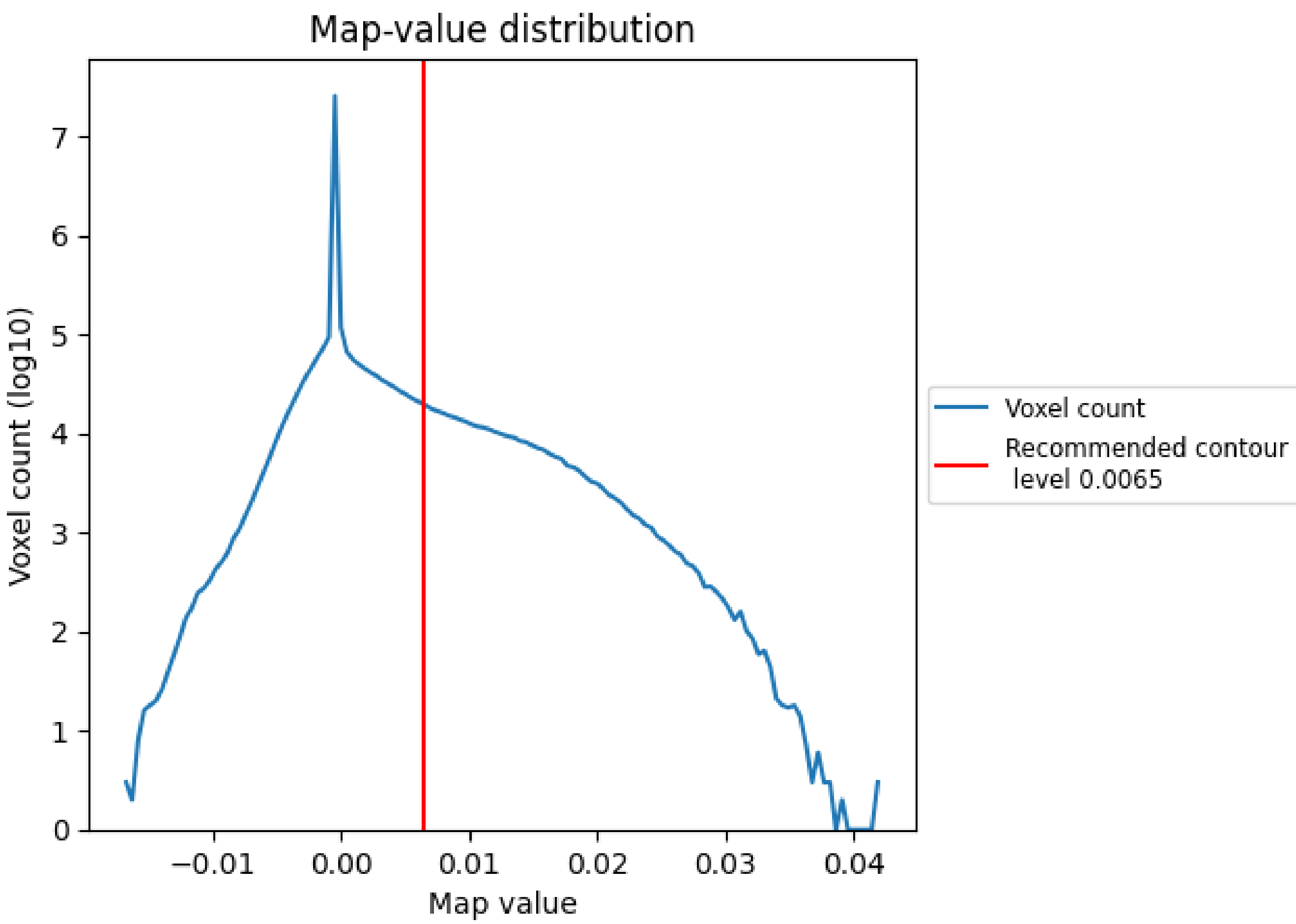


X Y Z

7 Map analysis (i)

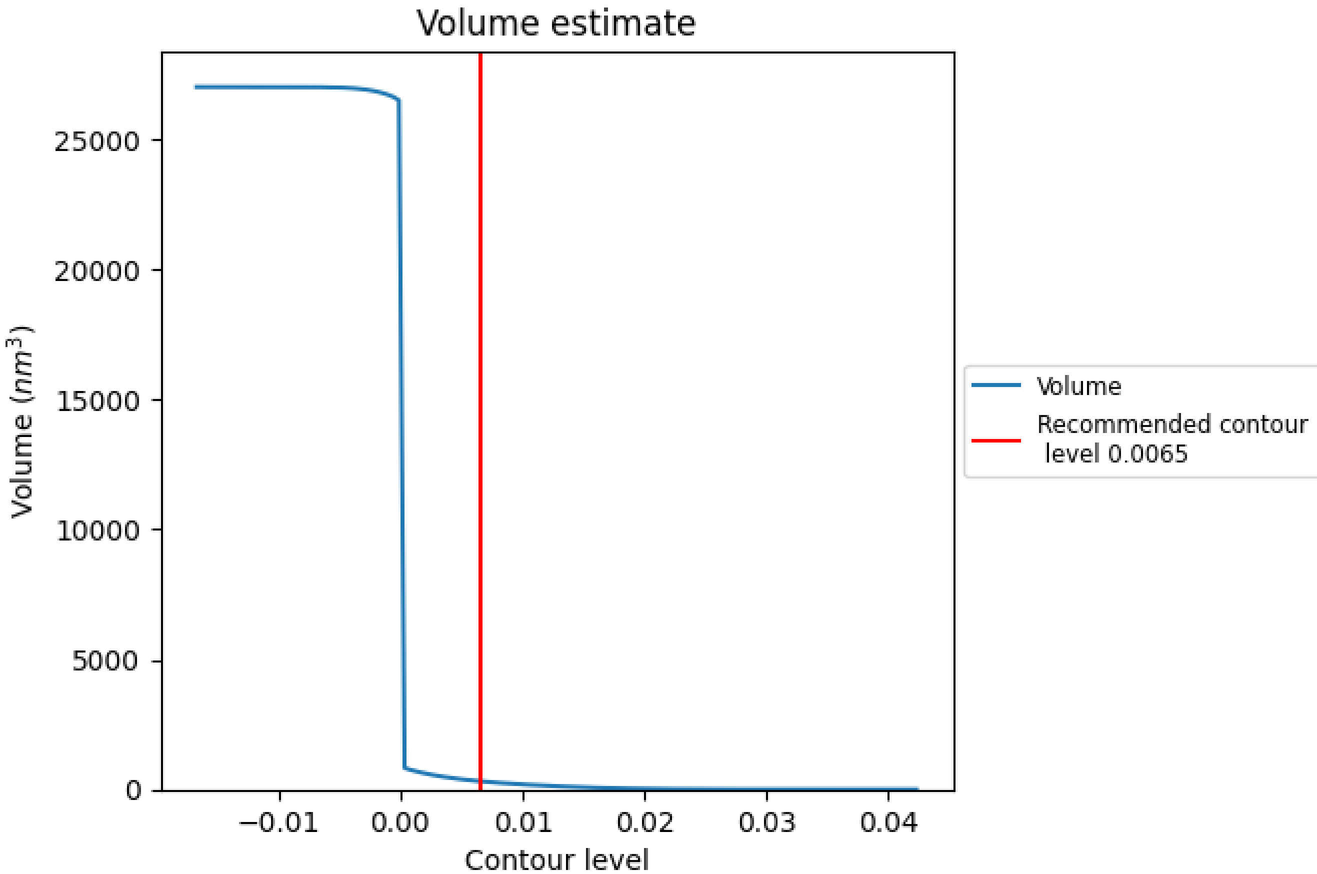
This section contains the results of statistical analysis of the map.

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.

