wwPDB EM Validation Summary Report ⓘ

Nov 26, 2022 – 07:43 PM EST

PDB ID	:	7SZJ
EMDB ID	:	EMD-25570
Title	:	Cryo-EM structure of Rifamycin bound to E. coli RNAP and rrnBP1 promoter complex
Authors	:	Shin, Y.; Murakami, K.S.
Deposited on	:	2021-11-28
Resolution	:	3.11 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:

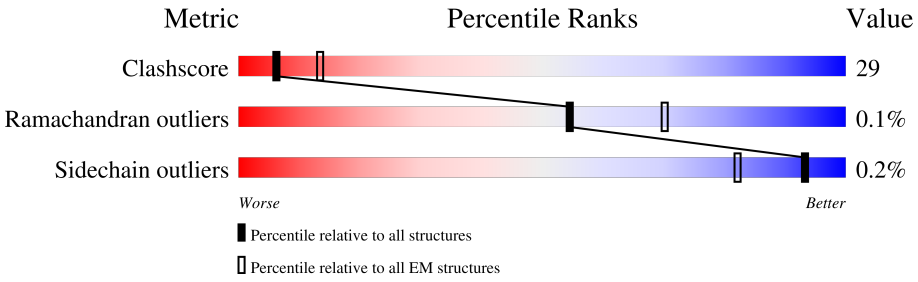
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.11 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	329	47% 24% 30%
1	B	329	38% 32% 30%
2	C	1342	63% 37%
3	D	1407	5% 56% 39% 5%
4	E	91	7% 54% 30% 16%
5	F	613	18% 34% 42% 24%
6	X	64	14% 19% 70% 11%
7	Y	64	12% 14% 62% 23%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 31231 atoms, of which 58 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
1	A	231	Total	C	N	O	S	0	0
			1794	1117	318	353	6		
1	B	230	Total	C	N	O	S	0	0
			1786	1112	317	351	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0
			10570	6631	1841	2055	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1340	Total	C	N	O	S	0	0
			10382	6522	1849	1962	49		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	76	Total	C	N	O	S	0	0
			605	368	115	121	1		

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	466	Total	C	N	O	S	0	0
			3799	2384	679	713	23		

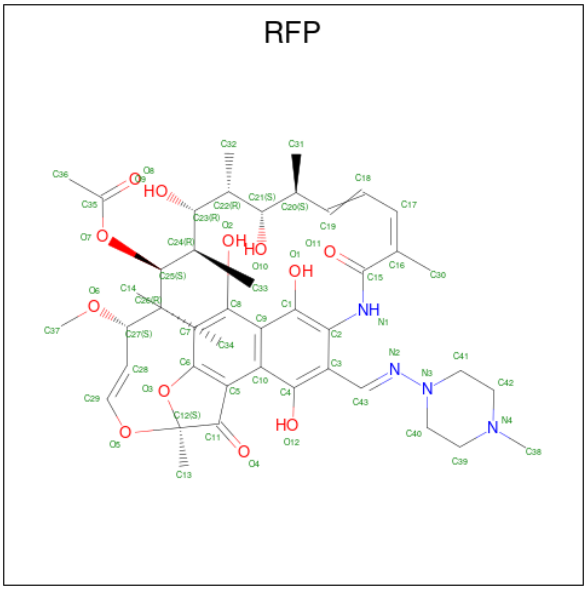
- Molecule 6 is a DNA chain called DNA (57-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	X	57	Total	C	N	O	P	0	0
			1162	554	208	343	57		

- Molecule 7 is a DNA chain called DNA (49-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	Y	49	Total	C	N	O	P	0	0
			1013	480	192	292	49		

- Molecule 8 is RIFAMPICIN (three-letter code: RFP) (formula: $C_{43}H_{58}N_4O_{12}$).



Mol	Chain	Residues	Atoms					AltConf
8	C	1	Total	C	H	N	O	0
			117	43	58	4	12	

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

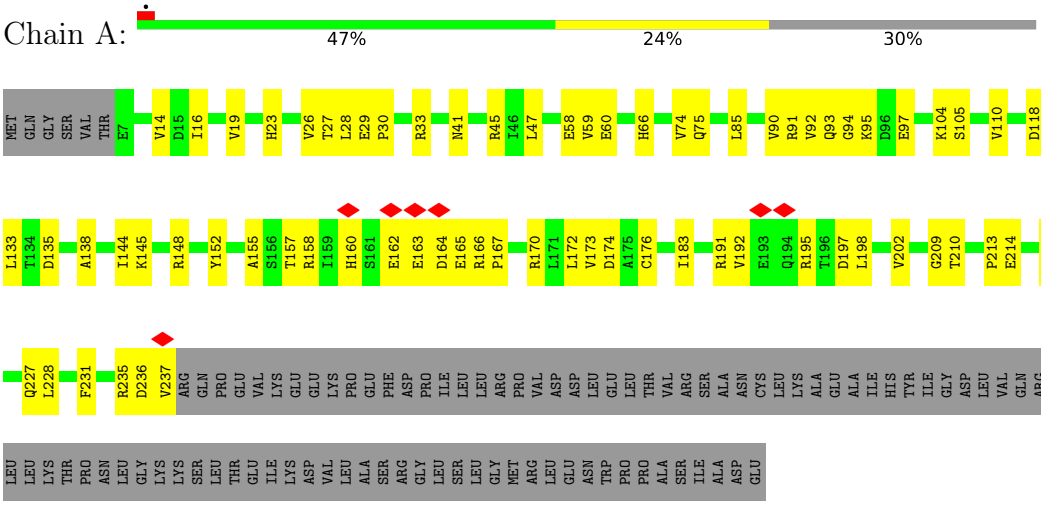
Mol	Chain	Residues	Atoms		AltConf
10	D	2	Total	Zn	0
			2	2	

