wwPDB EM Validation Summary Report (i)

Feb 27, 2024 – 12:52 AM EST

PDB ID	:	6VW0
EMDB ID	:	EMD-21409
Title	:	Mycobacterium tuberculosis RNAP S456L mutant open promoter complex
Authors	:	Lilic, M.; Boyaci, H.; Chen, J.; Darst, S.A.; Campbell, E.A.
Deposited on	:	2020-02-18
Resolution	:	3.59 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:

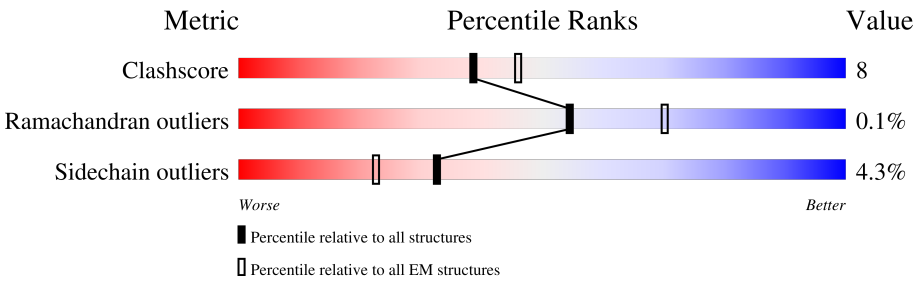
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.59 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	347	50% 15% 35%
1	B	347	53% 15% 32%
2	C	1179	73% 20% • 6%
3	D	1326	74% 20% • •
4	E	110	59% 15% • 25%
5	F	531	49% 11% • 39%
6	J	111	• 73% 24% • •
7	M	162	• 72% 23% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	O	90	57% 16% 28%
9	P	90	62% 9% 28%

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 29944 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	225	1716	1080	296	338	2	0	0
1	B	237	1765	1115	301	346	3	0	0

- 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	1111	8584	5382	1504	1659	39	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	456	LEU	SER	engineered mutation	UNP V9Z879
C	1179	LEU	-	expression tag	UNP V9Z879
C	1180	ALA	-	expression tag	UNP V9Z879
C	1181	ARG	-	expression tag	UNP V9Z879
C	1182	HIS	-	expression tag	UNP V9Z879
C	1183	GLY	-	expression tag	UNP V9Z879
C	1184	GLY	-	expression tag	UNP V9Z879
C	1185	SER	-	expression tag	UNP V9Z879

- 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	1270	9914	6208	1802	1862	42	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP A5U053

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	ALA	-	expression tag	UNP A5U053
D	1317	HIS	-	expression tag	UNP A5U053
D	1318	HIS	-	expression tag	UNP A5U053
D	1319	HIS	-	expression tag	UNP A5U053
D	1320	HIS	-	expression tag	UNP A5U053
D	1321	HIS	-	expression tag	UNP A5U053
D	1322	HIS	-	expression tag	UNP A5U053
D	1323	HIS	-	expression tag	UNP A5U053
D	1324	HIS	-	expression tag	UNP A5U053

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	83	649	414	108	127	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	GLY	-	expression tag	UNP A0A0T9N9K3

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	322	2527	1576	459	483	9	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	expression tag	UNP P9WGI0
F	-1	PRO	-	expression tag	UNP P9WGI0
F	0	HIS	-	expression tag	UNP P9WGI0

- Molecule 6 is a protein called RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	J	109	880	543	166	168	3	0	0

- Molecule 7 is a protein called RNA polymerase-binding transcription factor CarD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	M	159	1241	777	224	239	1	0	0

- Molecule 8 is a DNA chain called DNA (65-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	O	65	1336	633	243	395	65	0	0

- Molecule 9 is a DNA chain called DNA (65-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
9	P	65	1329	629	250	385	65	0	0

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

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

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

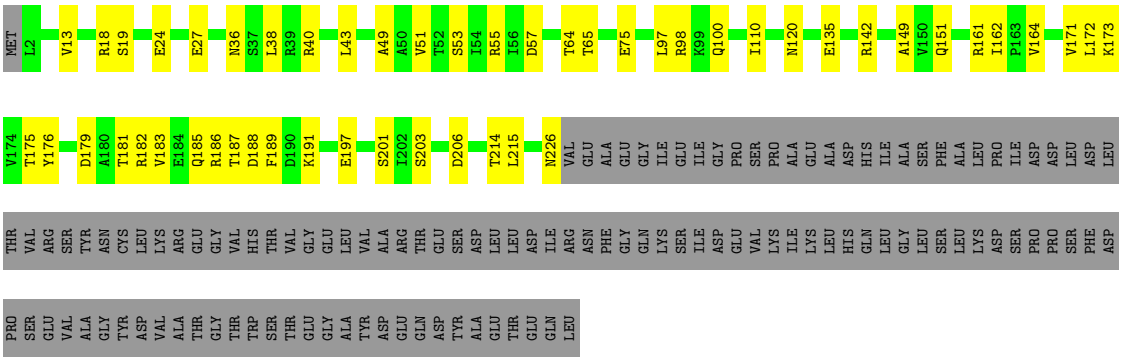
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
11	D	1	1	1	0

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.

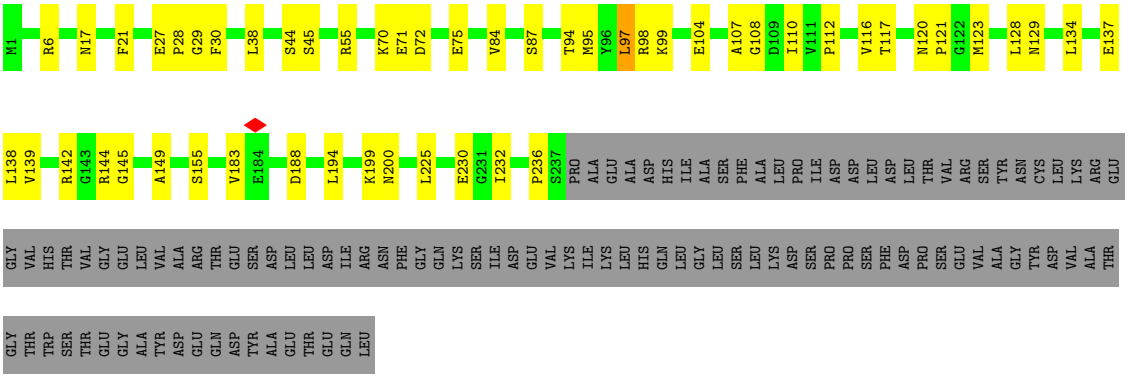
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A: 50% 15% 35%