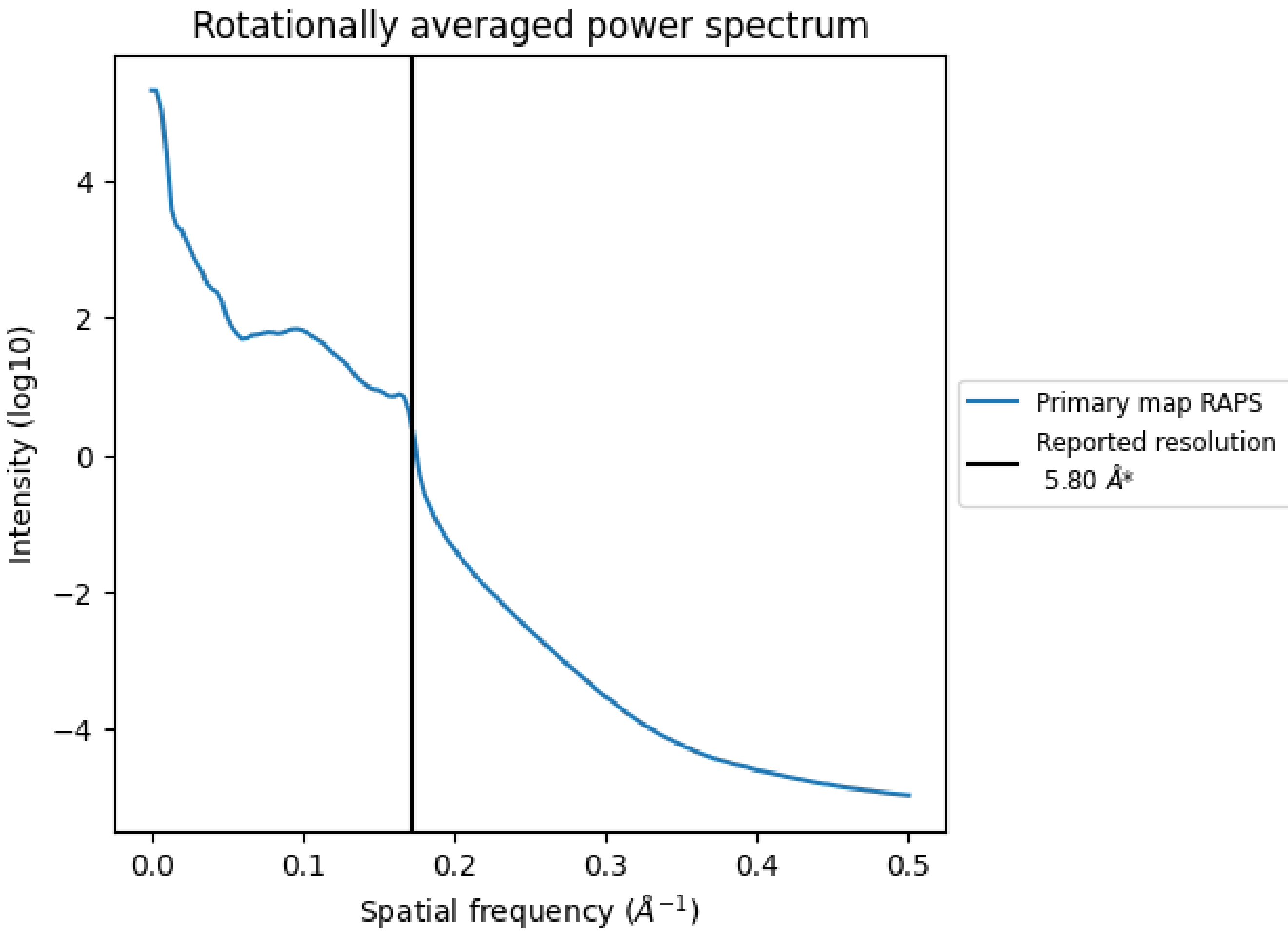
7.2 Volume estimate ⓘ



The volume at the recommended contour level is 316 nm^3; this corresponds to an approximate mass of 286 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)