3 Residue-property plots (i)

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

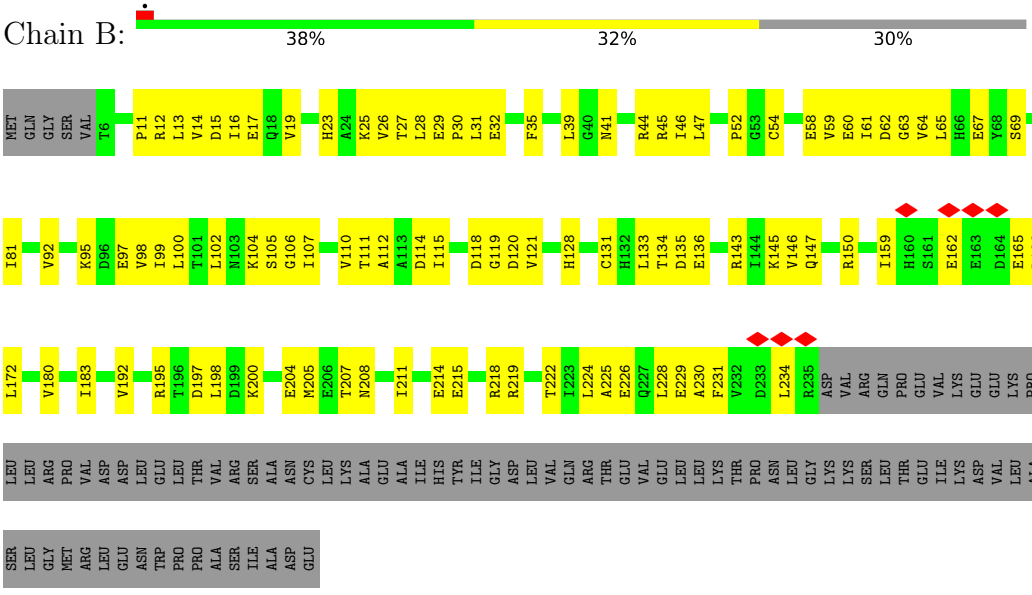
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:

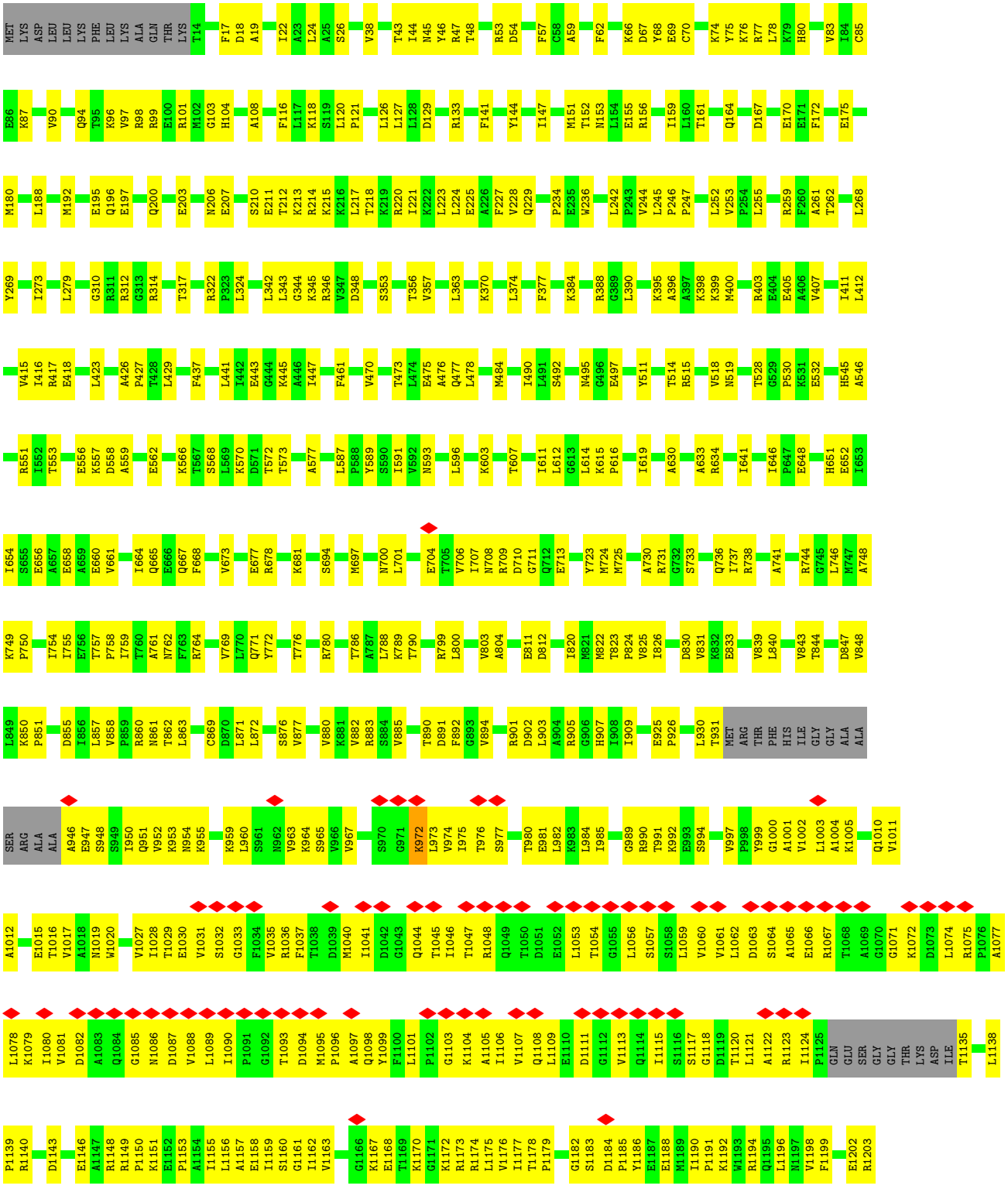


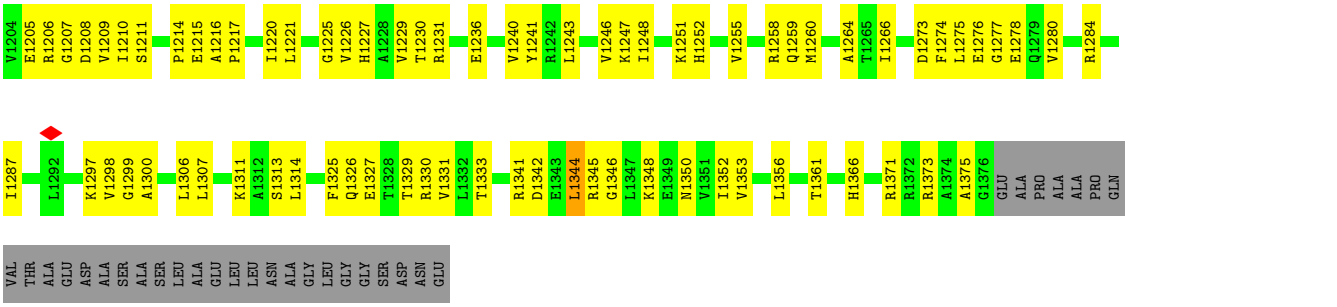
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:

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

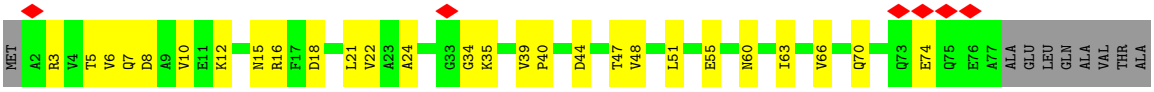
Chain D:





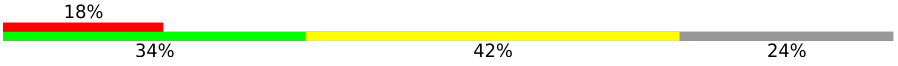
• Molecule 4: DNA-directed RNA polymerase subunit omega

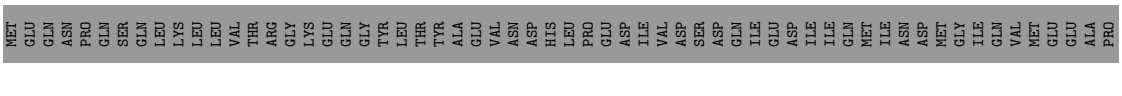
Chain E:

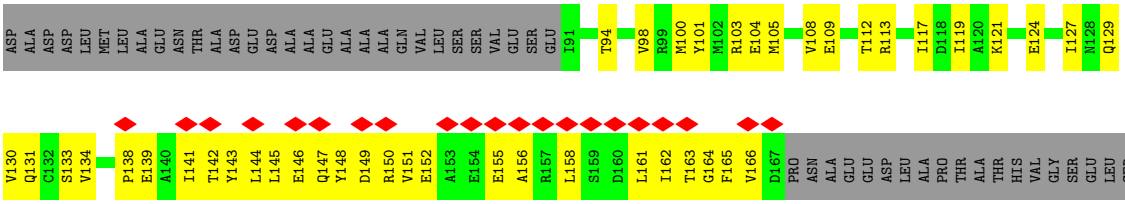


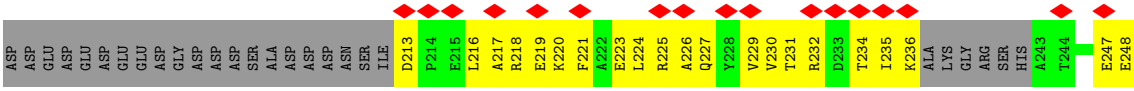
• Molecule 5: RNA polymerase sigma factor RpoD

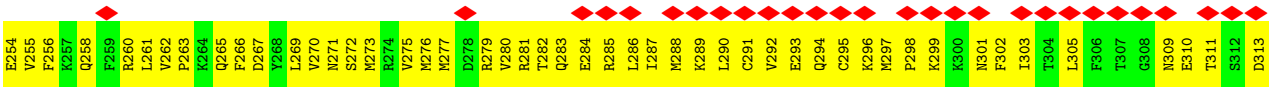
Chain F:

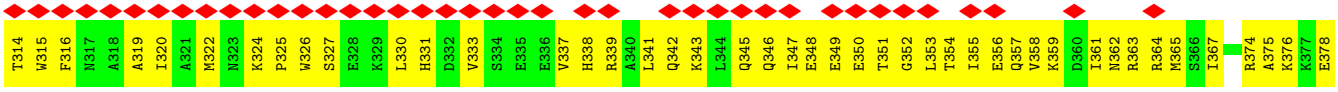


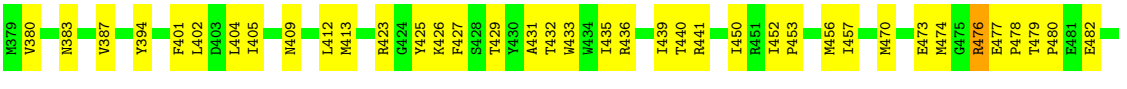


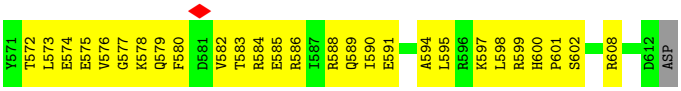




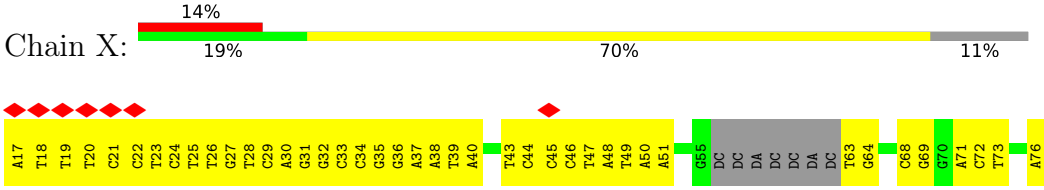




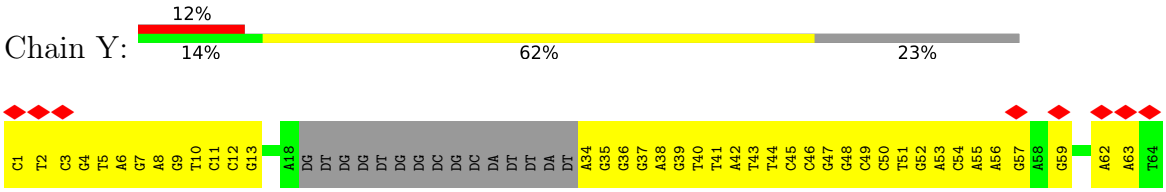




• Molecule 6: DNA (57-MER)



• Molecule 7: DNA (49-MER)



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	251427	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 ($e^-/Å^2$)	45	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	5.476	Depositor
Minimum map value	-3.679	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.114	Depositor
Recommended contour level	0.4	Depositor
Map size (Å)	403.2, 403.2, 403.2	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.12, 1.12, 1.12	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.28	0/1816	0.53	0/2461
1	B	0.26	0/1808	0.51	0/2450
2	C	0.29	0/10739	0.50	0/14489
3	D	0.28	0/10539	0.51	0/14234
4	E	0.25	0/607	0.50	0/817
5	F	0.24	0/3849	0.49	0/5171
6	X	0.46	0/1300	0.86	0/2000
7	Y	0.44	0/1137	0.83	0/1752
All	All	0.29	0/31795	0.55	0/43374

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
3	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	1344	LEU	Peptide

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	1794	0	1819	79	0
1	B	1786	0	1813	107	0
2	C	10570	0	10582	504	0
3	D	10382	0	10570	642	0
4	E	605	0	612	32	0
5	F	3799	0	3885	380	0
6	X	1162	0	644	71	0
7	Y	1013	0	552	56	0
8	C	59	58	58	0	0
9	D	1	0	0	0	0
10	D	2	0	0	0	0
All	All	31173	58	30535	1770	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 29.