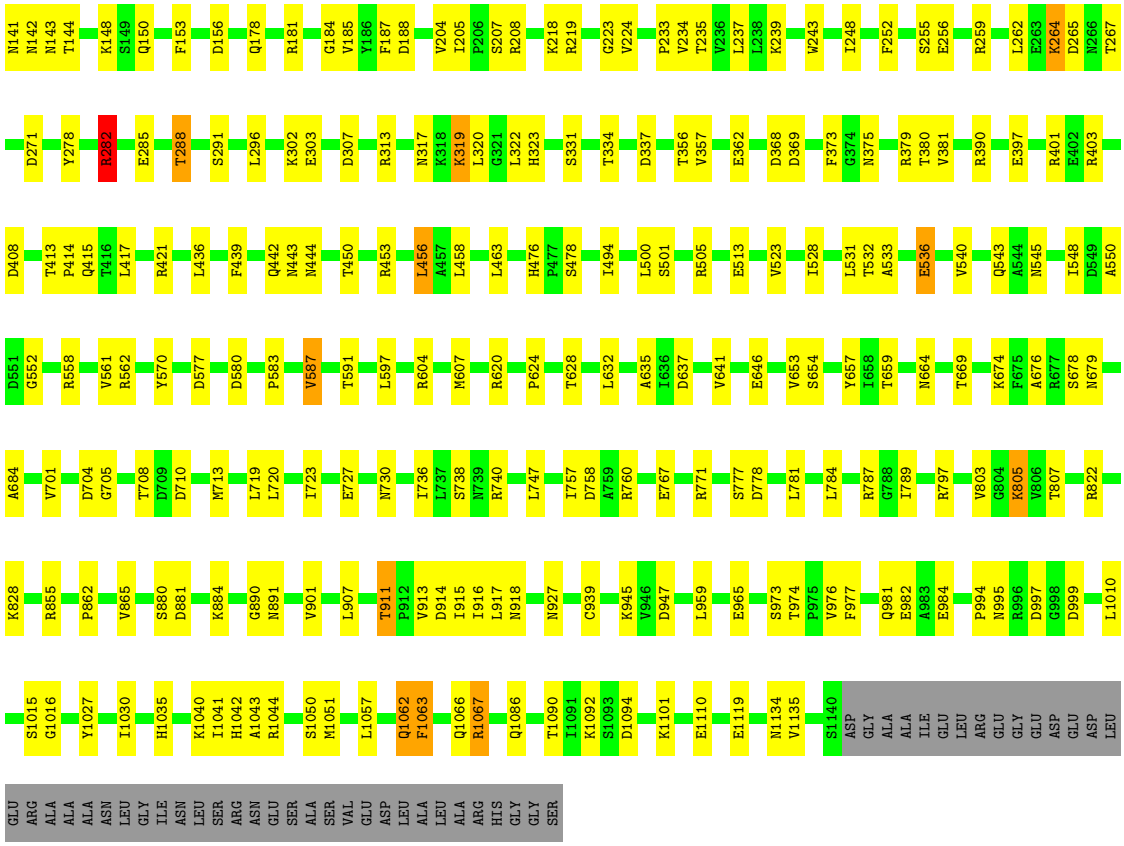
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B: 53% 15% 32%



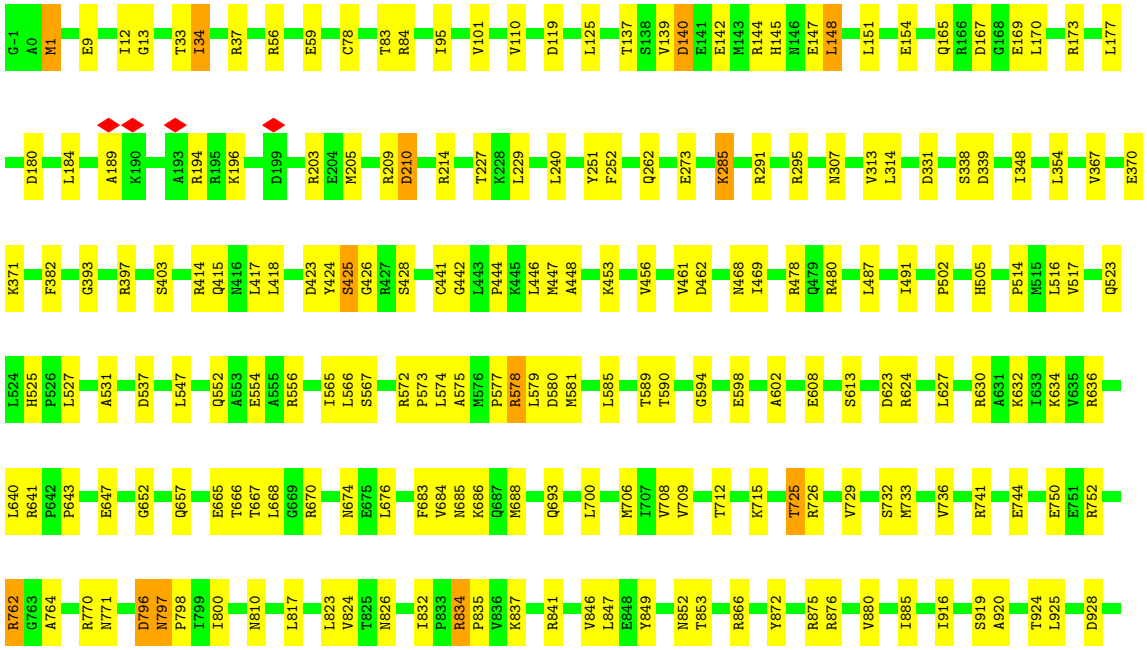
• Molecule 2: DNA-directed RNA polymerase subunit beta