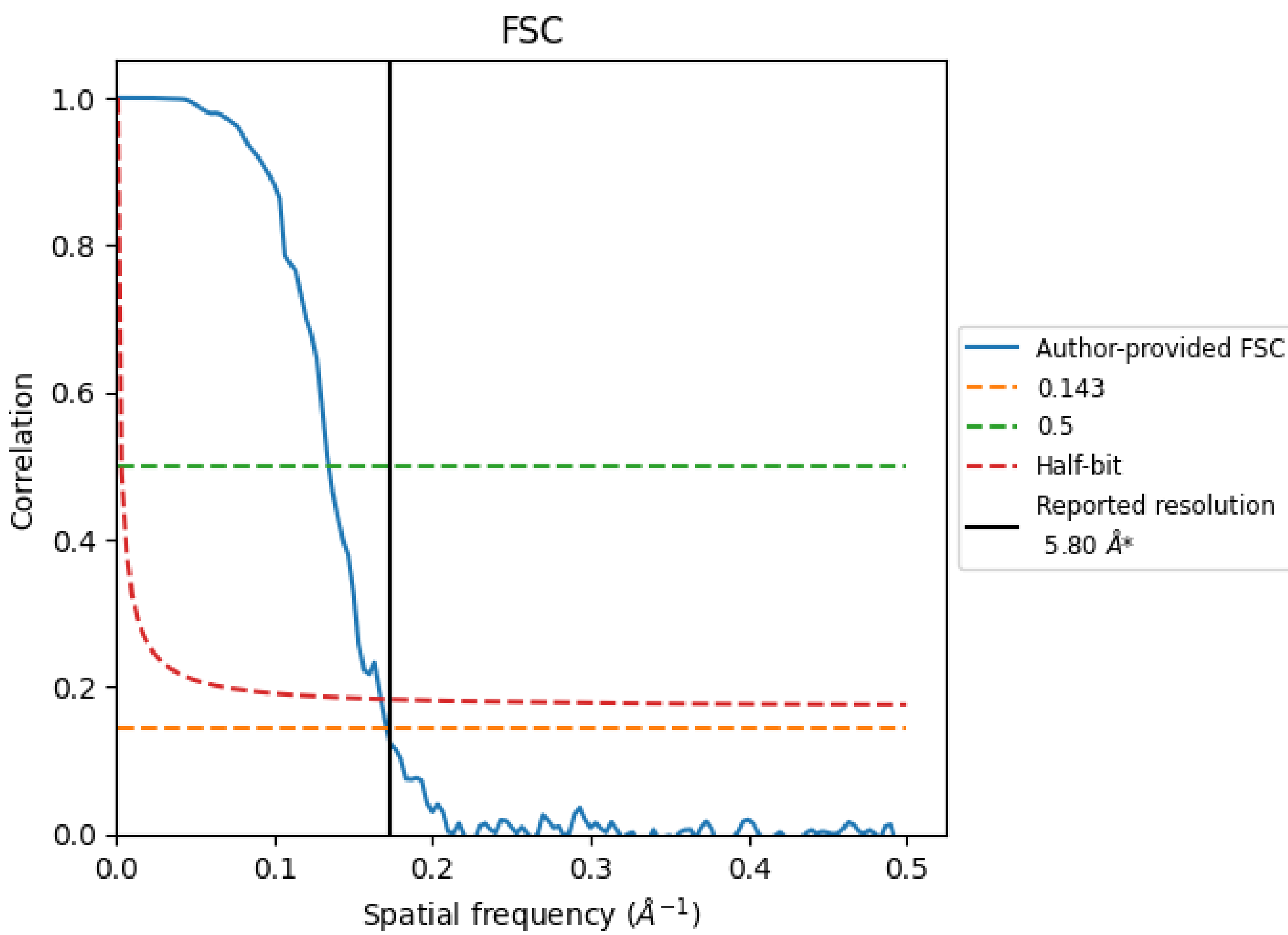


*Reported resolution corresponds to spatial frequency of 0.172 Å^{-1}

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 0.172 $Å^{-1}$

8.2 Resolution estimates (i)

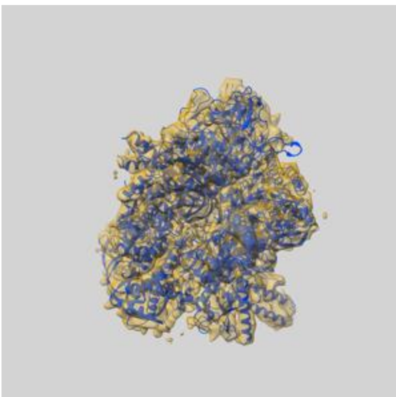
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	5.80	-	-
Author-provided FSC curve	5.84	7.45	5.97
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

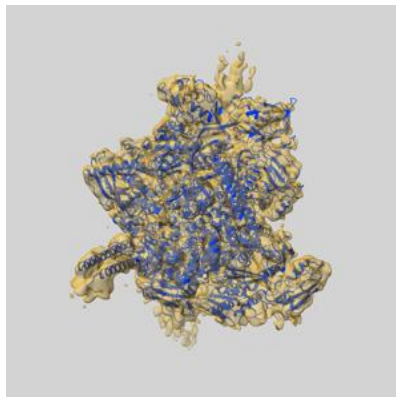
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-31403 and PDB model 7F0R. Per-residue inclusion information can be found in section 3 on page 8.