The worst 5 of 1770 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)
2:C:1025:PHE:HA	2:C:1028:LYS:HE2	1.19	1.16
3:D:126:LEU:HD11	3:D:223:LEU:HD13	1.21	1.12
2:C:1160:ASP:HB2	2:C:1161:LEU:HA	1.33	1.11
3:D:1056:LEU:HD12	3:D:1108:GLN:HB2	1.36	1.07
3:D:1046:ILE:HD12	3:D:1059:LEU:HD22	1.34	1.06

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	229/329 (70%)	217 (95%)	12 (5%)	0	100	100
1	B	228/329 (69%)	218 (96%)	10 (4%)	0	100	100
2	C	1338/1342 (100%)	1221 (91%)	116 (9%)	1 (0%)	51	83
3	D	1334/1407 (95%)	1244 (93%)	89 (7%)	1 (0%)	51	83
4	E	74/91 (81%)	70 (95%)	4 (5%)	0	100	100
5	F	458/613 (75%)	440 (96%)	18 (4%)	0	100	100
All	All	3661/4111 (89%)	3410 (93%)	249 (7%)	2 (0%)	54	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	1345	ARG
2	C	398	SER

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	199/286 (70%)	199 (100%)	0	100	100
1	B	198/286 (69%)	198 (100%)	0	100	100
2	C	1155/1157 (100%)	1152 (100%)	3 (0%)	92	96
3	D	1113/1168 (95%)	1112 (100%)	1 (0%)	93	98
4	E	65/75 (87%)	65 (100%)	0	100	100
5	F	415/540 (77%)	414 (100%)	1 (0%)	93	97
All	All	3145/3512 (90%)	3140 (100%)	5 (0%)	93	97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	272	ARG
2	C	283	LYS
2	C	478	ARG
3	D	972	LYS
5	F	476	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	342	GLN
5	F	362	ASN
2	C	799	ASN
3	D	157	GLN
3	D	1098	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)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

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	RFP	C	3001	-	63,63,63	3.36	25 (39%)	94,94,94	1.75	17 (18%)

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	RFP	C	3001	-	-	16/60/85/85	0/5/5/5

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	C	3001	RFP	C6-C7	10.27	1.57	1.39
8	C	3001	RFP	O3-C6	8.39	1.53	1.37
8	C	3001	RFP	C5-C10	7.57	1.59	1.43
8	C	3001	RFP	C38-N4	-7.48	1.29	1.46
8	C	3001	RFP	C3-C4	7.47	1.54	1.40

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	C	3001	RFP	O3-C6-C7	6.08	131.59	121.14
8	C	3001	RFP	C2-C3-C43	-5.42	118.66	124.17
8	C	3001	RFP	O7-C35-C36	4.63	119.60	111.09
8	C	3001	RFP	C4-C3-C43	4.14	121.47	116.52
8	C	3001	RFP	O4-C11-C5	-4.06	124.07	131.81

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

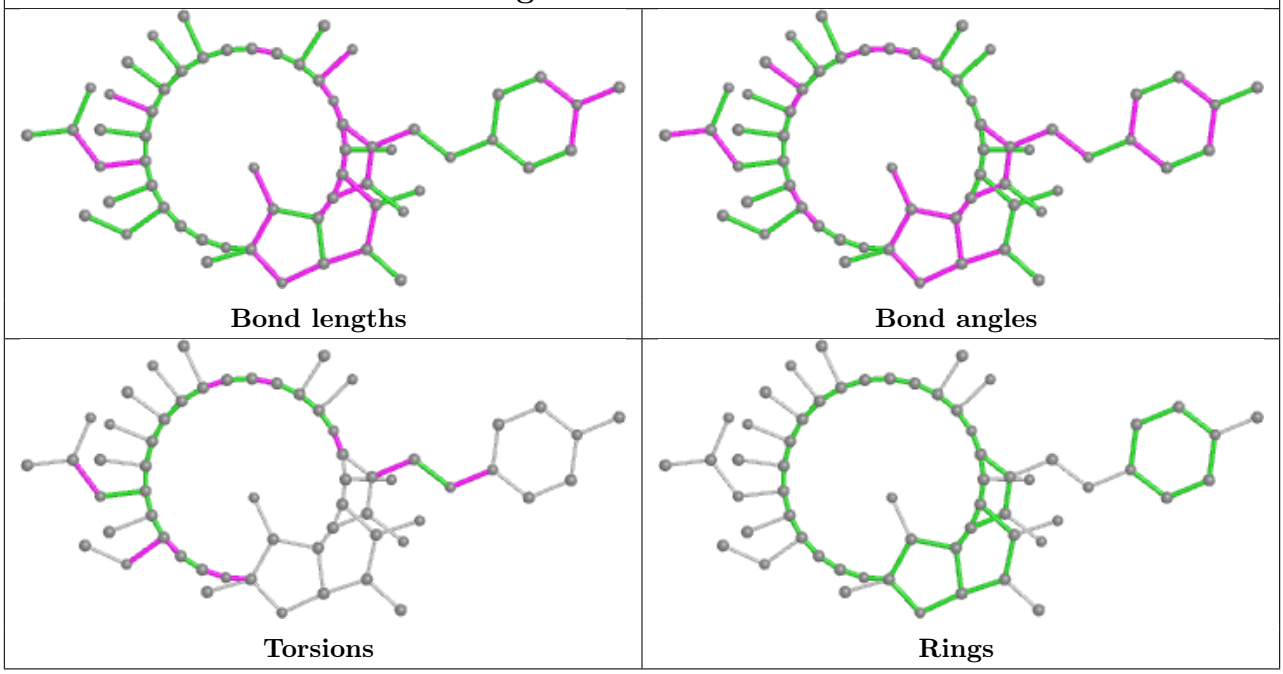
Mol	Chain	Res	Type	Atoms
8	C	3001	RFP	C4-C3-C43-N2
8	C	3001	RFP	C11-C12-O5-C29
8	C	3001	RFP	C13-C12-O5-C29
8	C	3001	RFP	O3-C12-O5-C29
8	C	3001	RFP	C16-C17-C18-C19

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 RFP C 3001



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.

