wwPDB EM Validation Summary Report

Nov 9, 2026 – 12:07 pm GMT

PDB ID	:	29OC / pdb_000029oc
EMDB ID	:	EMD-57262
Title	:	Cryo-EM structure of human fibrillar procollagen type I C-propeptide heterotrimer alpha1-K18A mutant complexed with PCPE-1
Authors	:	Carrique, L.; Lipinski, O.; Vadon-Le Goff, S.; Moali, C.
Deposited on	:	2026-03-26
Resolution	:	3.01 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:

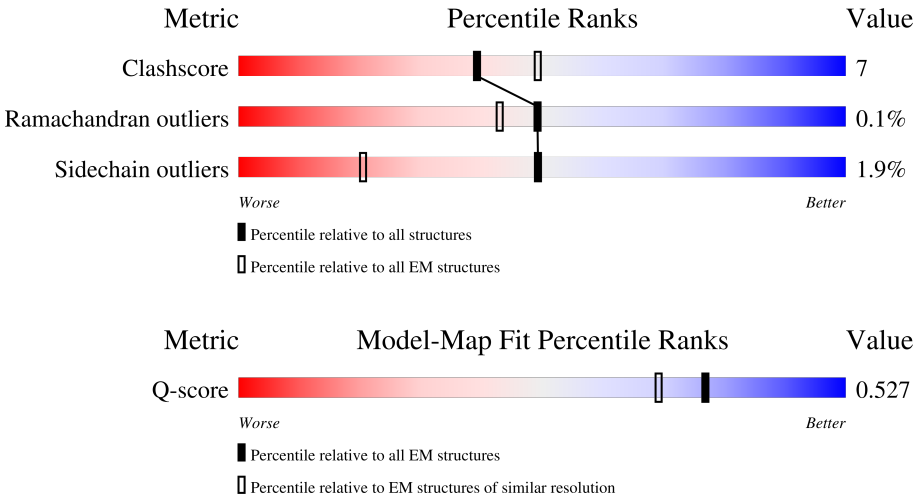
EMDB validation analysis	:	0.0.1.dev133
MolProbity	:	4-5-2 with Phenix2.2
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

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 3.01 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	13882 (2.51 - 3.51)

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	383	51% 12% 36%
1	B	383	53% 10% 36%
2	C	400	45% 14% 41%
3	D	433	48% 8% 44%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7569 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 Collagen alpha-1(I) chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	245	Total	C	N	O	S	0	0
			1920	1195	331	381	13		
1	B	245	Total	C	N	O	S	0	0
			1920	1195	331	381	13		

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-136	GLU	-	expression tag	UNP P02452
A	-135	THR	-	expression tag	UNP P02452
A	-134	GLY	-	expression tag	UNP P02452
A	-131	PRO	ALA	conflict	UNP P02452
A	-128	PRO	GLN	conflict	UNP P02452
A	-125	PRO	ARG	conflict	UNP P02452
A	18	ALA	LYS	conflict	UNP P02452
A	147	GLN	ASN	conflict	UNP P02452
A	216	SER	THR	variant	UNP P02452
B	-136	GLU	-	expression tag	UNP P02452
B	-135	THR	-	expression tag	UNP P02452
B	-134	GLY	-	expression tag	UNP P02452
B	-131	PRO	ALA	conflict	UNP P02452
B	-128	PRO	GLN	conflict	UNP P02452
B	-125	PRO	ARG	conflict	UNP P02452
B	18	ALA	LYS	conflict	UNP P02452
B	147	GLN	ASN	conflict	UNP P02452
B	216	SER	THR	variant	UNP P02452

- Molecule 2 is a protein called Collagen alpha-2(I) chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	236	Total	C	N	O	S	0	0
			1874	1180	316	367	11		

There are 43 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-152	GLU	-	expression tag	UNP P08123
C	-151	THR	-	expression tag	UNP P08123
C	-150	GLY	-	expression tag	UNP P08123
C	-149	TRP	-	expression tag	UNP P08123
C	-148	SER	-	expression tag	UNP P08123
C	-147	HIS	-	expression tag	UNP P08123
C	-146	PRO	-	expression tag	UNP P08123
C	-145	GLN	-	expression tag	UNP P08123
C	-144	PHE	-	expression tag	UNP P08123
C	-143	GLU	-	expression tag	UNP P08123
C	-142	LYS	-	expression tag	UNP P08123
C	-141	GLY	-	expression tag	UNP P08123
C	-140	GLY	-	expression tag	UNP P08123
C	-139	GLY	-	expression tag	UNP P08123
C	-138	SER	-	expression tag	UNP P08123
C	-137	GLY	-	expression tag	UNP P08123
C	-136	GLY	-	expression tag	UNP P08123
C	-135	GLY	-	expression tag	UNP P08123
C	-134	SER	-	expression tag	UNP P08123
C	-133	GLY	-	expression tag	UNP P08123
C	-132	GLY	-	expression tag	UNP P08123
C	-131	SER	-	expression tag	UNP P08123
C	-130	ALA	-	expression tag	UNP P08123
C	-129	TRP	-	expression tag	UNP P08123
C	-128	SER	-	expression tag	UNP P08123
C	-127	HIS	-	expression tag	UNP P08123
C	-126	PRO	-	expression tag	UNP P08123
C	-125	GLN	-	expression tag	UNP P08123
C	-124	PHE	-	expression tag	UNP P08123
C	-123	GLU	-	expression tag	UNP P08123
C	-122	LYS	-	expression tag	UNP P08123
C	-121	SER	-	expression tag	UNP P08123
C	-120	ALA	-	expression tag	UNP P08123
C	-119	GLY	-	expression tag	UNP P08123
C	-118	PRO	-	expression tag	UNP P08123
C	-117	PRO	-	expression tag	UNP P08123
C	-116	GLY	-	expression tag	UNP P08123
C	-115	PRO	-	expression tag	UNP P08123
C	-114	PRO	-	expression tag	UNP P08123
C	-113	GLY	-	expression tag	UNP P08123
C	-112	PRO	-	expression tag	UNP P08123
C	-111	PRO	-	expression tag	UNP P08123

