wwPDB EM Validation Summary Report (i)

Nov 15, 2022 – 12:40 PM JST

PDB ID	:	6L42
EMDB ID	:	EMD-0828
Title	:	Structure of severe fever with thrombocytopenia syndrome virus L protein
Authors	:	Wang, P.; Lou, Z.
Deposited on	:	2019-10-15
Resolution	:	3.40 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:

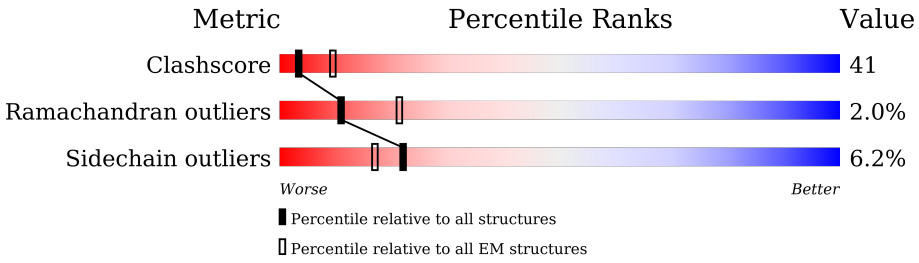
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.40 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	2109	49% 28% 7% 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14827 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 RNA polymerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1864	14826	9388	2574	2773	91	0	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	initiating methionine	UNP I0DF35
A	-23	SER	-	expression tag	UNP I0DF35
A	-22	TYR	-	expression tag	UNP I0DF35
A	-21	TYR	-	expression tag	UNP I0DF35
A	-20	HIS	-	expression tag	UNP I0DF35
A	-19	HIS	-	expression tag	UNP I0DF35
A	-18	HIS	-	expression tag	UNP I0DF35
A	-17	HIS	-	expression tag	UNP I0DF35
A	-16	HIS	-	expression tag	UNP I0DF35
A	-15	HIS	-	expression tag	UNP I0DF35
A	-14	ASP	-	expression tag	UNP I0DF35
A	-13	TYR	-	expression tag	UNP I0DF35
A	-12	ASP	-	expression tag	UNP I0DF35
A	-11	ILE	-	expression tag	UNP I0DF35
A	-10	PRO	-	expression tag	UNP I0DF35
A	-9	THR	-	expression tag	UNP I0DF35
A	-8	THR	-	expression tag	UNP I0DF35
A	-7	GLU	-	expression tag	UNP I0DF35
A	-6	ASN	-	expression tag	UNP I0DF35
A	-5	LEU	-	expression tag	UNP I0DF35
A	-4	TYR	-	expression tag	UNP I0DF35
A	-3	PHE	-	expression tag	UNP I0DF35
A	-2	GLN	-	expression tag	UNP I0DF35
A	-1	GLY	-	expression tag	UNP I0DF35
A	0	ALA	-	expression tag	UNP I0DF35
A	1321	GLU	GLN	engineered mutation	UNP I0DF35