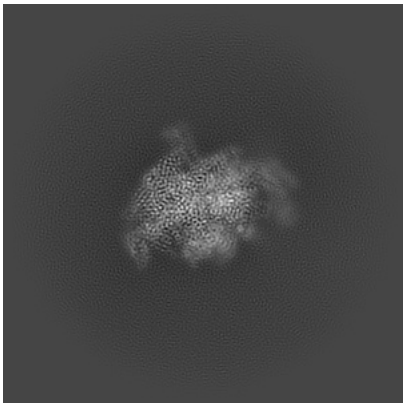
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-25570. These allow visual inspection of the internal detail of the map and identification of artifacts.

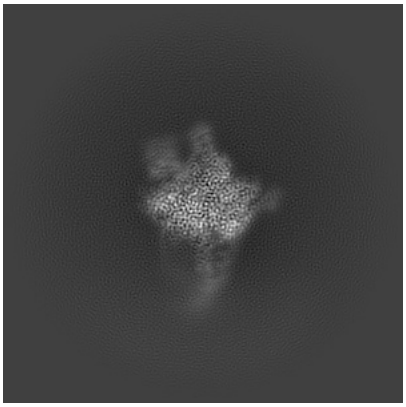
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

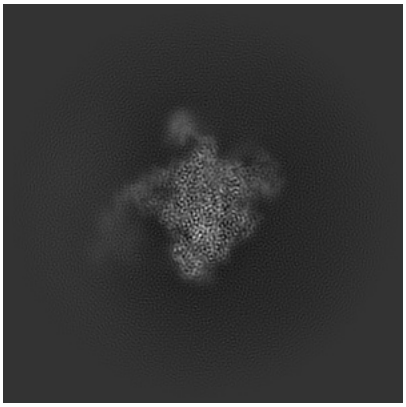
6.1.1 Primary map



X



Y

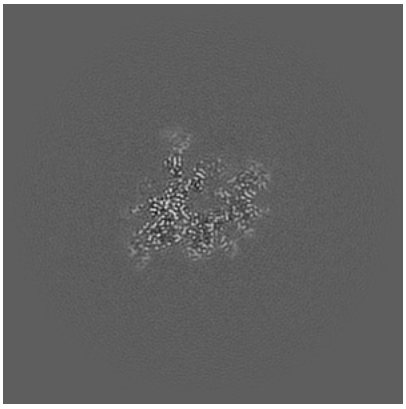


Z

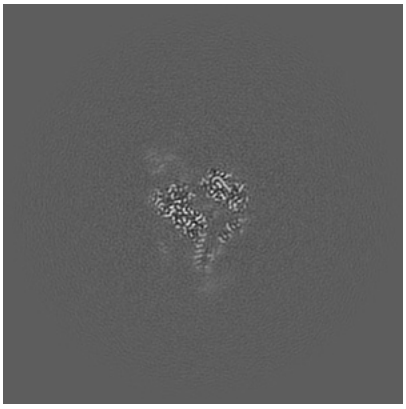
The images above show the map projected in three orthogonal directions.

6.2 Central slices

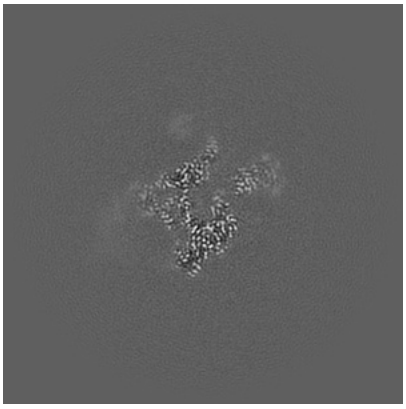
6.2.1 Primary map



X Index: 180



Y Index: 180

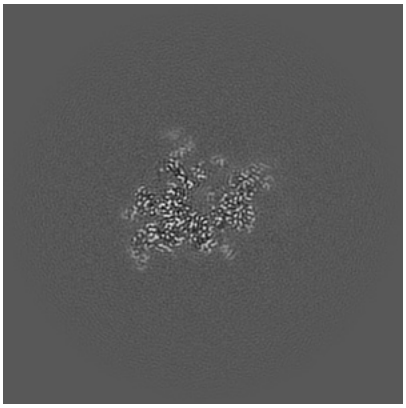


Z Index: 180

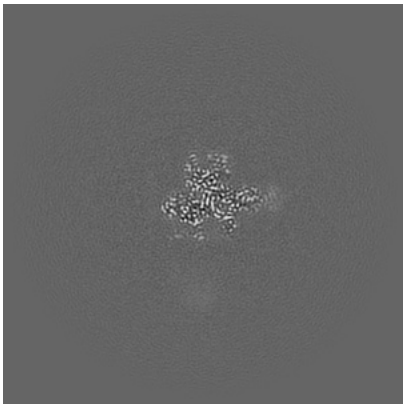
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

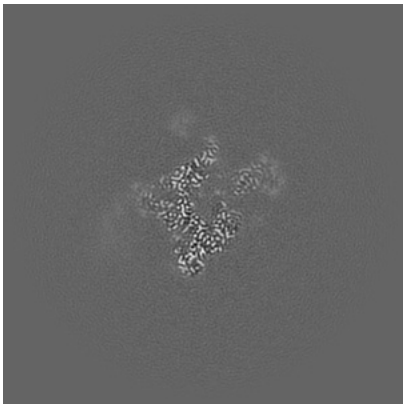
6.3.1 Primary map



X Index: 176



Y Index: 157



Z Index: 178

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

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.4. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