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	148	GLN	ASN	conflict	UNP P08123

- Molecule 3 is a protein called Procollagen C-endopeptidase enhancer 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	243	Total	C	N	O	S	0	0
			1850	1170	311	360	9		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	425	HIS	-	expression tag	UNP Q15113
D	426	HIS	-	expression tag	UNP Q15113
D	427	HIS	-	expression tag	UNP Q15113
D	428	HIS	-	expression tag	UNP Q15113
D	429	HIS	-	expression tag	UNP Q15113
D	430	HIS	-	expression tag	UNP Q15113
D	431	HIS	-	expression tag	UNP Q15113
D	432	HIS	-	expression tag	UNP Q15113
D	433	GLY	-	expression tag	UNP Q15113

- Molecule 4 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

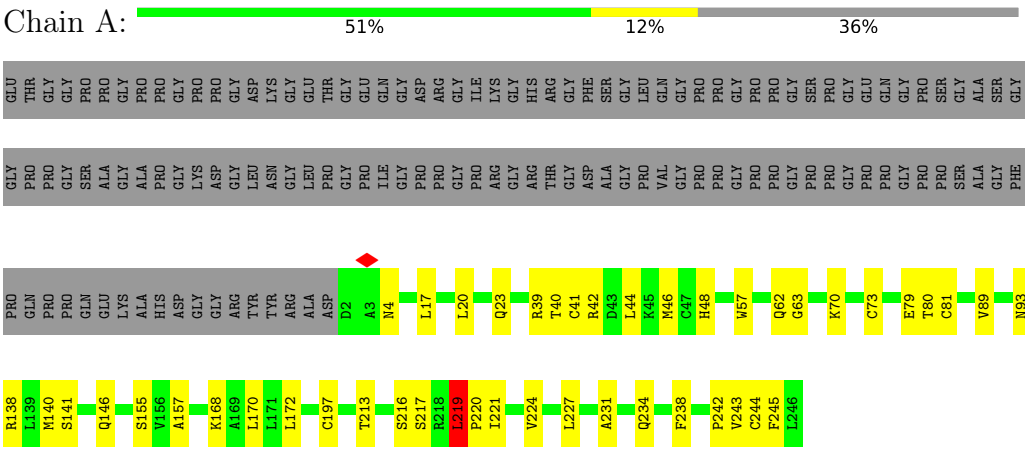
Mol	Chain	Residues	Atoms		AltConf
4	A	1	Total	Ca	0
			1	1	
4	B	1	Total	Ca	0
			1	1	
4	C	1	Total	Ca	0
			1	1	
4	D	2	Total	Ca	0
			2	2	

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.

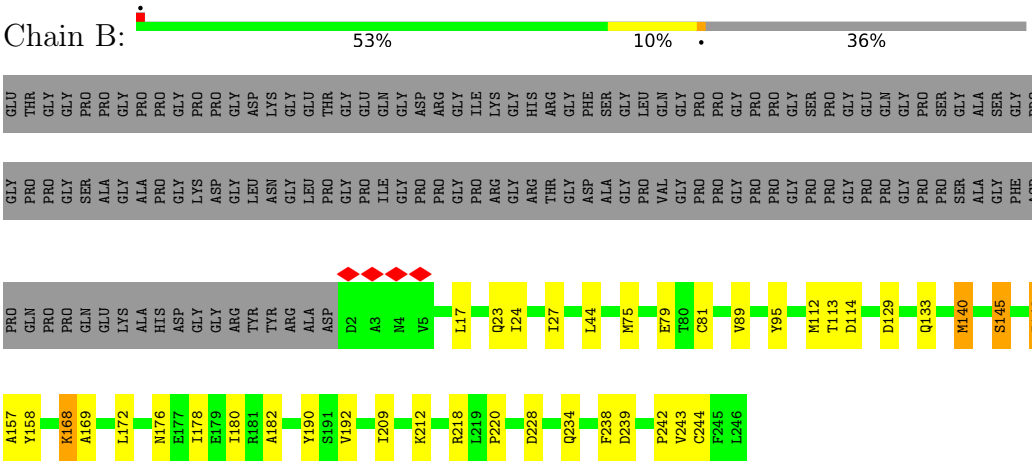
• Molecule 1: Collagen alpha-1(I) chain

Chain A:



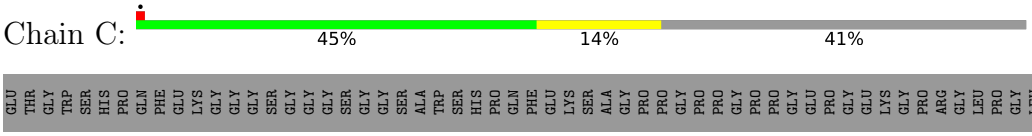
• Molecule 1: Collagen alpha-1(I) chain

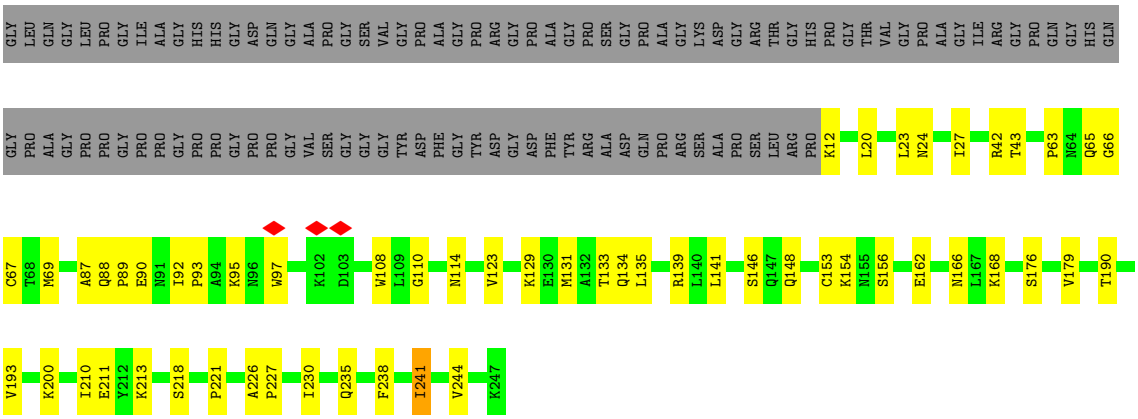
Chain B:



• Molecule 2: Collagen alpha-2(I) chain

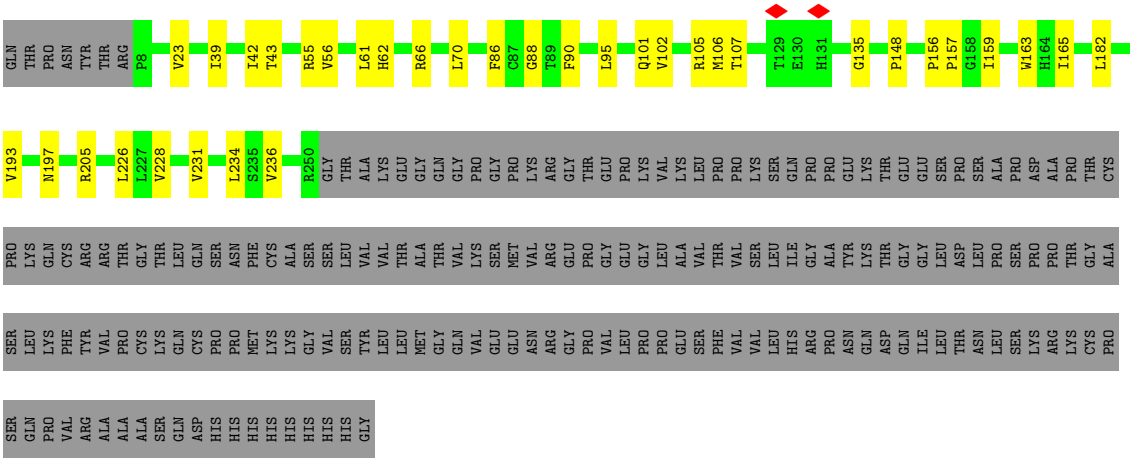
Chain C:





• Molecule 3: Procollagen C-endopeptidase enhancer 1





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	586000	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI/PHILIPS CM300FEG/HE	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	9.26	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	157.920	Depositor
Minimum map value	-119.422	Depositor
Average map value	0.001	Depositor
Map value standard deviation	3.503	Depositor
Recommended contour level	6.9	Depositor
Map size (Å)	219.08992, 219.08992, 219.08992	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.85582, 0.85582, 0.85582	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.14	0/1961	0.34	0/2657
1	B	0.13	0/1961	0.31	0/2657
2	C	0.13	0/1915	0.29	0/2593
3	D	0.12	0/1901	0.31	0/2588
All	All	0.13	0/7738	0.32	0/10495

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
1	A	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
1	A	219	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	1920	0	1837	34	0
1	B	1920	0	1837	28	0
2	C	1874	0	1815	29	0
3	D	1850	0	1754	20	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
All	All	7569	0	7243	106	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 7.

The worst 5 of 106 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:172:LEU:HD23	1:A:221:ILE:HG21	1.70	0.72
1:A:168:LYS:HA	1:A:227:LEU:HD23	1.76	0.68
1:A:138:ARG:HH21	1:A:220:PRO:HG2	1.58	0.68
1:A:213:THR:HG22	1:A:219:LEU:HD12	1.75	0.68
3:D:23:VAL:HG21	3:D:42:ILE:HD11	1.76	0.67

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	243/383 (63%)	231 (95%)	12 (5%)	0	100	100
1	B	243/383 (63%)	233 (96%)	10 (4%)	0	100	100
2	C	234/400 (58%)	227 (97%)	7 (3%)	0	100	100
3	D	241/433 (56%)	227 (94%)	13 (5%)	1 (0%)	30	63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	961/1599 (60%)	918 (96%)	42 (4%)	1 (0%)	49 79

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	236	VAL

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	213/303 (70%)	208 (98%)	5 (2%)	44 72
1	B	213/303 (70%)	209 (98%)	4 (2%)	50 75
2	C	206/310 (66%)	202 (98%)	4 (2%)	50 75
3	D	201/364 (55%)	198 (98%)	3 (2%)	57 79
All	All	833/1280 (65%)	817 (98%)	16 (2%)	49 75

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	228	VAL
3	D	107	THR
1	B	168	LYS
2	C	241	ILE
1	B	148	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	161	GLN
1	B	173	GLN
2	C	143	ASN
1	B	176	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	165	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 5 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.