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

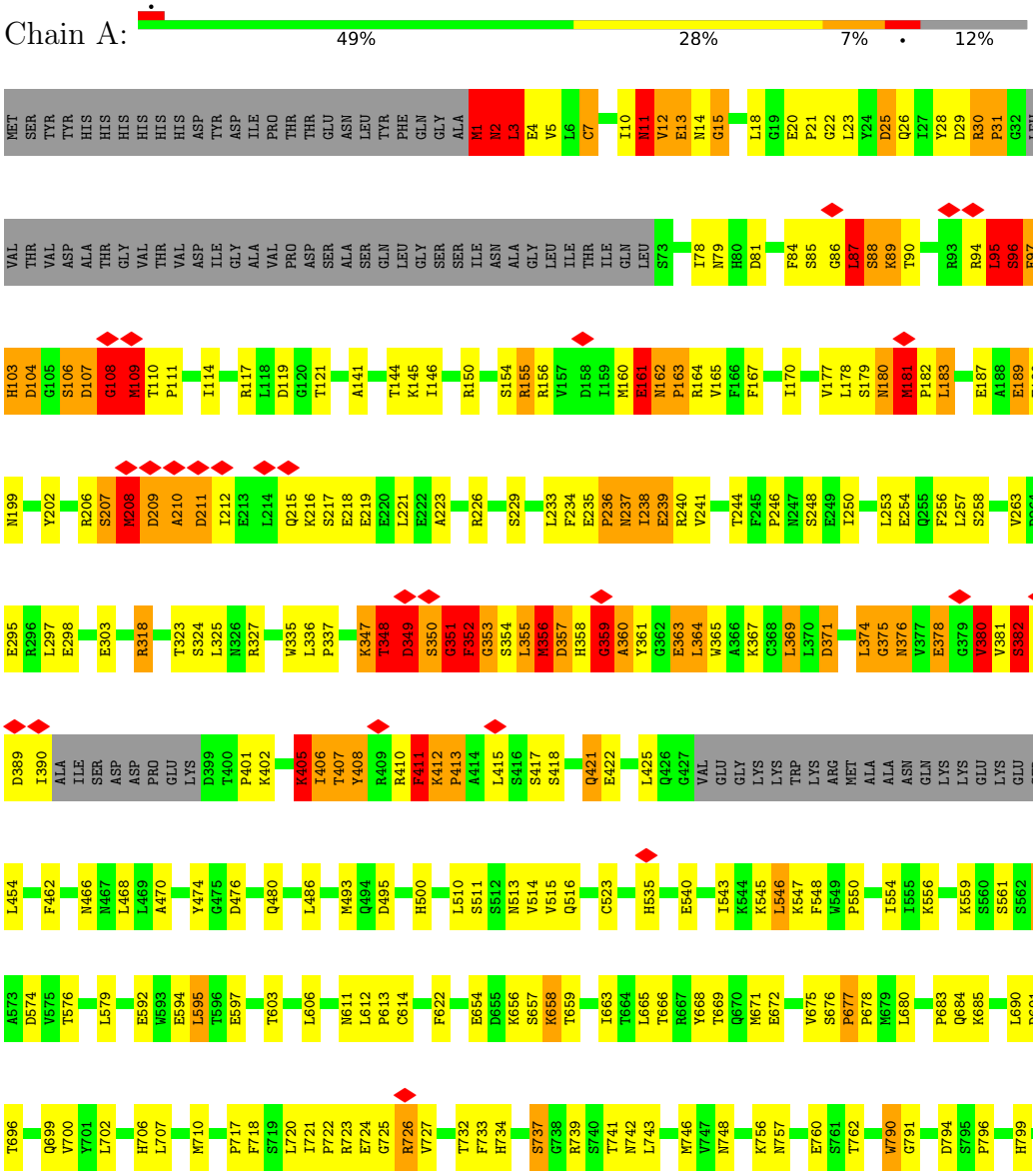
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
2	A	1	1	1	0

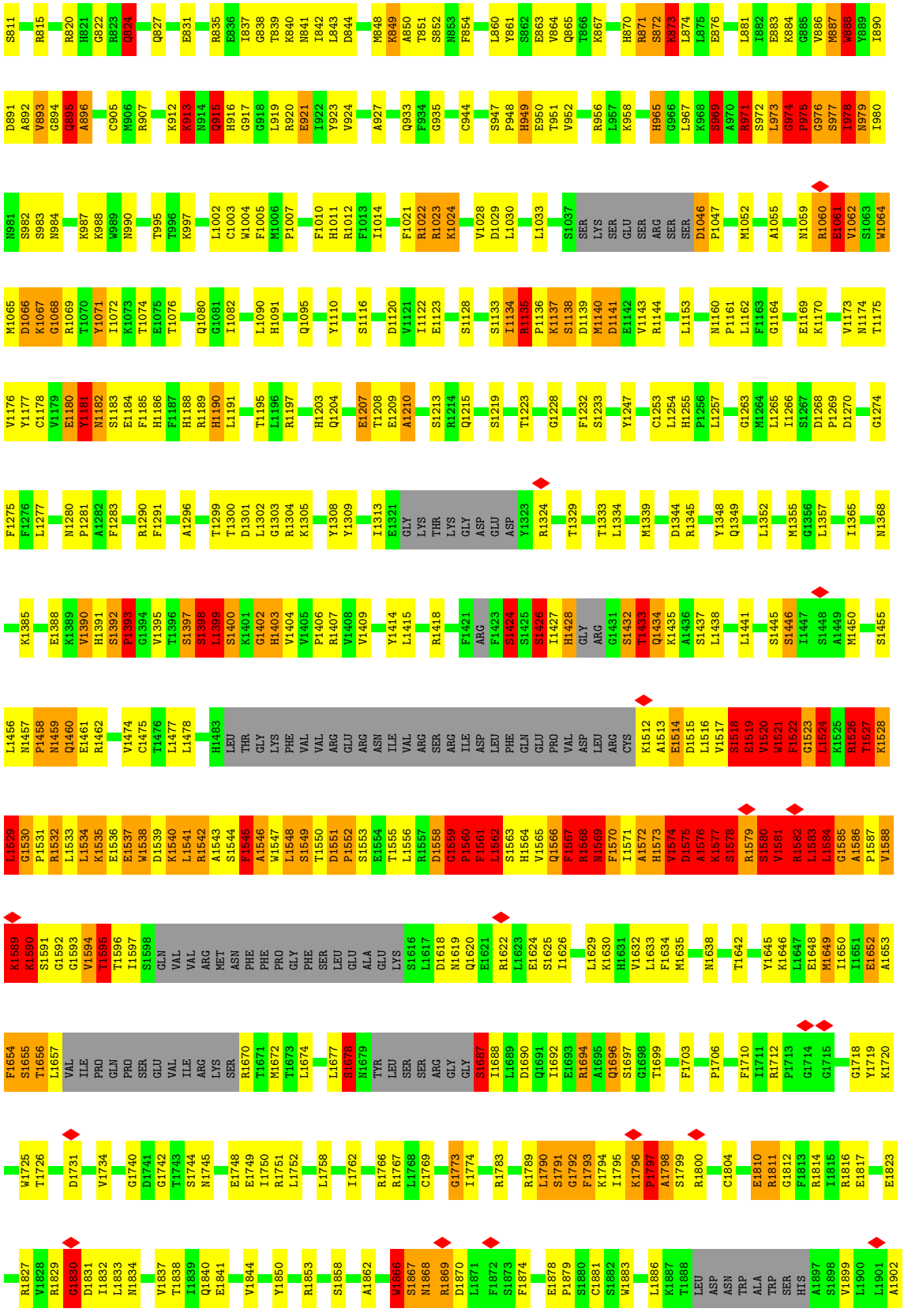
3 Residue-property plots (i)