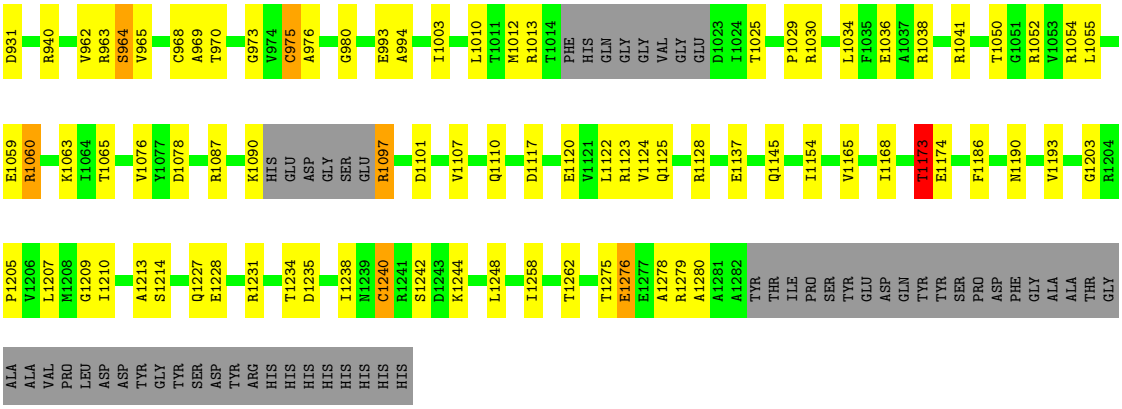
Chain C: 73% 20% • 6%



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

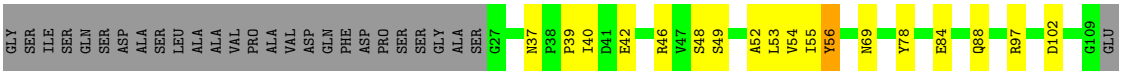
Chain D:





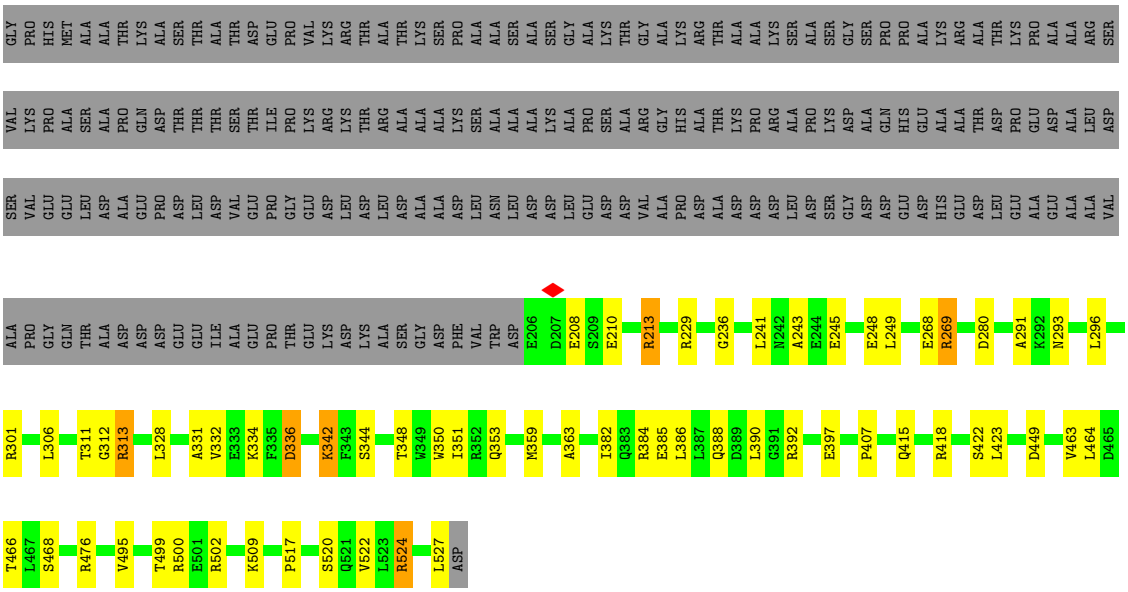
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:



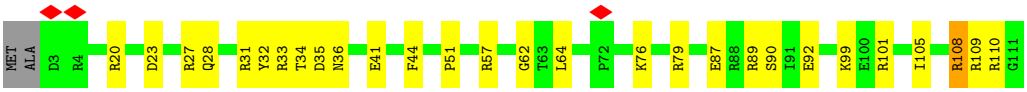
- Molecule 5: RNA polymerase sigma factor SigA

Chain F:



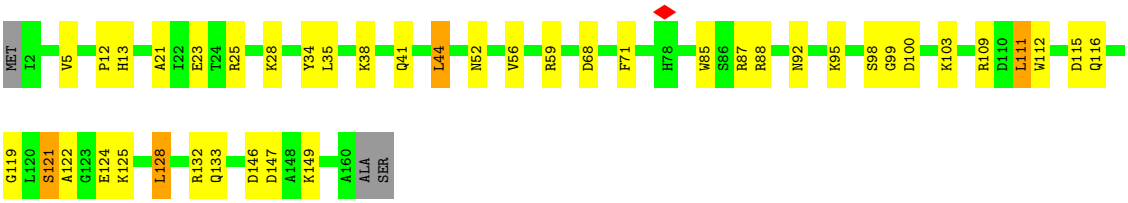
- Molecule 6: RNA polymerase-binding protein RbpA

Chain J:



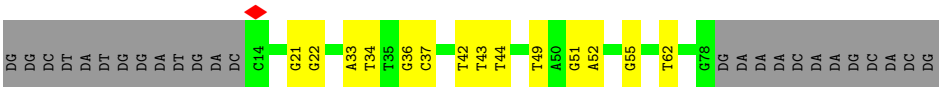
- Molecule 7: RNA polymerase-binding transcription factor CarD

Chain M:



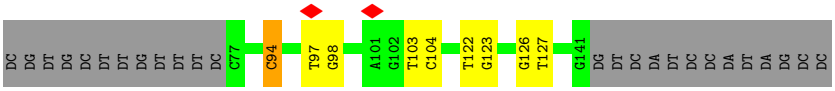
- Molecule 8: DNA (65-MER)