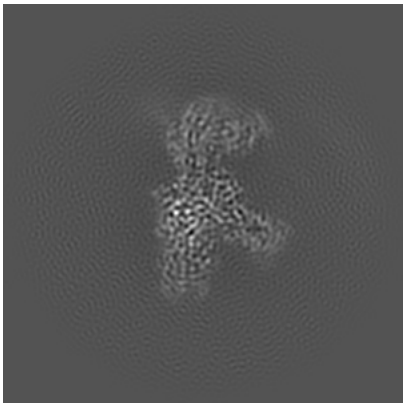
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-57262. These allow visual inspection of the internal detail of the map and identification of artifacts.

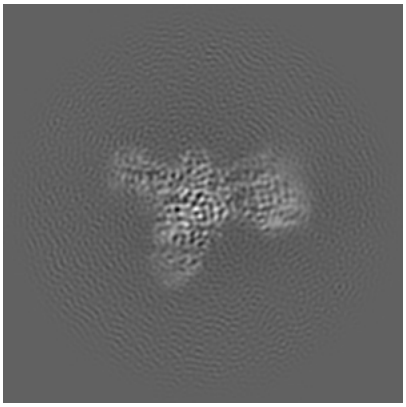
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

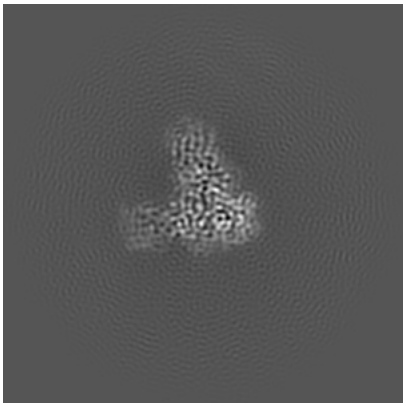
6.1.1 Primary map



X



Y

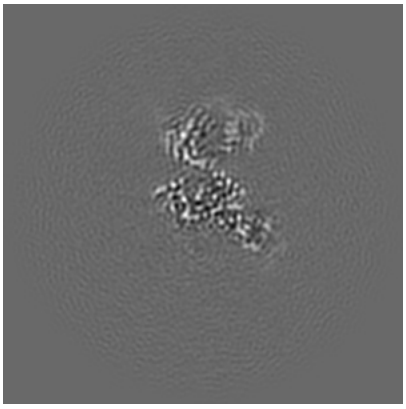


Z

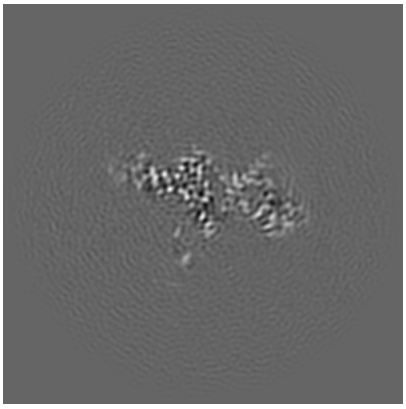
The images above show the map projected in three orthogonal directions.

6.2 Central slices

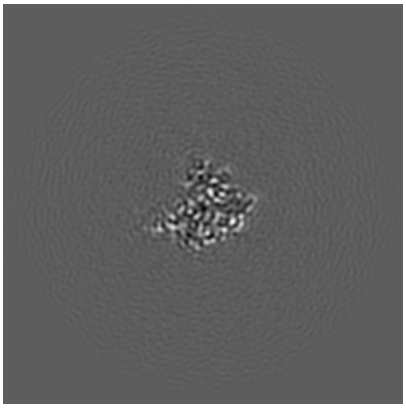
6.2.1 Primary map



X Index: 128



Y Index: 128

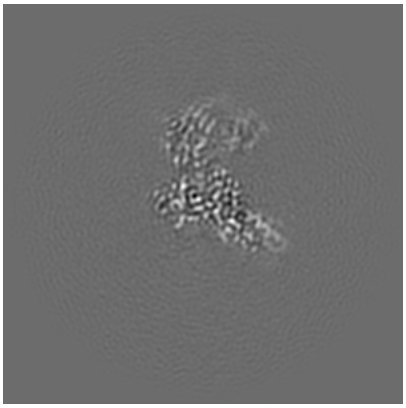


Z Index: 128

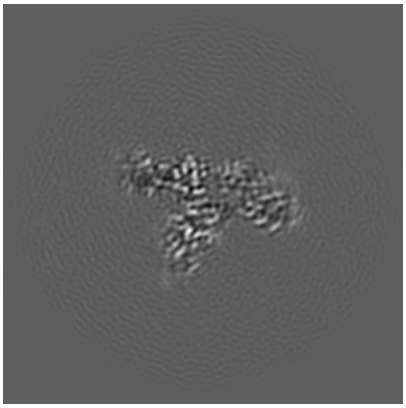
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