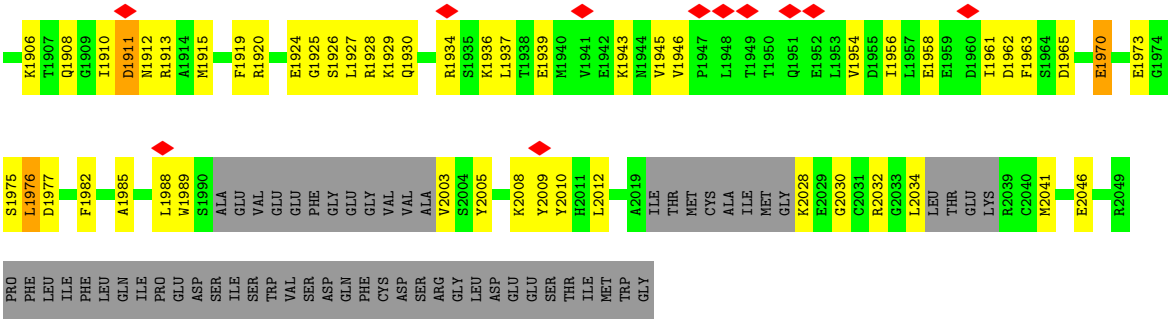
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: RNA polymerase

Chain A:






K1906 T1907 Q1908 G1909 I1910 D1911 N1912 R1913 A1914 M1915 F1919 R1920 E1924 G1925 S1926 L1927 R1928 K1929 Q1930 R1934 S1935 K1936 L1937 T1938 E1939 M1940 V1941 E1942 K1943 N1944 V1945 V1946 P1947 L1948 T1949 T1950 Q1951 E1952 L1953 V1954 D1955 I1956 L1957 E1958 E1959 D1960 I1961 D1962 F1963 S1964 D1965 E1970 E1973 G1974
S1975 L1976 D1977 F1982 A1985 L1988 W1989 S1990 ALA GLU VAL GLU GLU PHE GLY GLU GLY VAL VAL ALA V2003 S2004 Y2005 K2008 Y2009 Y2010 H2011 L2012 A2019 ILE THR MET CYS ALA ILE MET GLY K2028 E2029 G2030 G2031 R2032 G2033 L2034 LEU THR GLU LYS R2039 C2040 M2041 E2046 R2049
PRO PHE LEU ILE PHE LEU GLN ILE PRO GLU ASP SER ILE SER TRP VAL SER ASP GLN PHE CYS ASP SER ARG GLY LEU ASP GLU GLU SER THR ILE MET TRP GLY

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	147344	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 ($e^-/Å^2$)	40, 40	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k), GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.061	Depositor
Minimum map value	-0.028	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0085	Depositor
Map size (Å)	237.6, 237.6, 237.6	wwPDB
Map dimensions	220, 220, 220	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.08, 1.08, 1.08	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

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	0.63	50/15117 (0.3%)	1.58	342/20379 (1.7%)

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	94

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1575	ASP	CB-CG	-14.65	1.21	1.51
1	A	1520	VAL	CA-C	-13.50	1.17	1.52
1	A	1139	ASP	C-O	-11.93	1.00	1.23
1	A	348	THR	C-N	11.29	1.60	1.34
1	A	892	ALA	C-N	11.11	1.59	1.34

The worst 5 of 342 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1580	SER	O-C-N	-27.81	78.20	122.70
1	A	1	MET	O-C-N	-27.06	79.40	122.70
1	A	348	THR	O-C-N	-26.74	79.92	122.70
1	A	1567	PHE	O-C-N	-24.76	83.09	122.70
1	A	1687	SER	O-C-N	-24.62	83.31	122.70

There are no chirality outliers.