Chain O:



- Molecule 9: DNA (65-MER)

Chain P:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	55833	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 ($e^-/Å^2$)	71	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	2.591	Depositor
Minimum map value	-1.047	Depositor
Average map value	0.004	Depositor
Map value standard deviation	0.100	Depositor
Recommended contour level	0.387	Depositor
Map size (Å)	332.8, 332.8, 332.8	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.3, 1.3, 1.3	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: MG, 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	0.49	0/1742	0.57	0/2370
1	B	0.47	0/1792	0.55	0/2442
2	C	0.52	1/8743 (0.0%)	0.58	0/11860
3	D	0.51	0/10078	0.58	2/13624 (0.0%)
4	E	0.51	0/662	0.55	0/901
5	F	0.39	0/2558	0.50	0/3451
6	J	0.36	0/896	0.51	0/1210
7	M	0.32	0/1257	0.55	0/1700
8	O	0.89	0/1497	1.04	1/2310 (0.0%)
9	P	0.88	0/1491	1.02	1/2297 (0.0%)
All	All	0.54	1/30716 (0.0%)	0.63	4/42165 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1
3	D	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	723	ILE	C-N	-5.55	1.21	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	834	ARG	C-N-CD	-5.67	108.12	120.60
3	D	140	ASP	CB-CG-OD2	5.16	122.94	118.30
8	O	42	DT	P-O3'-C3'	5.14	125.87	119.70
9	P	94	DC	O4'-C4'-C3'	-5.08	102.47	104.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	282	ARG	Peptide
3	D	1173	THR	Peptide

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	1716	0	1756	33	0
1	B	1765	0	1794	31	0
2	C	8584	0	8509	156	0
3	D	9914	0	9986	177	0
4	E	649	0	645	10	0
5	F	2527	0	2535	39	0
6	J	880	0	852	18	0
7	M	1241	0	1259	26	0
8	O	1336	0	732	10	0
9	P	1329	0	727	7	0
10	D	2	0	0	0	0
11	D	1	0	0	0	0
All	All	29944	0	28795	463	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 8.

The worst 5 of 463 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:144:ARG:O	3:D:148:LEU:HB2	1.79	0.82

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:M:112:TRP:O	7:M:116:GLN:HB2	1.88	0.74
3:D:880:VAL:HG21	3:D:1210:ILE:HB	1.71	0.73
3:D:641:ARG:HA	3:D:657:GLN:HG3	1.70	0.72
5:F:386:LEU:O	5:F:390:LEU:HB3	1.94	0.68

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	223/347 (64%)	202 (91%)	21 (9%)	0	100	100
1	B	235/347 (68%)	204 (87%)	31 (13%)	0	100	100
2	C	1109/1179 (94%)	1007 (91%)	101 (9%)	1 (0%)	51	83
3	D	1264/1326 (95%)	1174 (93%)	89 (7%)	1 (0%)	51	83
4	E	81/110 (74%)	78 (96%)	3 (4%)	0	100	100
5	F	320/531 (60%)	309 (97%)	11 (3%)	0	100	100
6	J	107/111 (96%)	96 (90%)	11 (10%)	0	100	100
7	M	157/162 (97%)	148 (94%)	9 (6%)	0	100	100
All	All	3496/4113 (85%)	3218 (92%)	276 (8%)	2 (0%)	54	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	1276	GLU
2	C	982	GLU

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	194/297 (65%)	188 (97%)	6 (3%)	40	71
1	B	195/297 (66%)	191 (98%)	4 (2%)	53	78
2	C	932/997 (94%)	891 (96%)	41 (4%)	28	63
3	D	1048/1103 (95%)	1004 (96%)	44 (4%)	30	63
4	E	69/89 (78%)	65 (94%)	4 (6%)	20	55
5	F	262/429 (61%)	249 (95%)	13 (5%)	24	59
6	J	92/97 (95%)	87 (95%)	5 (5%)	22	57
7	M	129/131 (98%)	119 (92%)	10 (8%)	12	44
All	All	2921/3440 (85%)	2794 (96%)	127 (4%)	33	63

5 of 127 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	209	ARG
5	F	522	VAL
3	D	725	THR
5	F	509	LYS
7	M	35	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	351	ASN
5	F	461	GLN
3	D	494	HIS
3	D	1227	GLN
3	D	416	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.

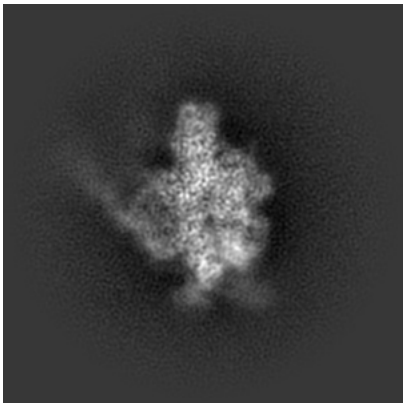
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-21409. These allow visual inspection of the internal detail of the map and identification of artifacts.

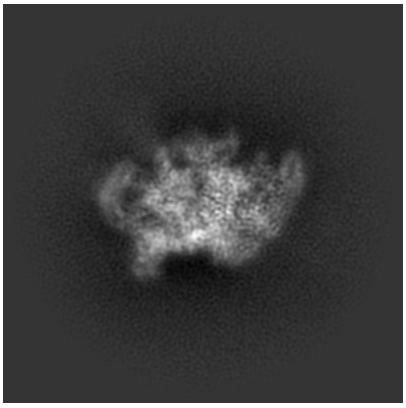
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