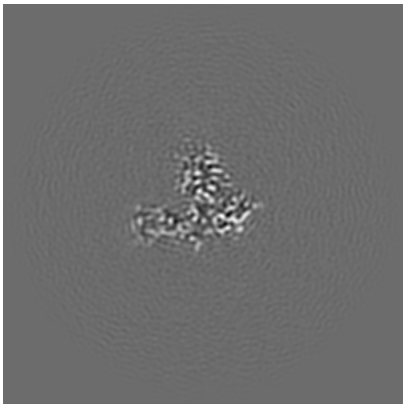
6.3.1 Primary map



X Index: 126



Y Index: 118

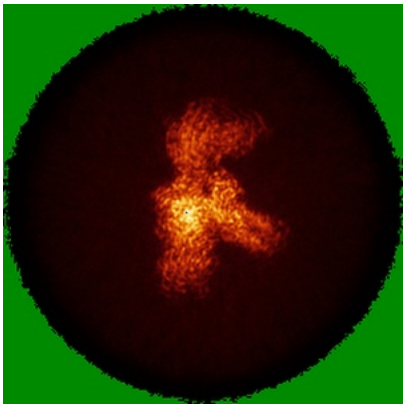


Z Index: 122

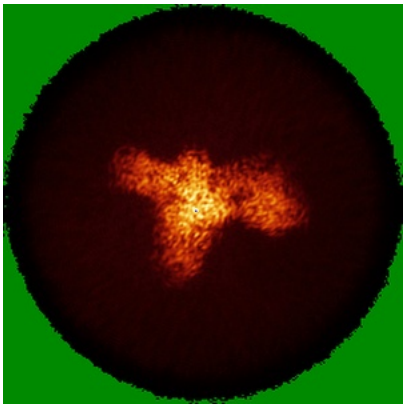
The images above show the largest variance slices of the map in three orthogonal directions.

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

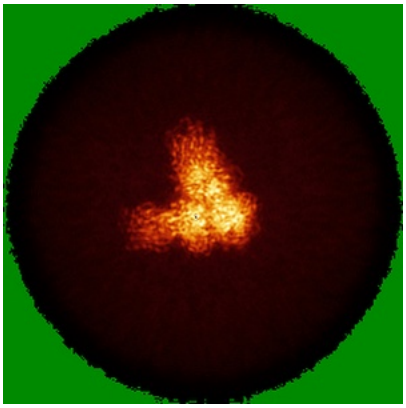
6.4.1 Primary map



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

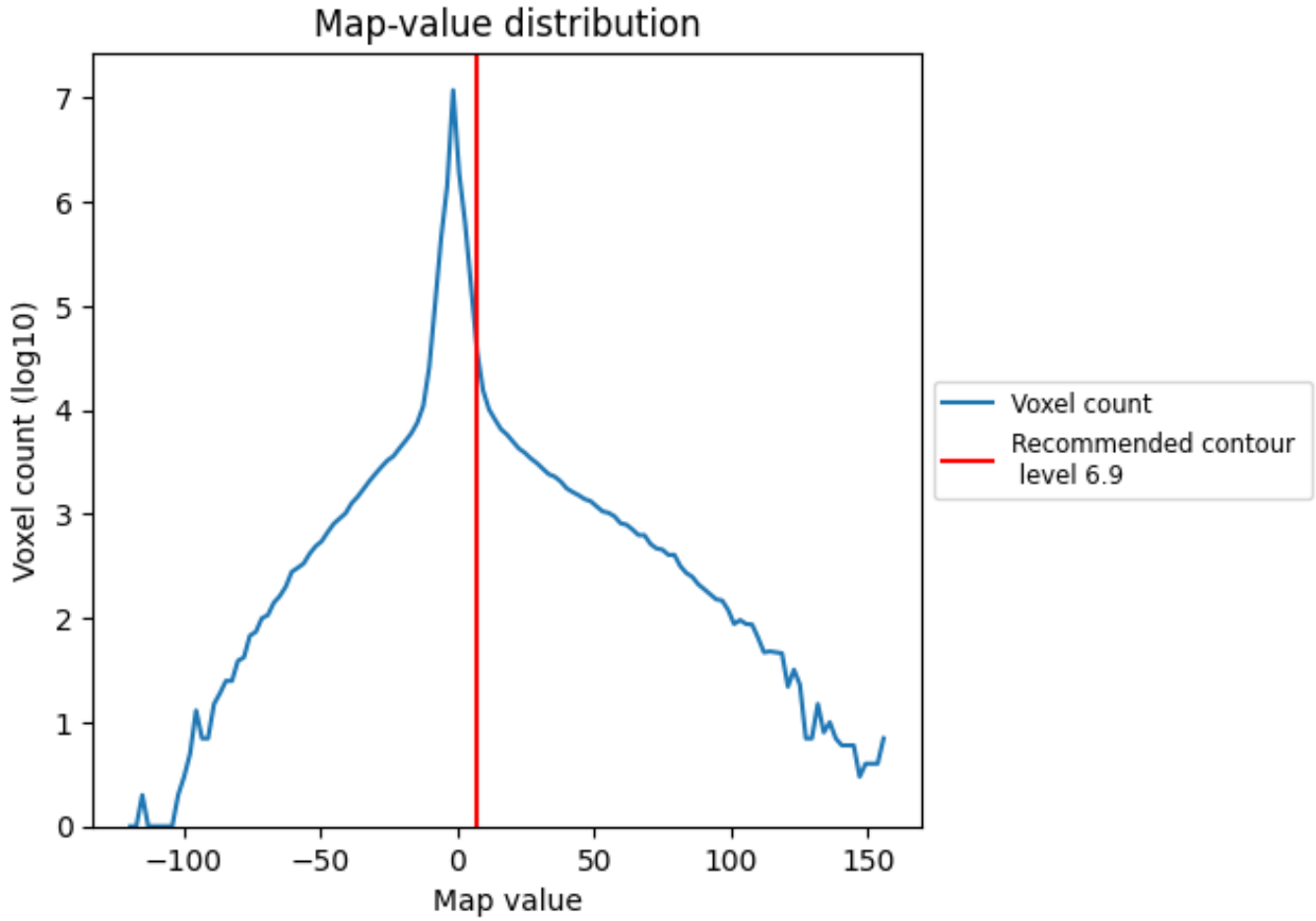
6.6 Mask visualisation (i)