5 of 94 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1	MET	Mainchain
1	A	11	ASN	Mainchain
1	A	2	ASN	Mainchain
1	A	3	LEU	Mainchain
1	A	95	LEU	Mainchain

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	14826	0	14790	1202	0
2	A	1	0	0	0	0
All	All	14827	0	14790	1202	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 41.

The worst 5 of 1202 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:1556:LEU:HD11	1:A:1564:HIS:CD2	1.27	1.67
1:A:917:GLY:CA	1:A:920:ARG:HD3	1.26	1.65
1:A:1568:ARG:CA	1:A:1568:ARG:N	1.67	1.55
1:A:193:TYR:CE1	1:A:197:ILE:HD11	1.40	1.55
1:A:380:VAL:CG2	1:A:411:PHE:CE1	1.83	1.55

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	1832/2109 (87%)	1586 (87%)	209 (11%)	37 (2%)	7 30

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	163	PRO
1	A	356	MET
1	A	978	ILE
1	A	1424	SER
1	A	1426	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	1635/1848 (88%)	1533 (94%)	102 (6%)	18 48

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1177	TYR
1	A	1528	LYS
1	A	1911	ASP
1	A	1181	TYR
1	A	1433	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	895	GLN
1	A	915	GLN
1	A	1564	HIS
1	A	1204	GLN
1	A	1403	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is 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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

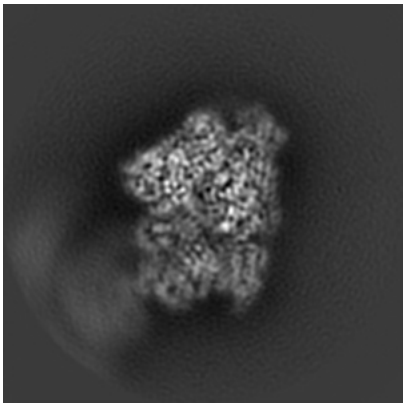
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-0828. These allow visual inspection of the internal detail of the map and identification of artifacts.

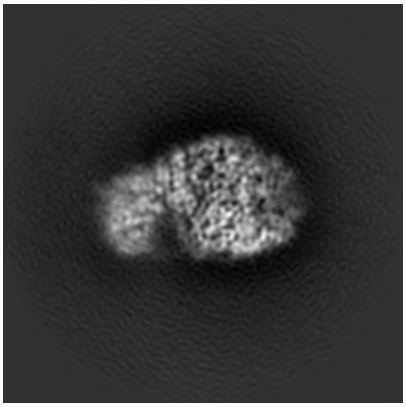
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

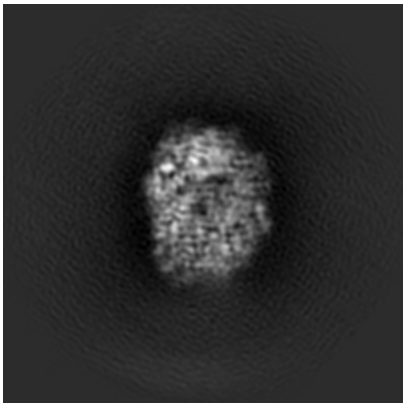
6.1.1 Primary map



X



Y

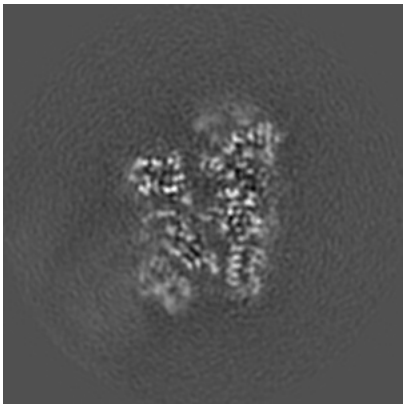


Z

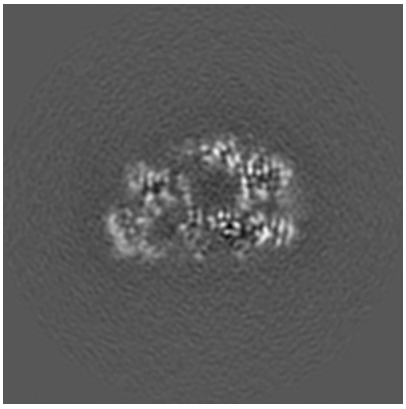
The images above show the map projected in three orthogonal directions.