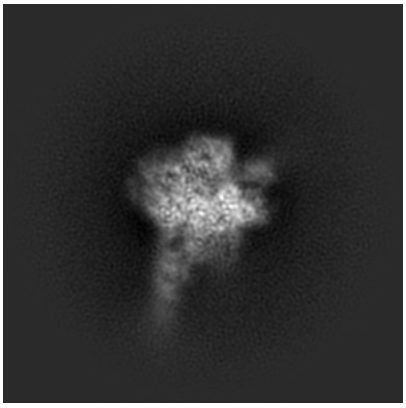
6.1.1 Primary map



X



Y

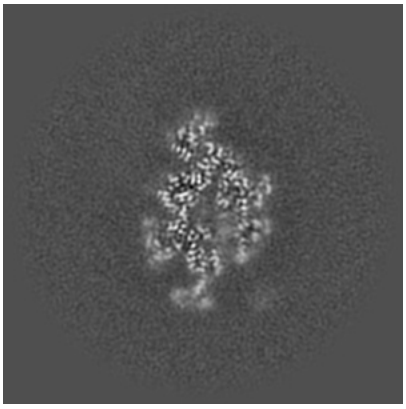


Z

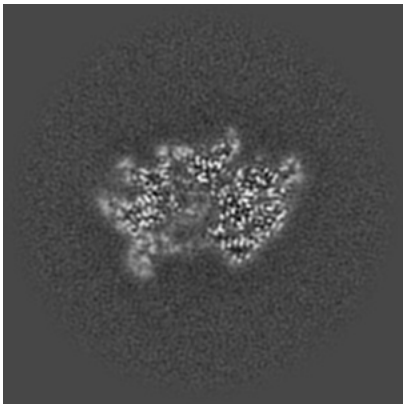
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

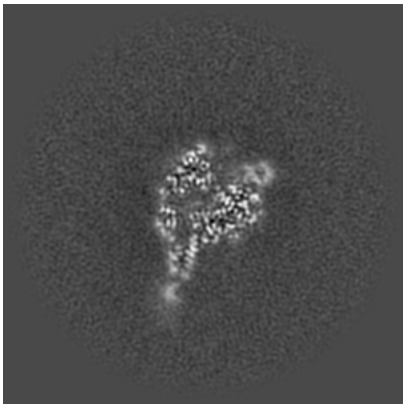
6.2.1 Primary map



X Index: 128



Y Index: 128

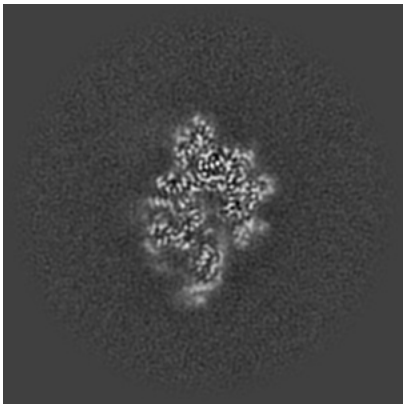


Z Index: 128

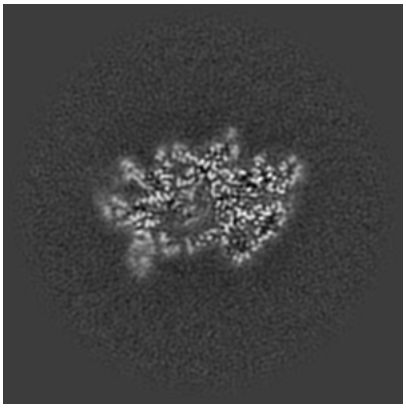
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

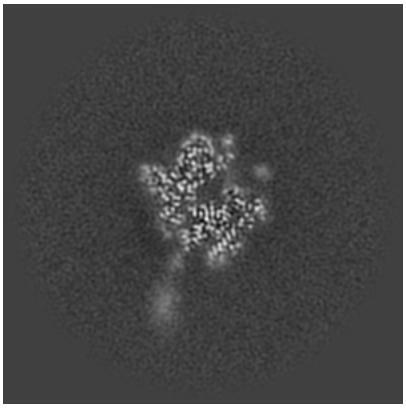
6.3.1 Primary map



X Index: 122



Y Index: 126

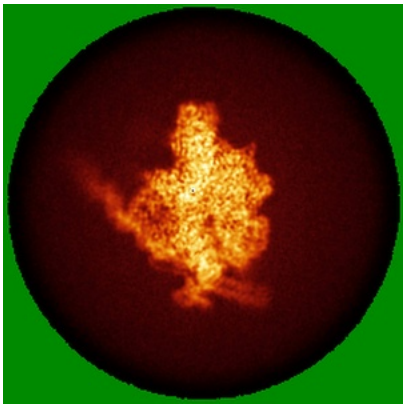


Z Index: 137

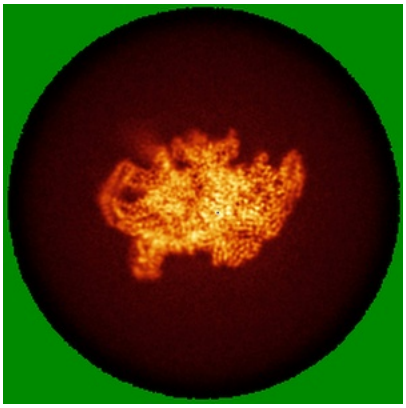
The images above show the largest variance slices of the map in three orthogonal directions.