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

7 Map analysis

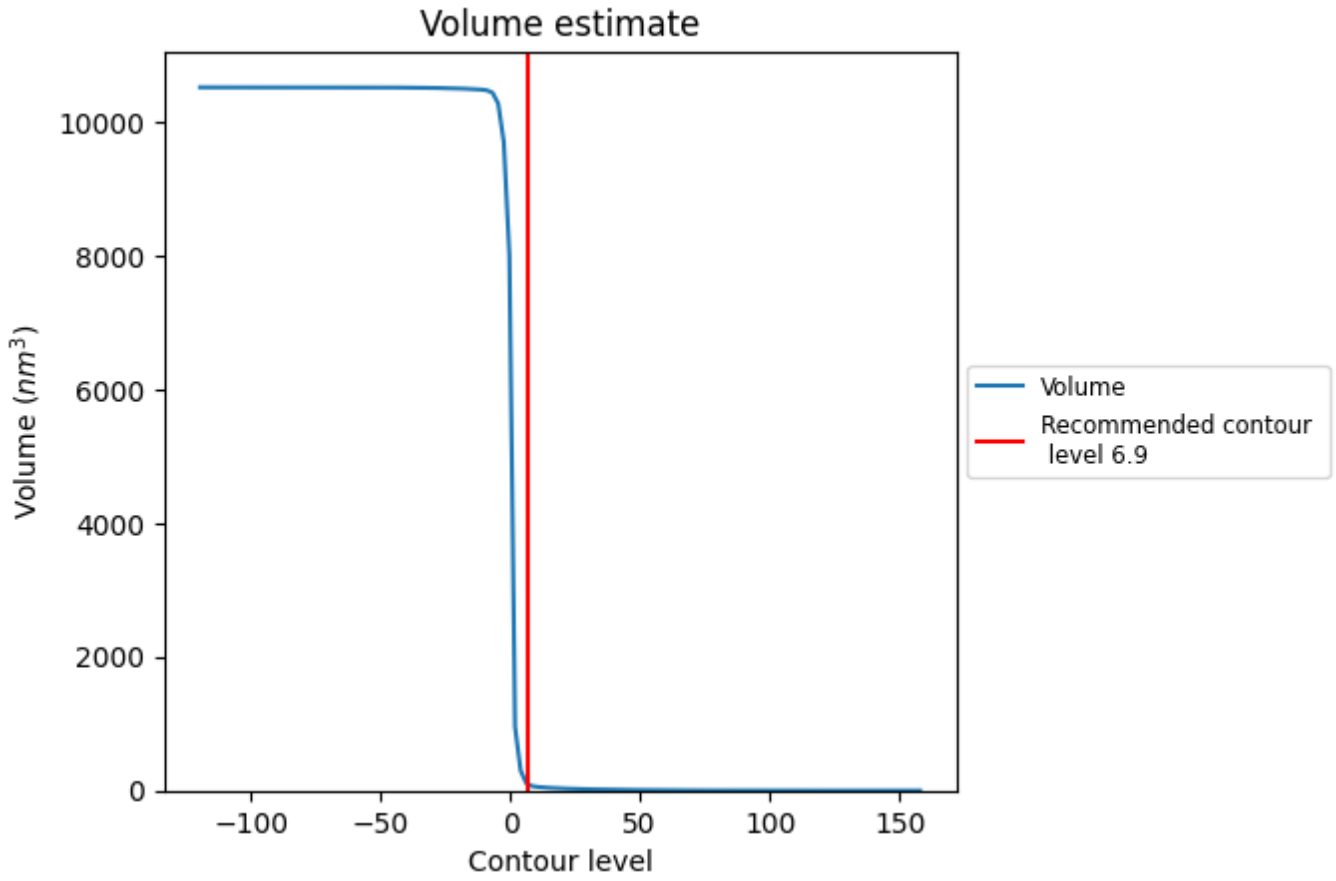
This section contains the results of statistical analysis of the map.

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.