6.2 Central slices

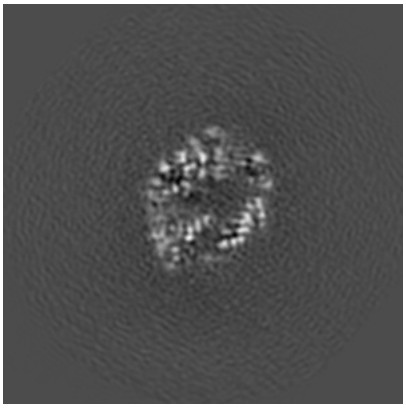
6.2.1 Primary map



X Index: 110



Y Index: 110

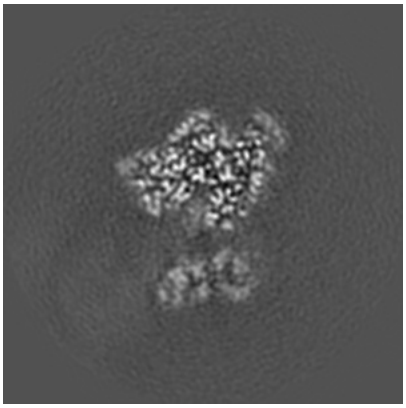


Z Index: 110

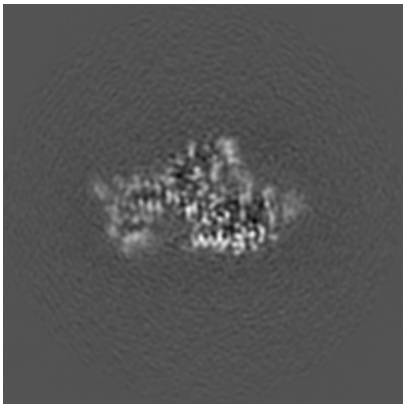
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

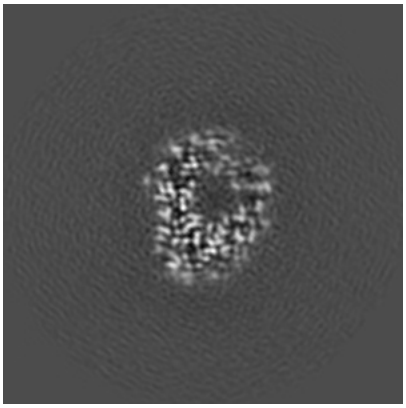
6.3.1 Primary map



X Index: 94



Y Index: 129

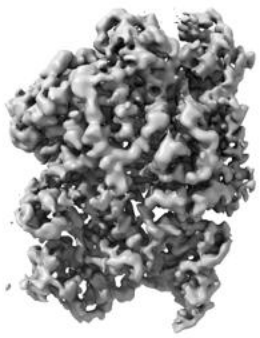


Z Index: 121

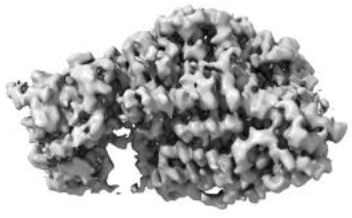
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

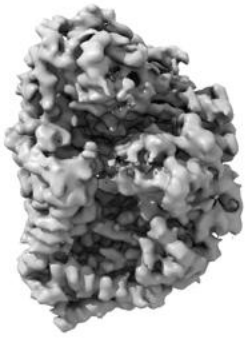
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0085. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

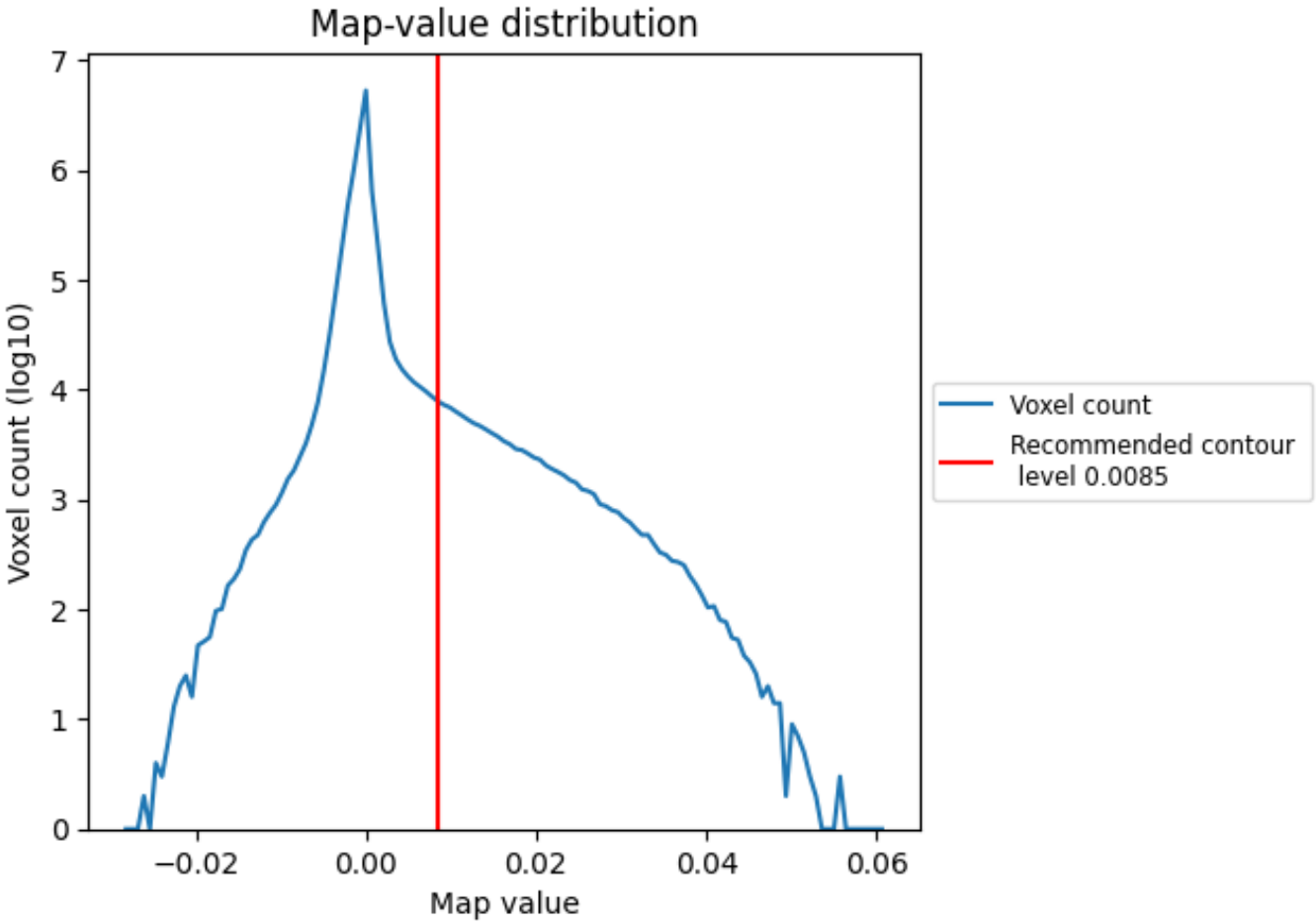
6.5 Mask visualisation (i)