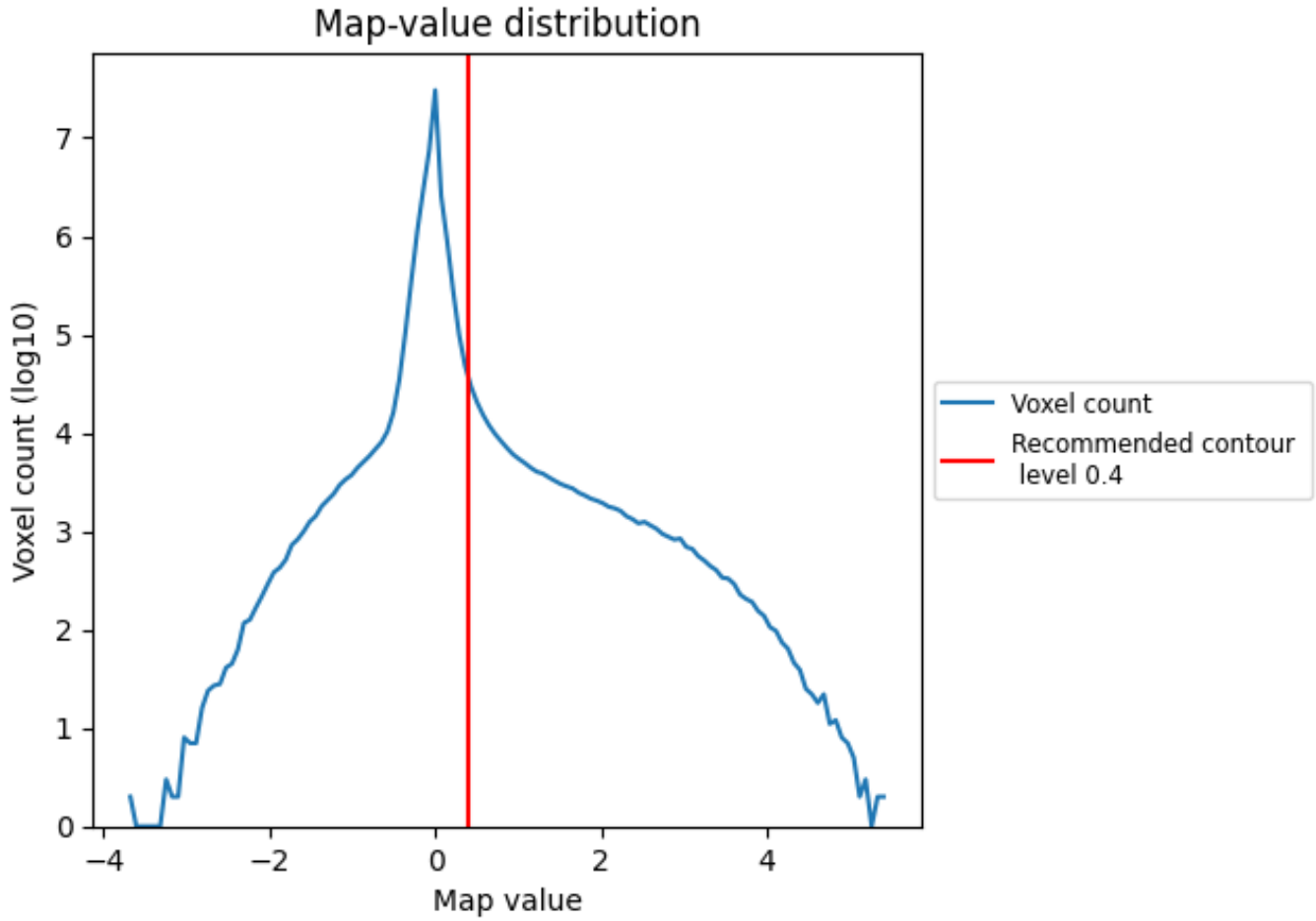
6.5 Mask visualisation (i)

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

7 Map analysis

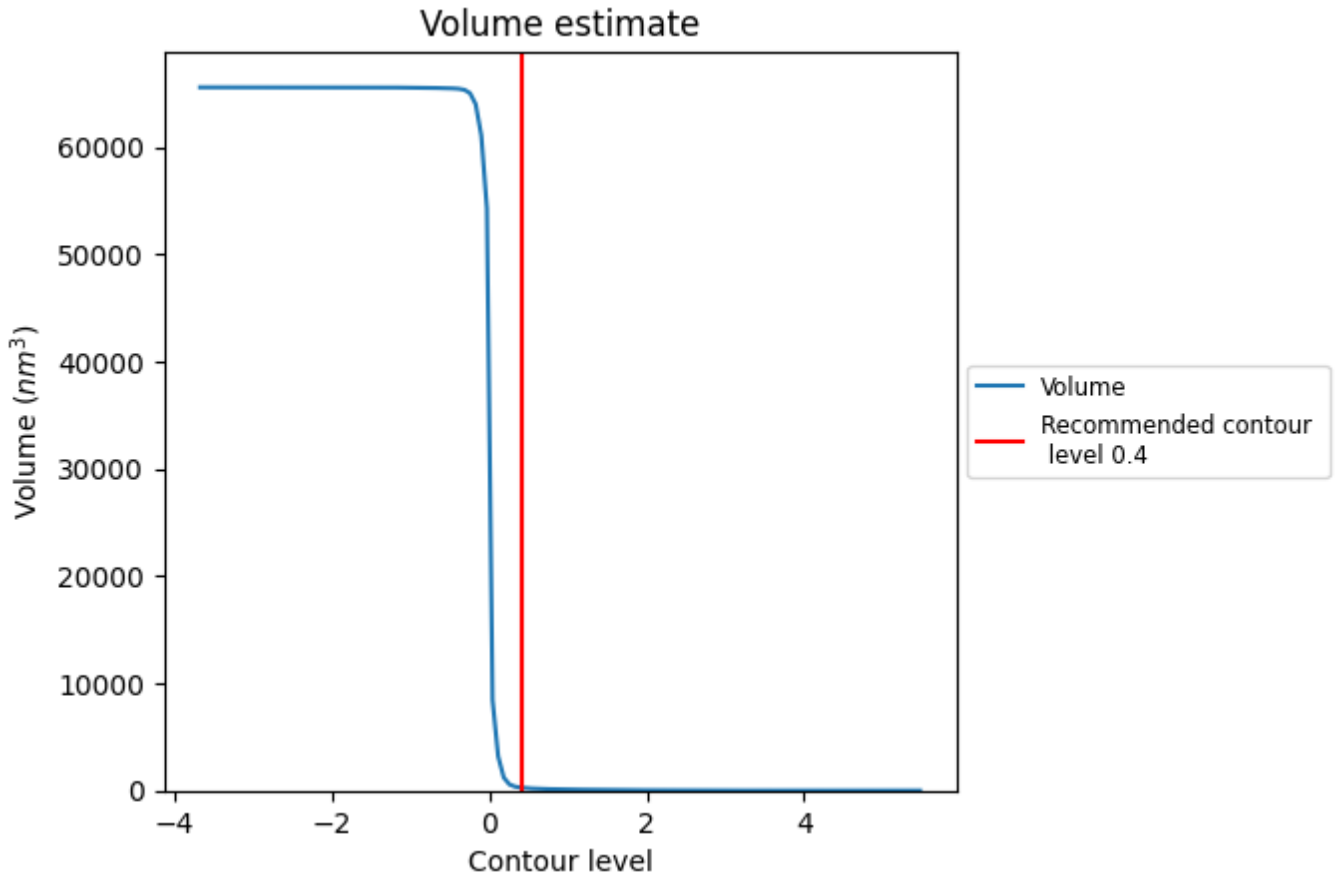
This section contains the results of statistical analysis of the map.

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.