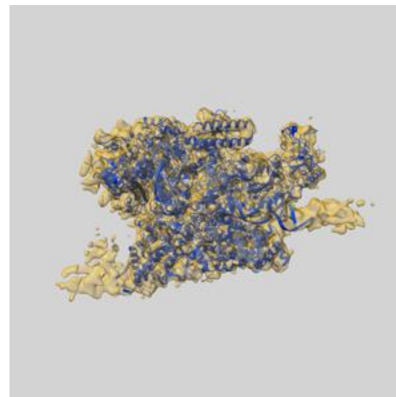
9.1 Map-model overlay (i)



X



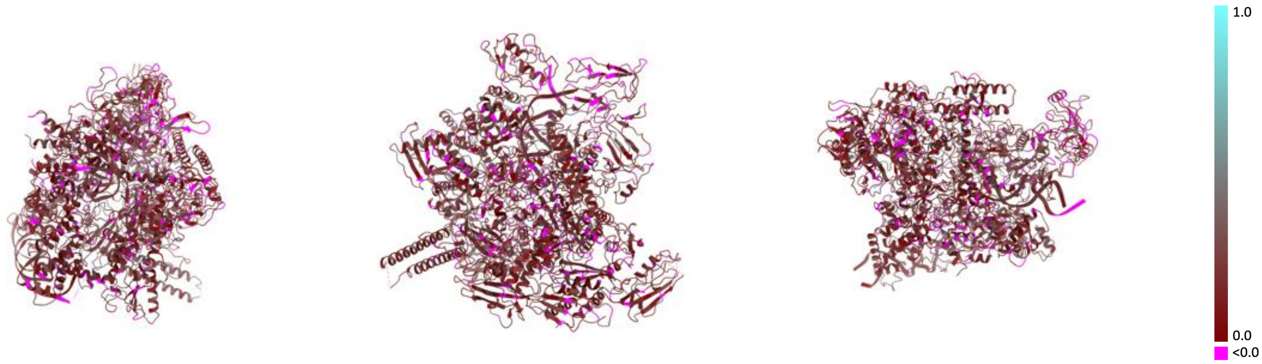
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0065 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.

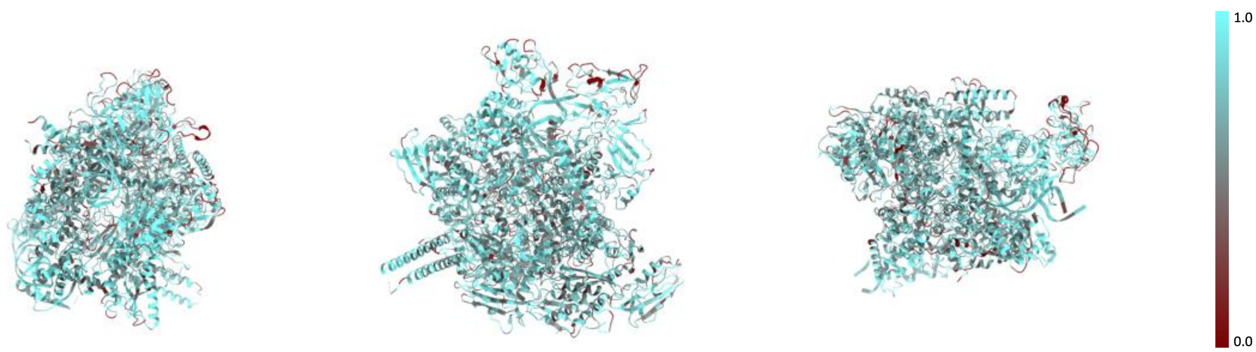
9.2 Q-score mapped to coordinate model (i)



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.