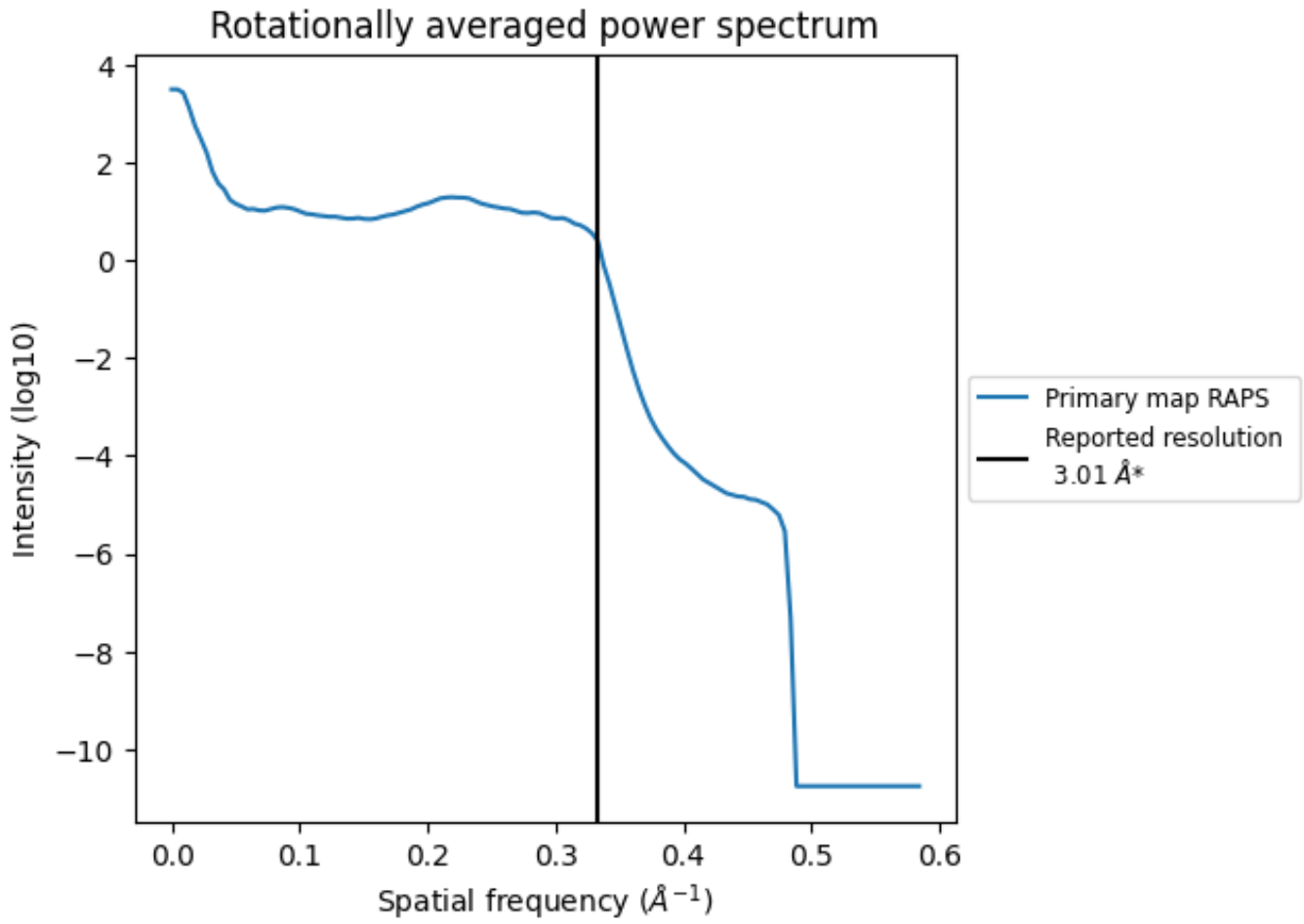
7.2 Volume estimate (i)



The volume at the recommended contour level is 97 nm^3; this corresponds to an approximate mass of 88 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 0.332 Å^{-1}

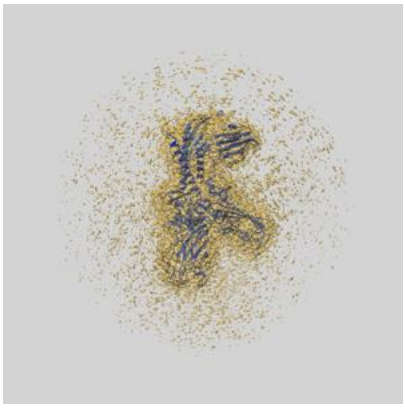
8 Fourier-Shell correlation (i)

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

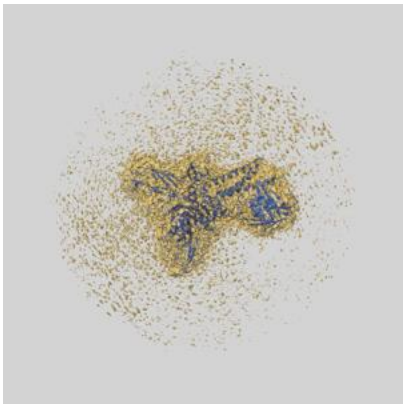
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-57262 and PDB model 29OC. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay



X



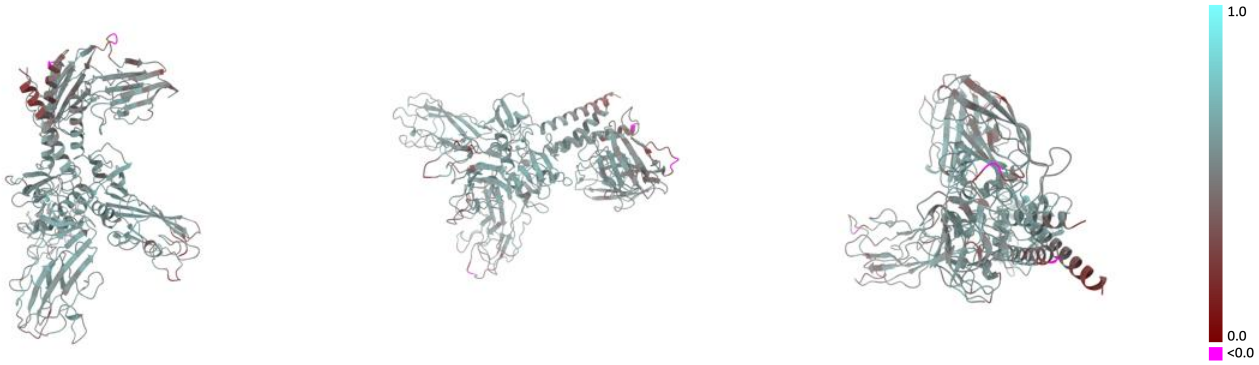
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 6.9 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.