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

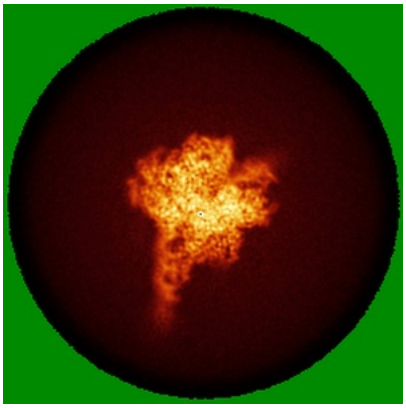
6.4.1 Primary map



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

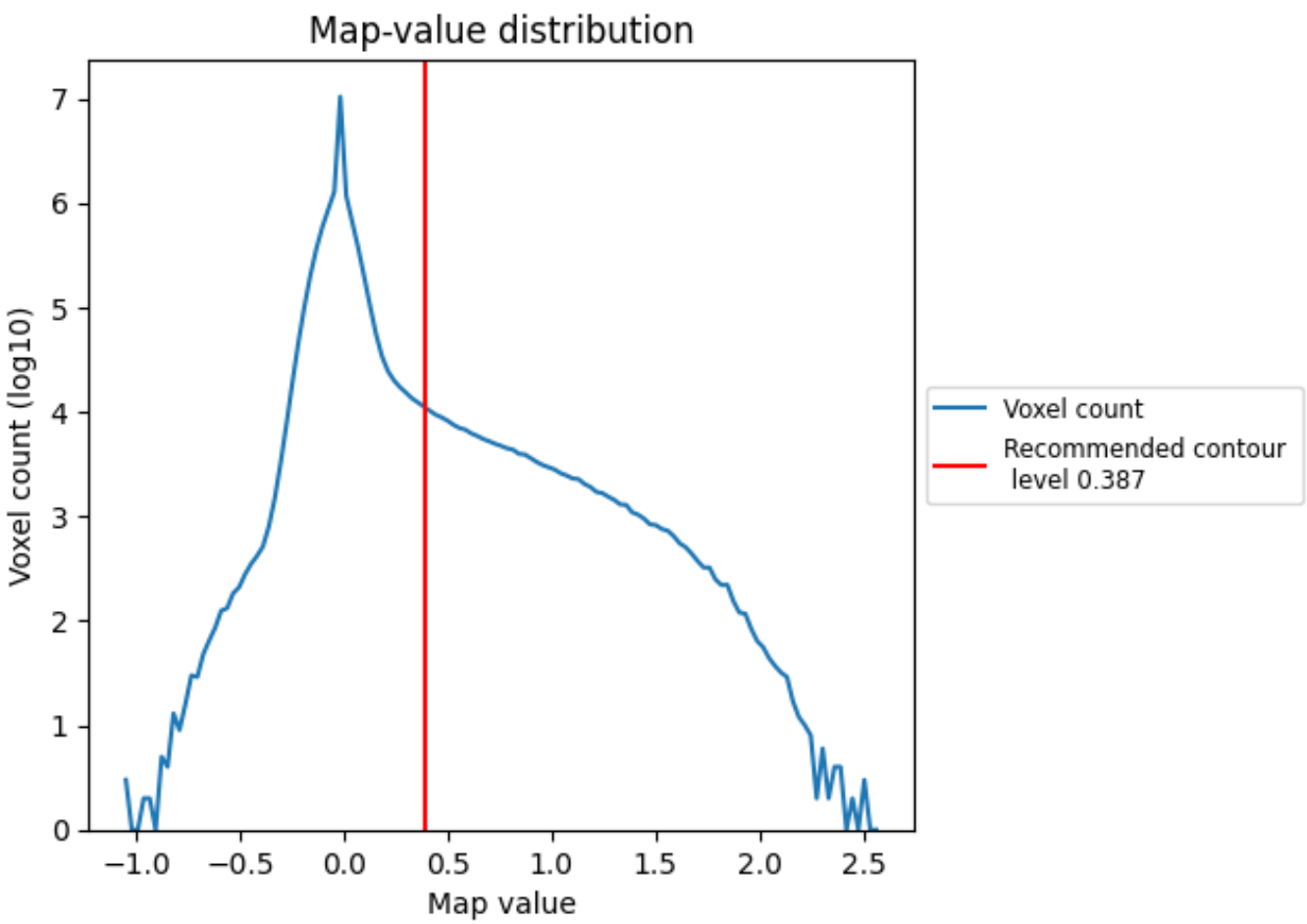
6.6 Mask visualisation (i)

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

7 Map analysis (i)

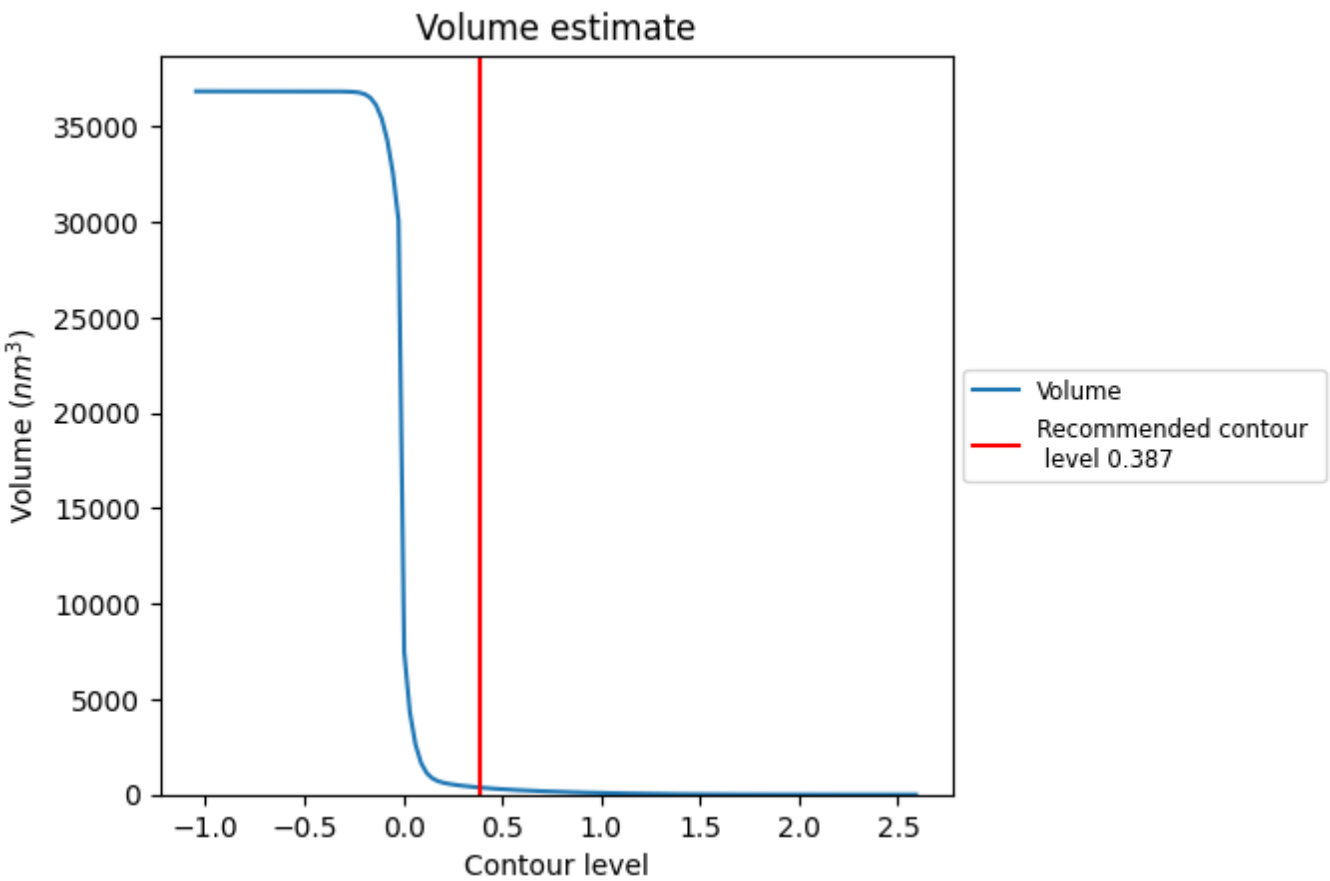
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