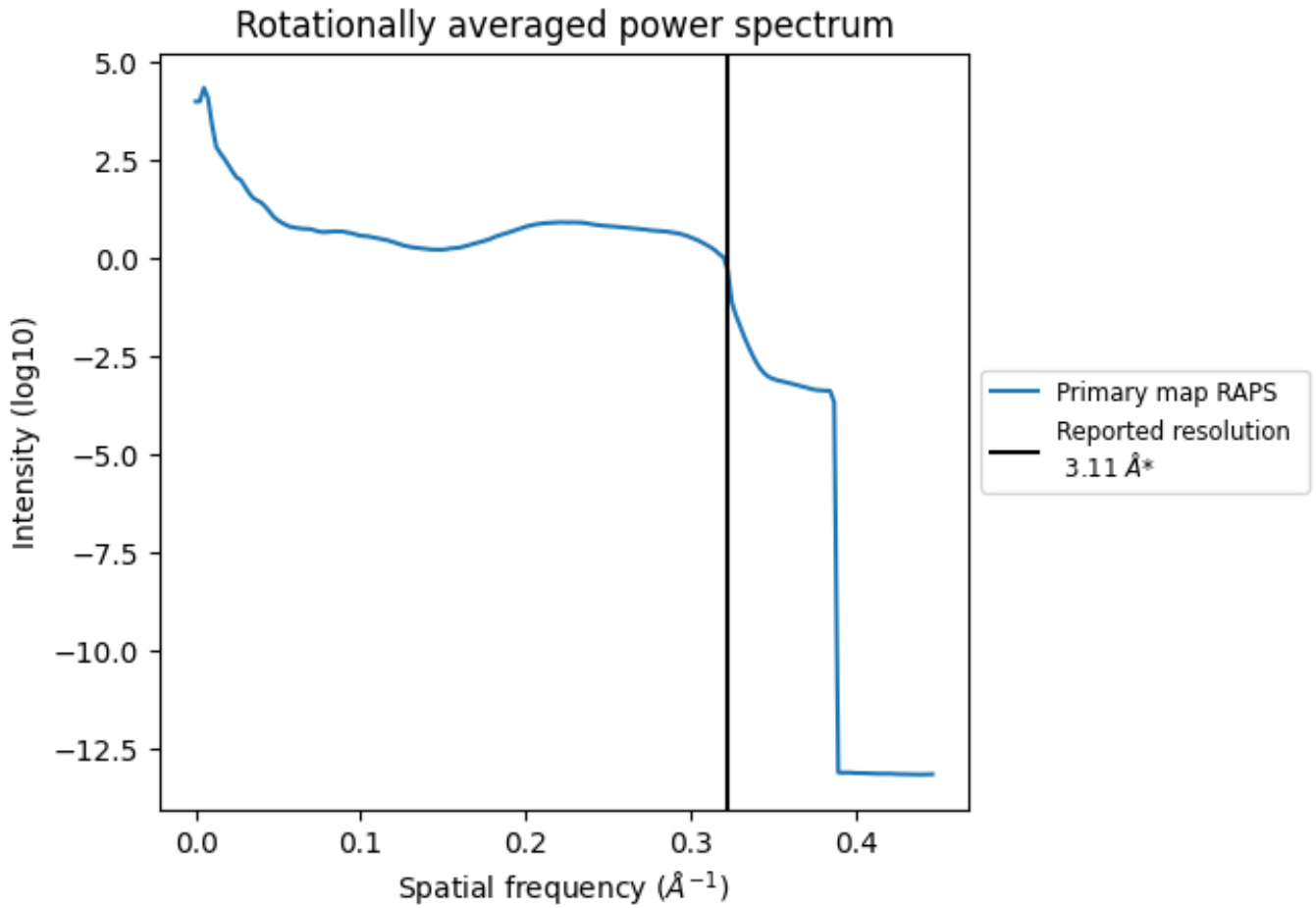
7.2 Volume estimate (i)



The volume at the recommended contour level is 275 nm^3; this corresponds to an approximate mass of 248 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 0.322 $Å^{-1}$

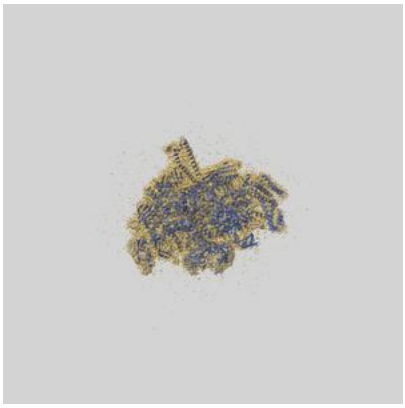
8 Fourier-Shell correlation

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

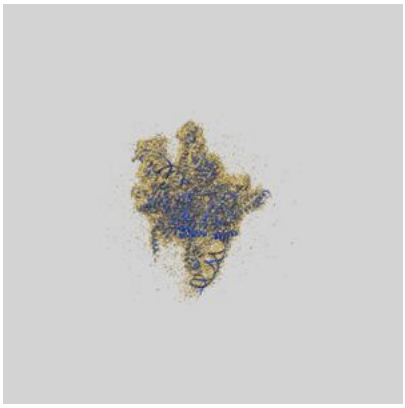
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-25570 and PDB model 7SZJ. Per-residue inclusion information can be found in section 3 on page 5.

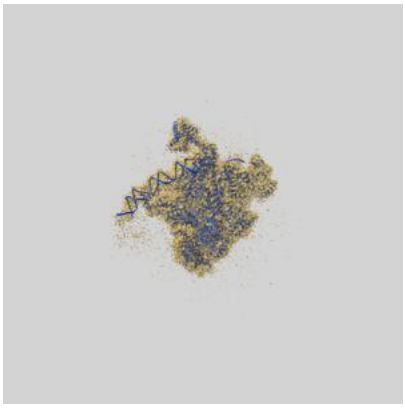
9.1 Map-model overlay



X



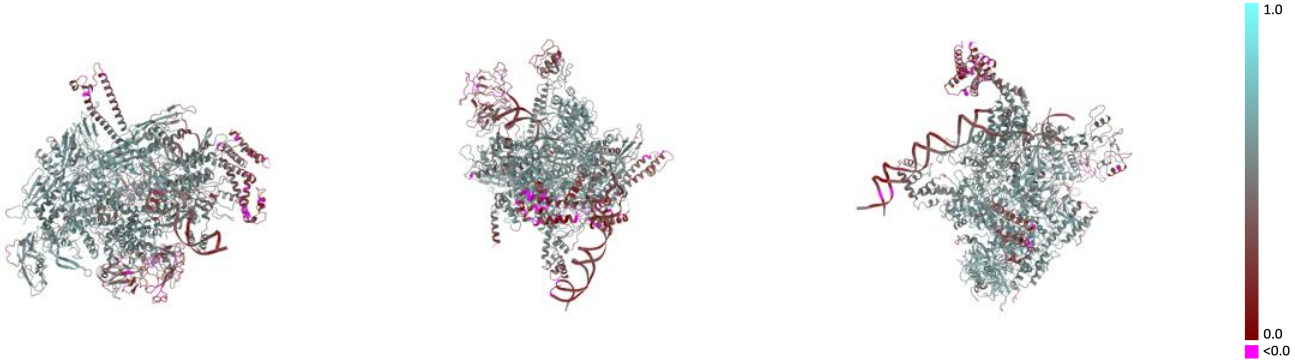
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.4 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.