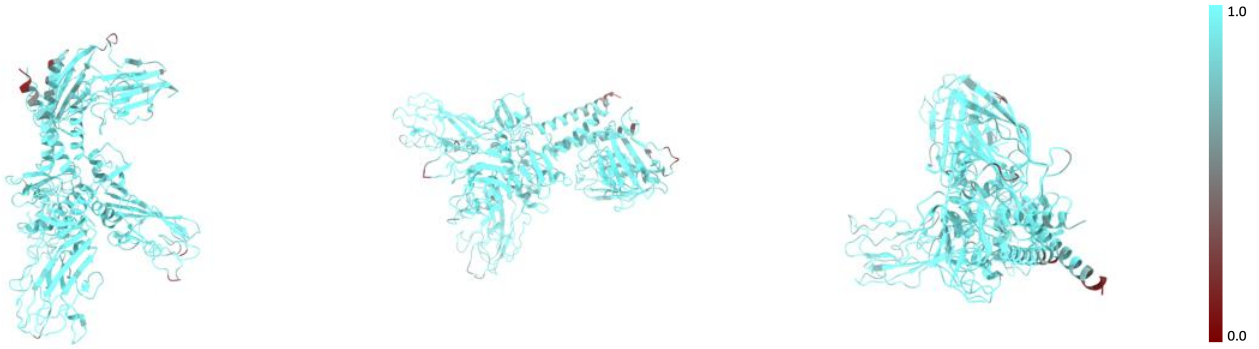
9.2 Q-score mapped to coordinate model (i)



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.

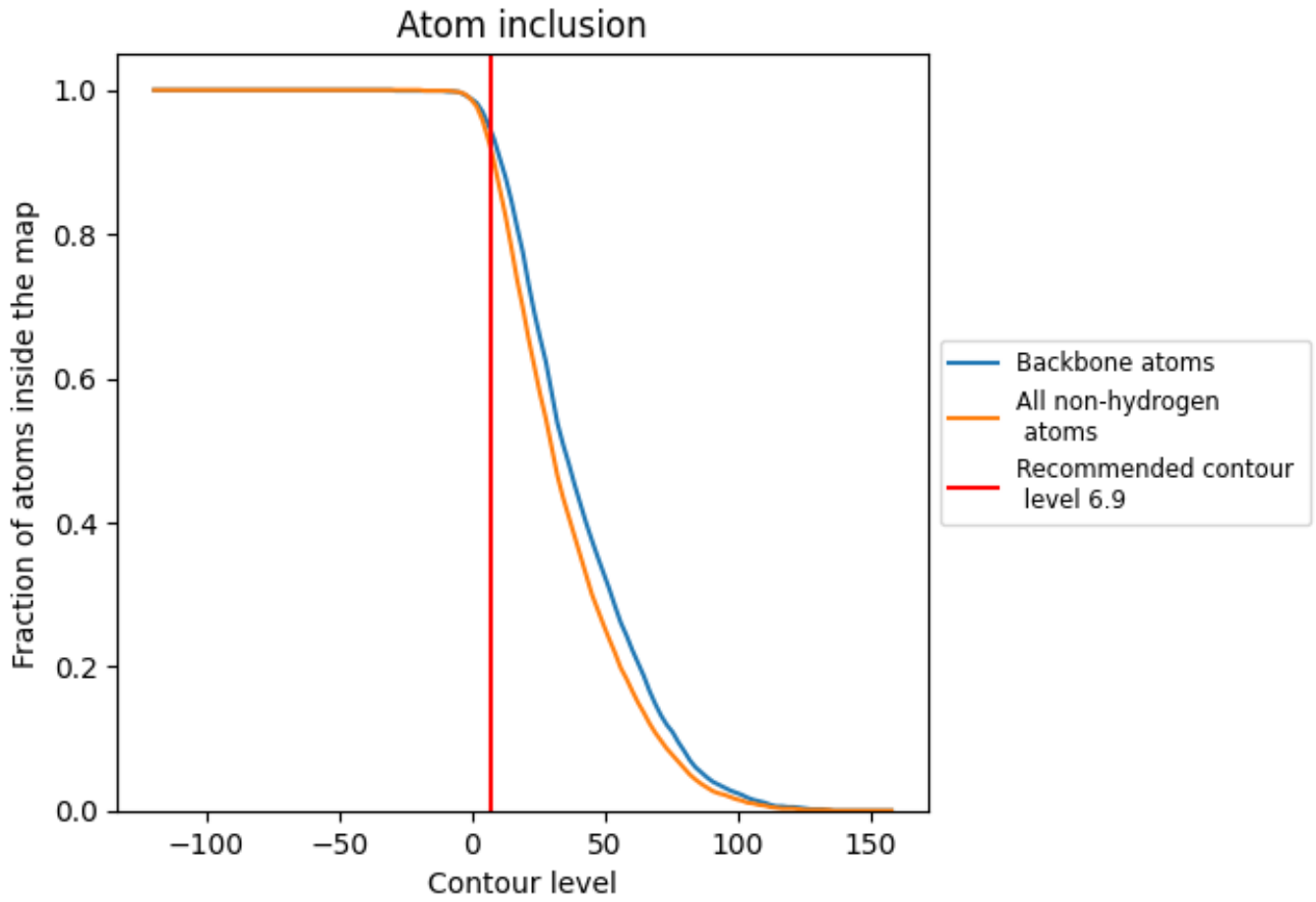
9.3 Atom inclusion mapped to coordinate model (i)



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 (6.9).

9.4 Atom inclusion (i)



At the recommended contour level, 95% of all backbone atoms, 92% 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 (6.9) and Q-score for the entire model and for each chain.



Chain	Atom inclusion	Q-score
All	0.9230	0.5270
A	0.9100	0.5060
B	0.9340	0.5450
C	0.9310	0.5430
D	0.9160	0.5150