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

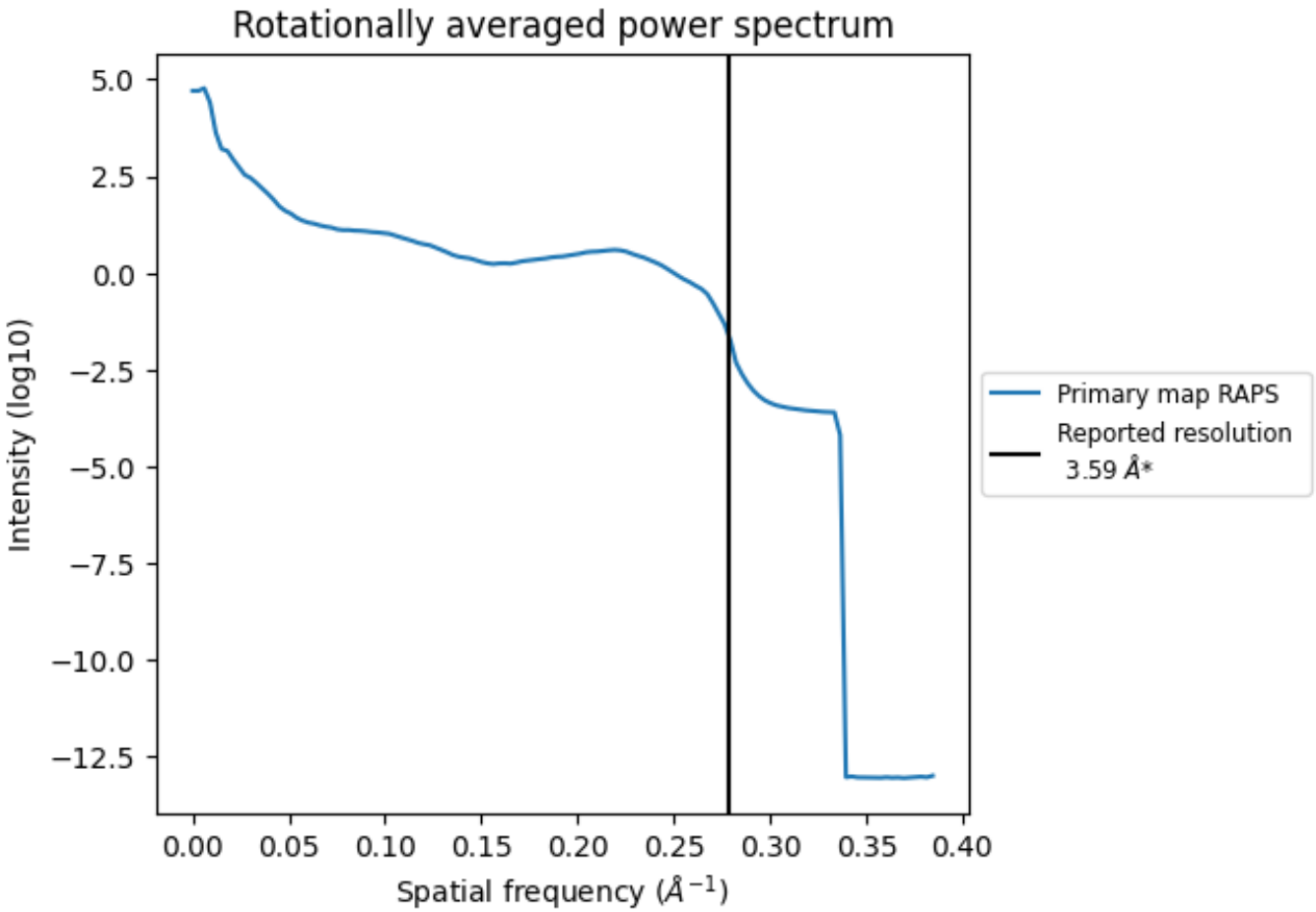
7.2 Volume estimate (i)



The volume at the recommended contour level is 367 nm^3; this corresponds to an approximate mass of 332 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 0.279 Å^{-1}

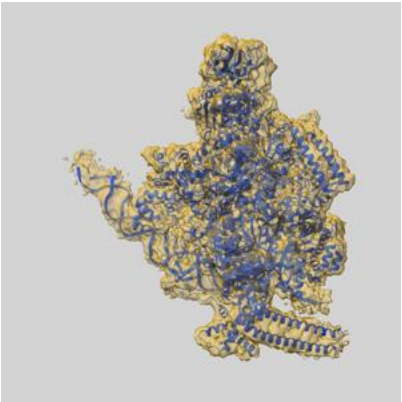
8 Fourier-Shell correlation (i)

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

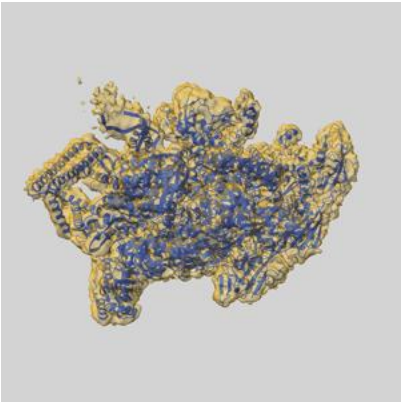
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-21409 and PDB model 6VW0. Per-residue inclusion information can be found in section 3 on page 7.