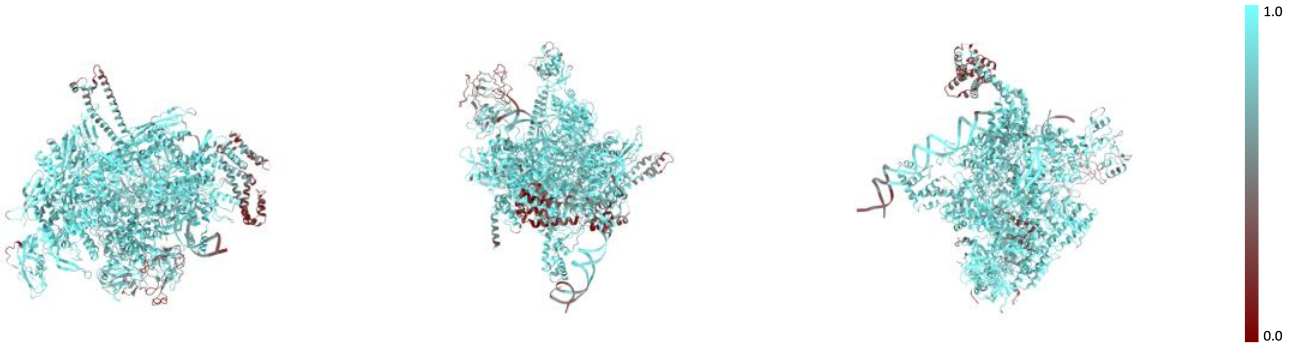
9.2 Q-score mapped to coordinate model



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.

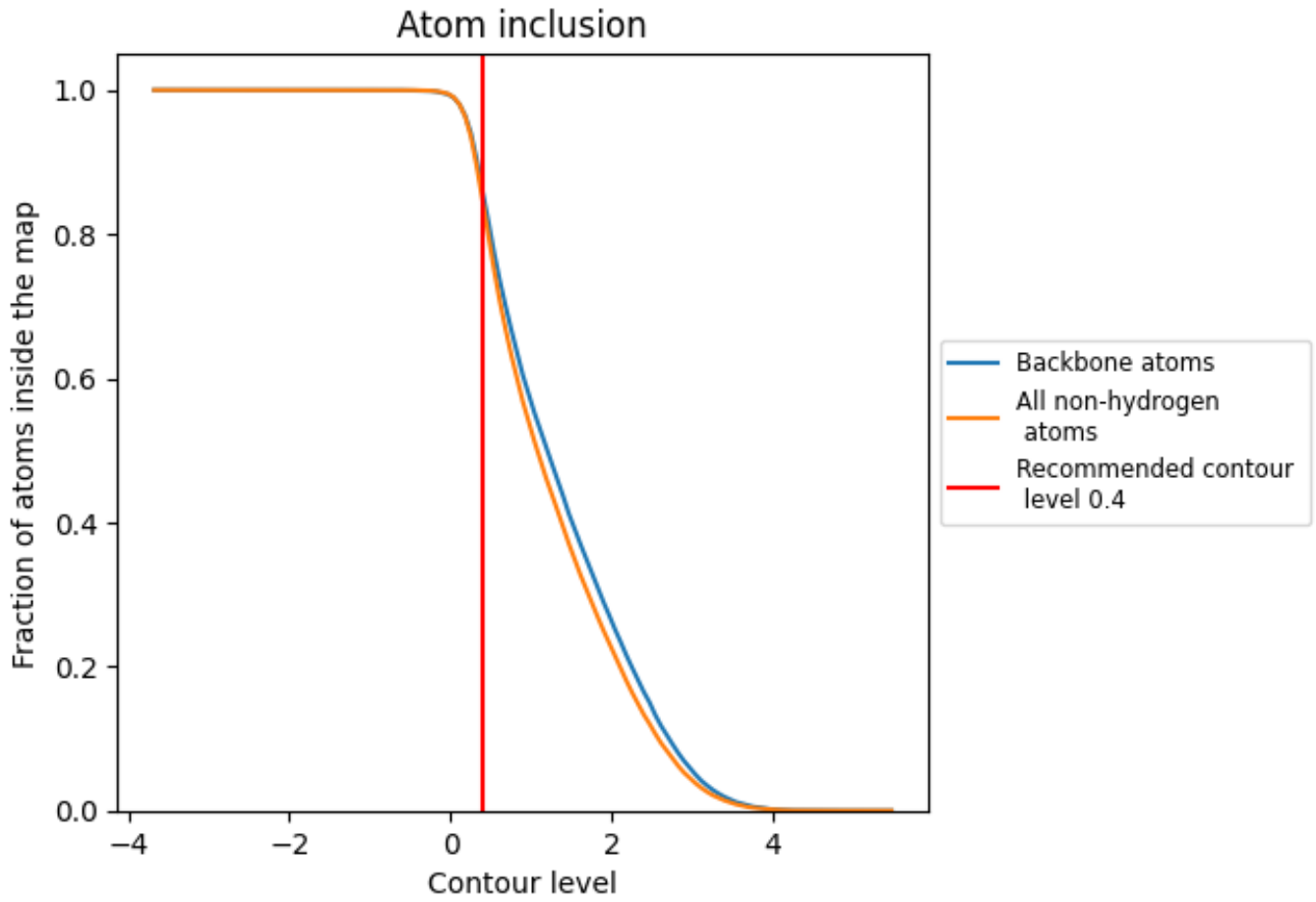
9.3 Atom inclusion mapped to coordinate model



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 (0.4).

9.4 Atom inclusion (i)



At the recommended contour level, 86% of all backbone atoms, 85% 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 (0.4) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8457	0.4890
A	0.9146	0.5540
B	0.8799	0.5240
C	0.9019	0.5320
D	0.8653	0.5070
E	0.7742	0.4890
F	0.6751	0.3660
X	0.7212	0.3270
Y	0.7196	0.3060