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

7 Map analysis

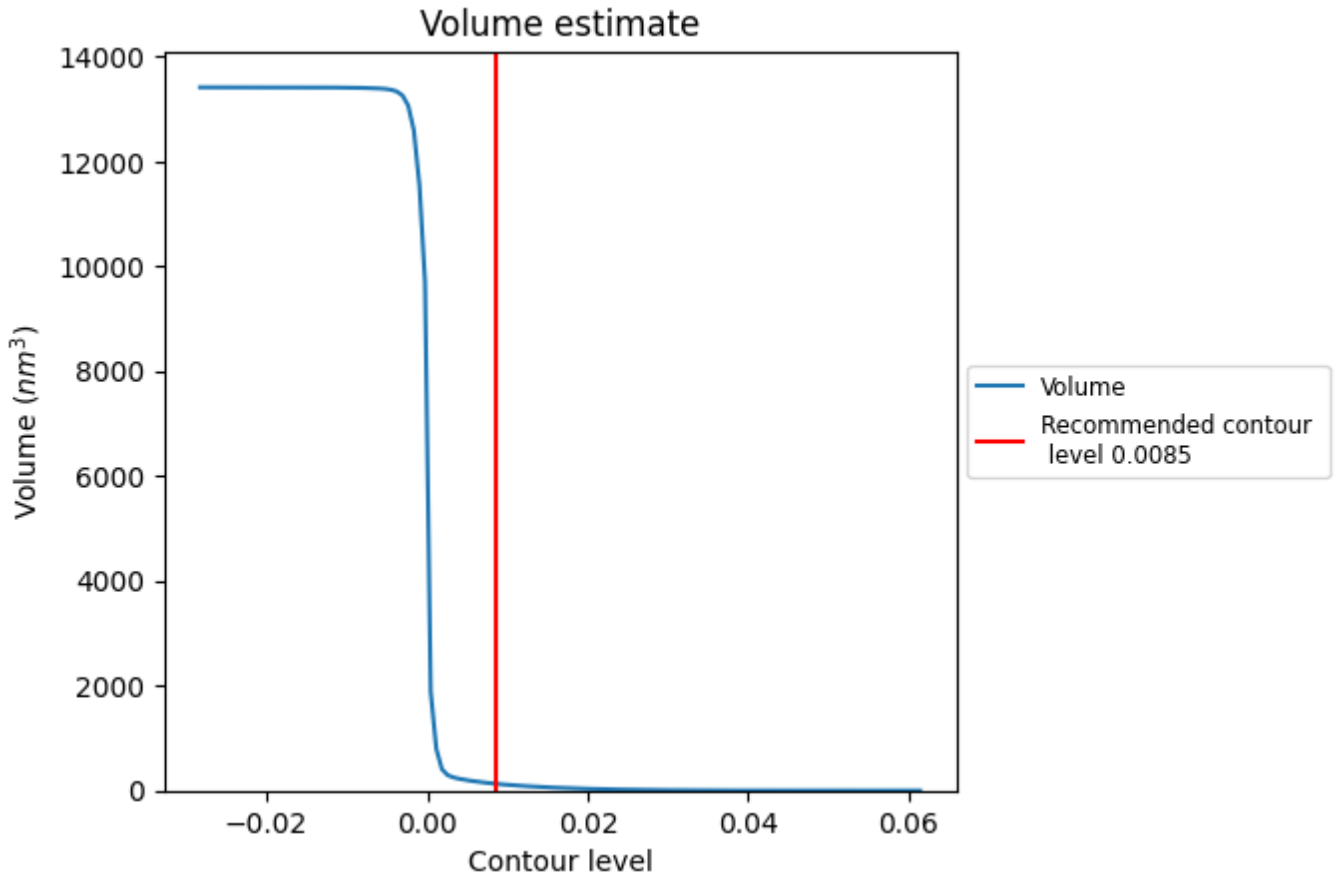
This section contains the results of statistical analysis of the map.