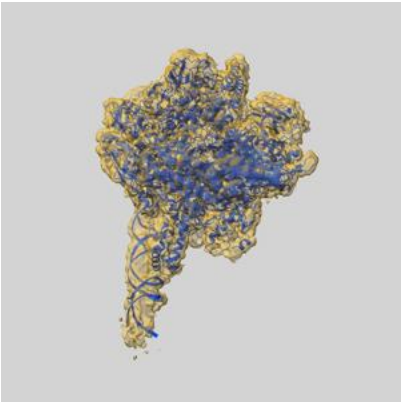
9.1 Map-model overlay (i)



X



Y



Z

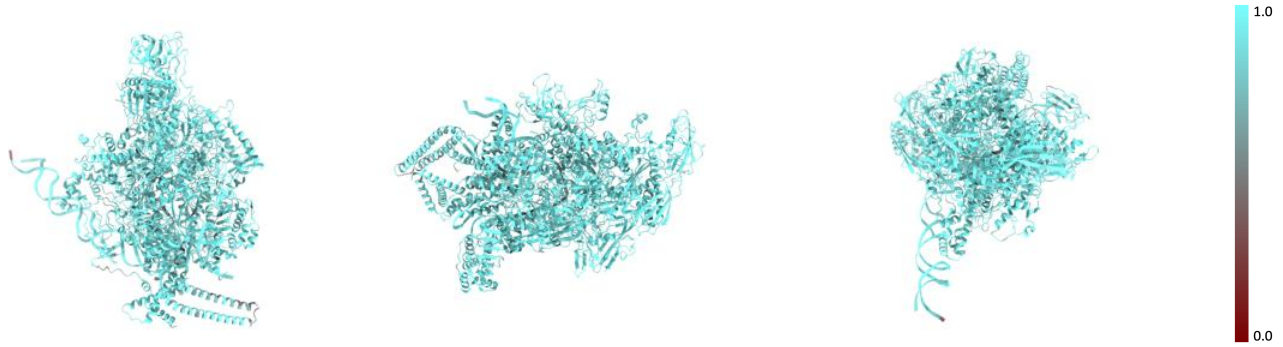
The images above show the 3D surface view of the map at the recommended contour level 0.387 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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.

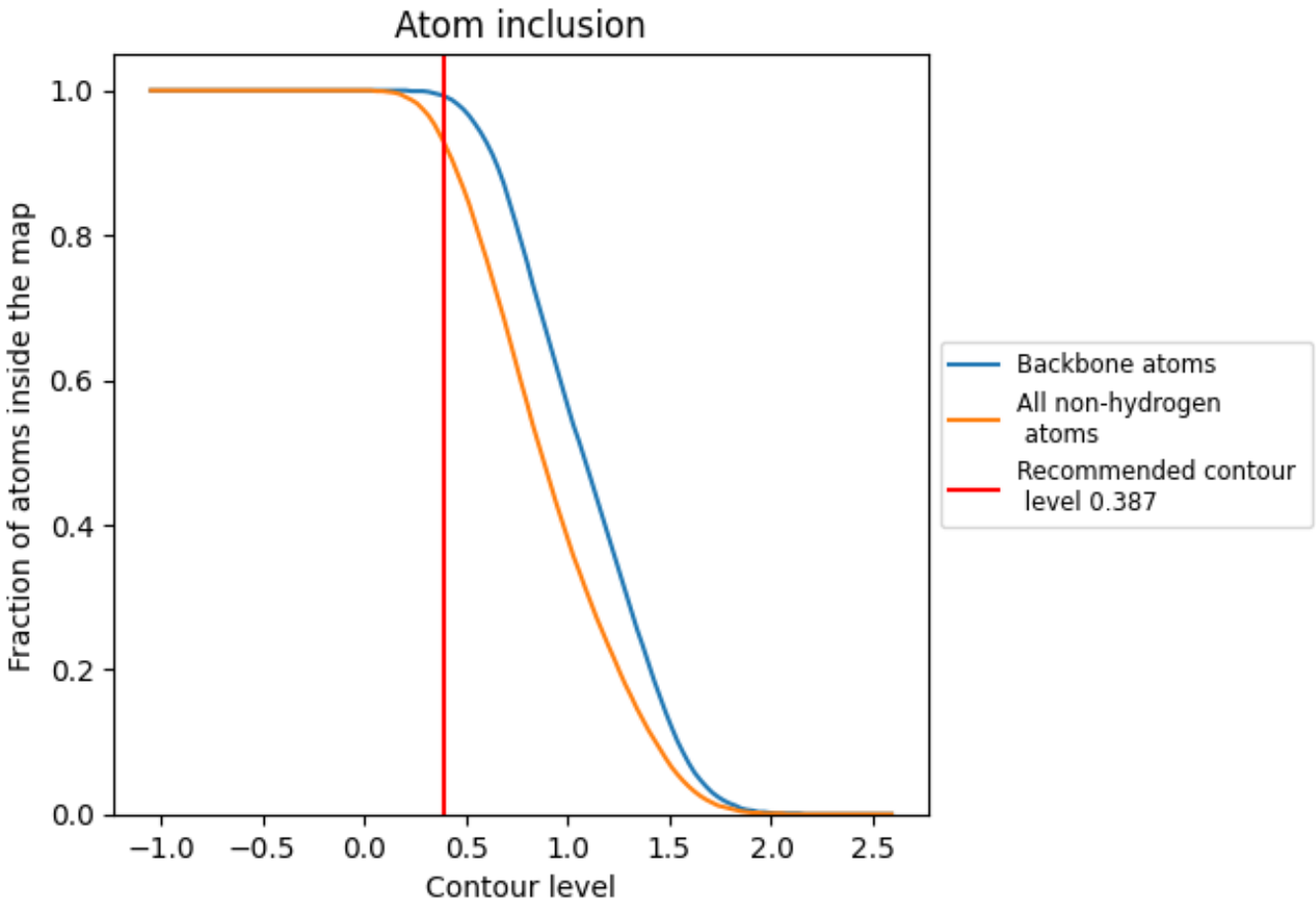
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.387).

9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 93% 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 (0.387) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9310	0.4200
A	0.9580	0.4650
B	0.9400	0.4320
C	0.9420	0.4560
D	0.9290	0.4360
E	0.9430	0.4620
F	0.9100	0.3720
J	0.8630	0.3690
M	0.8980	0.3230
O	0.9520	0.3080
P	0.9030	0.3050