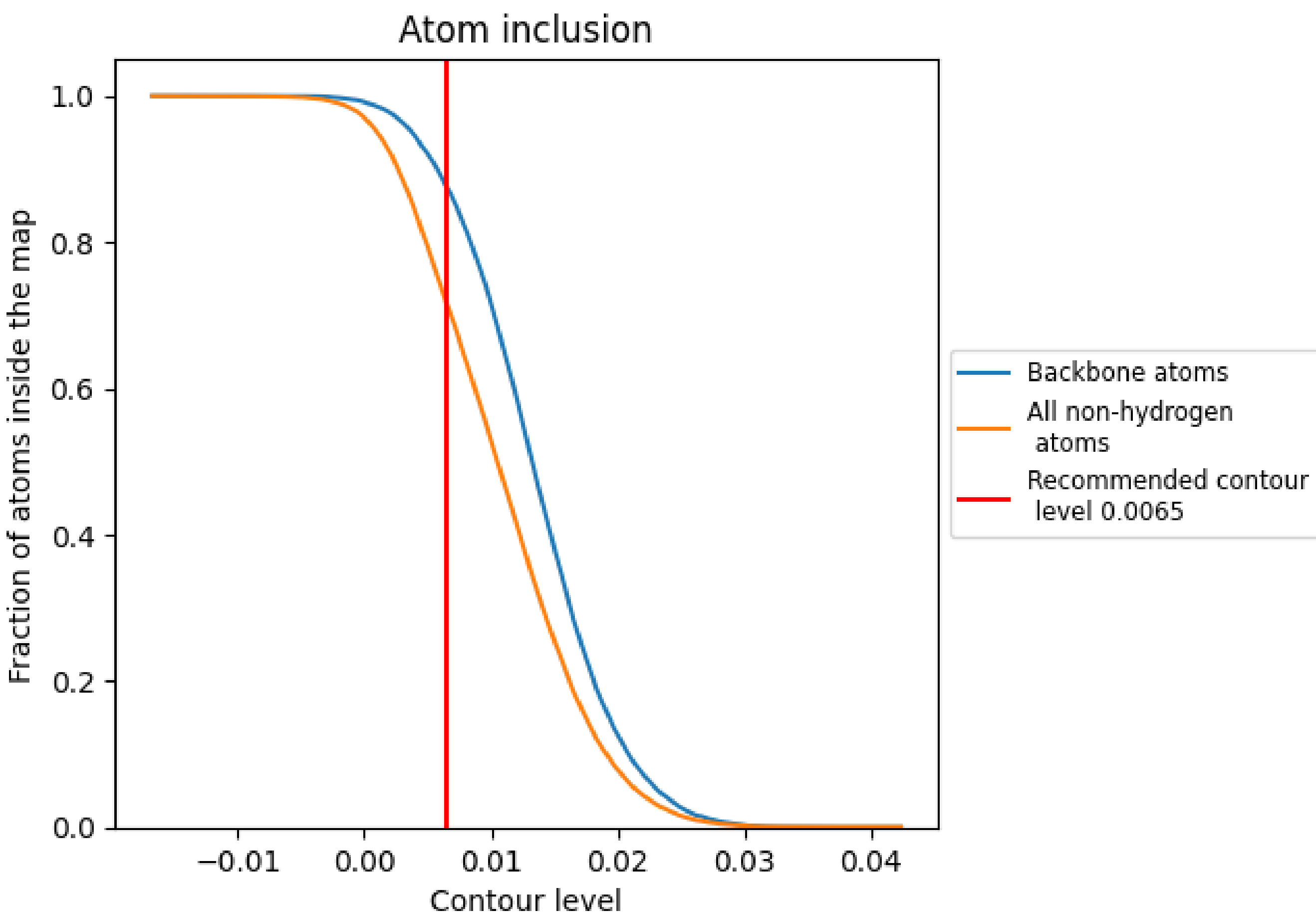
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0065).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 72% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.0065) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7154	0.1600
A	0.7079	0.1630
B	0.7119	0.1520
C	0.7044	0.1550
D	0.6992	0.1560
E	0.7184	0.1730
F	0.7320	0.1730
G	0.7989	0.2490
H	0.8386	0.1900
I	0.8649	0.1950