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.

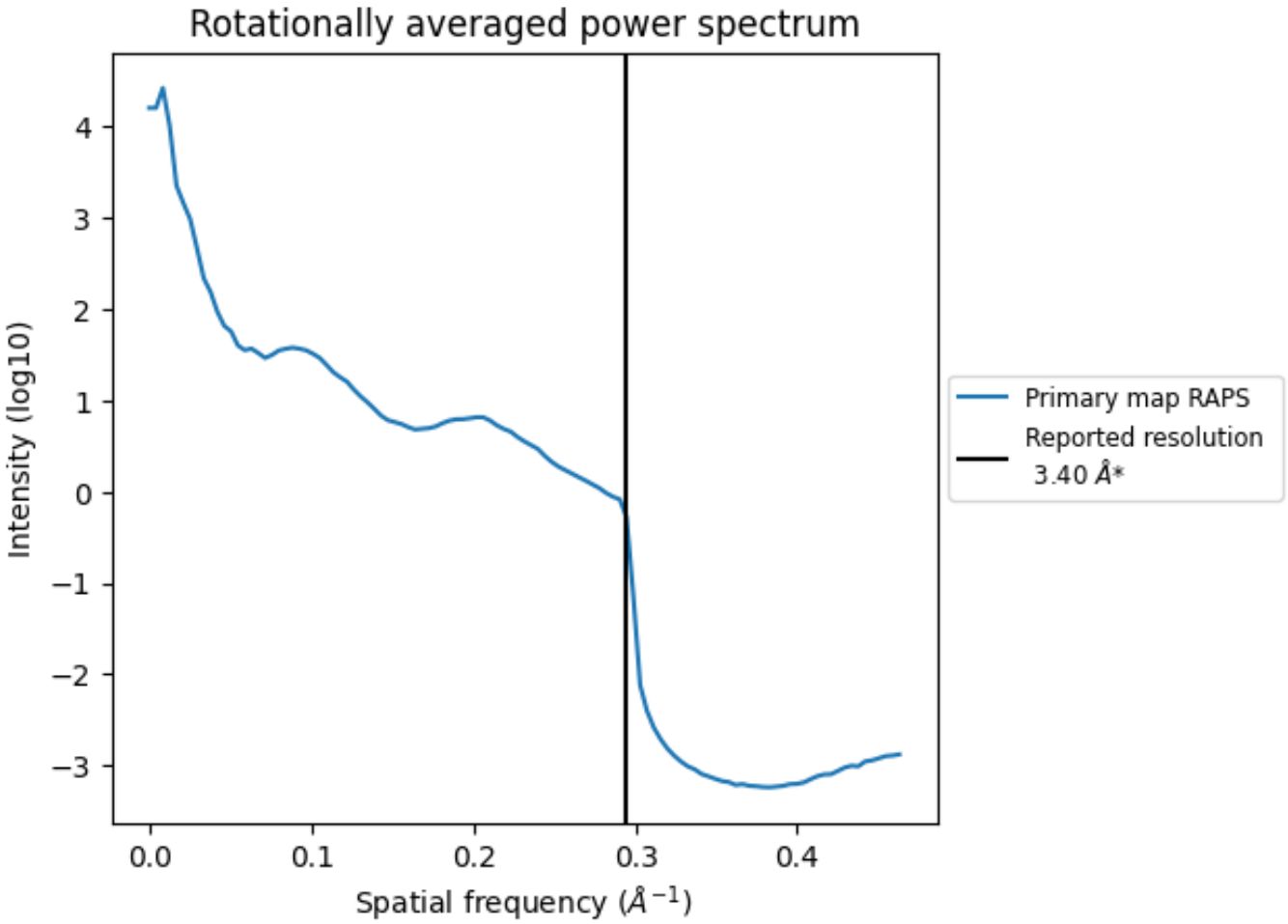
7.2 Volume estimate (i)



The volume at the recommended contour level is 130 nm^3; this corresponds to an approximate mass of 118 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 0.294 Å^{-1}

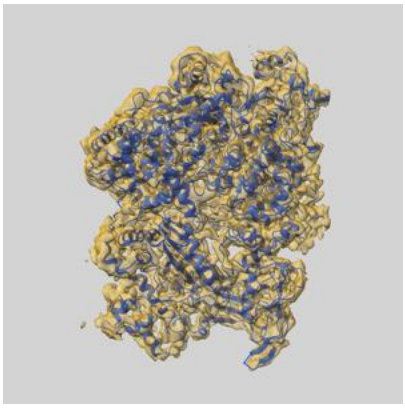
8 Fourier-Shell correlation (i)

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

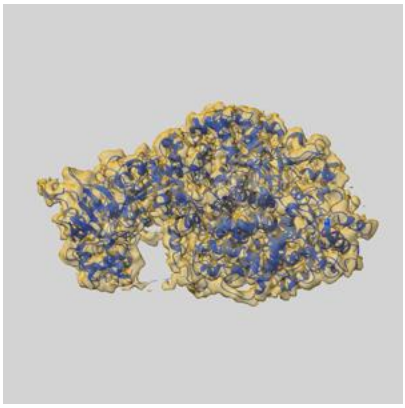
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-0828 and PDB model 6L42. Per-residue inclusion information can be found in section 3 on page 5.

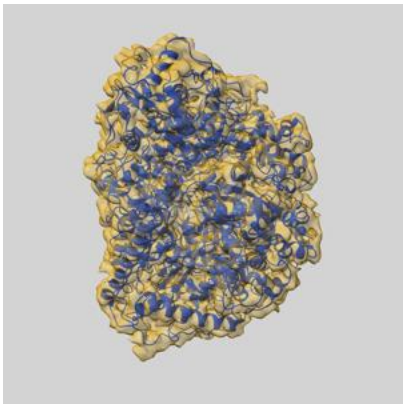
9.1 Map-model overlay



X



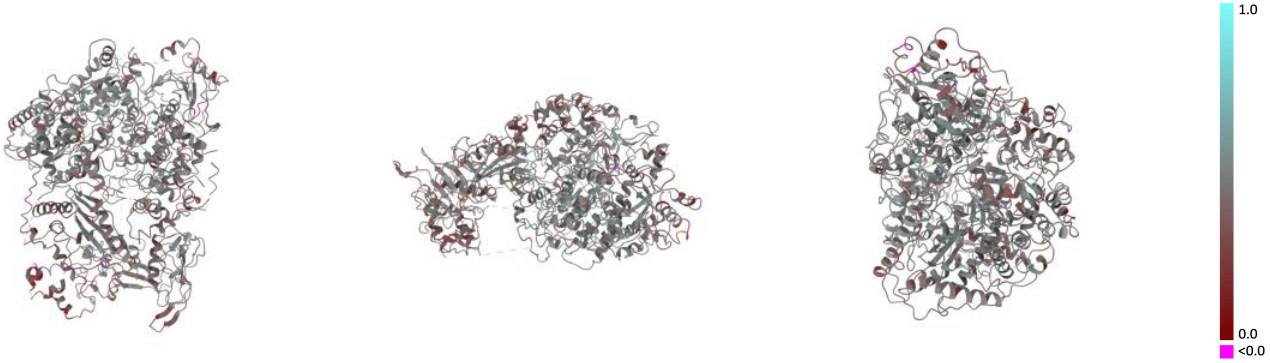
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0085 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.

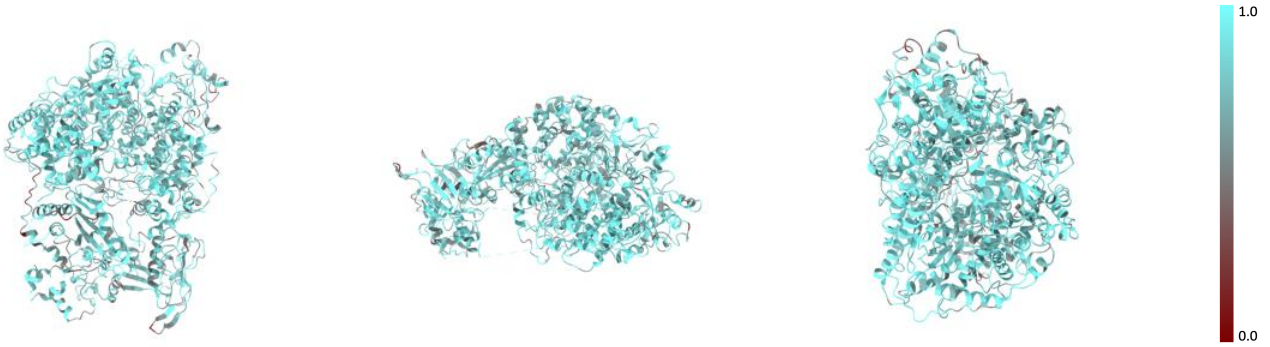
9.2 Q-score mapped to coordinate model (i)



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.

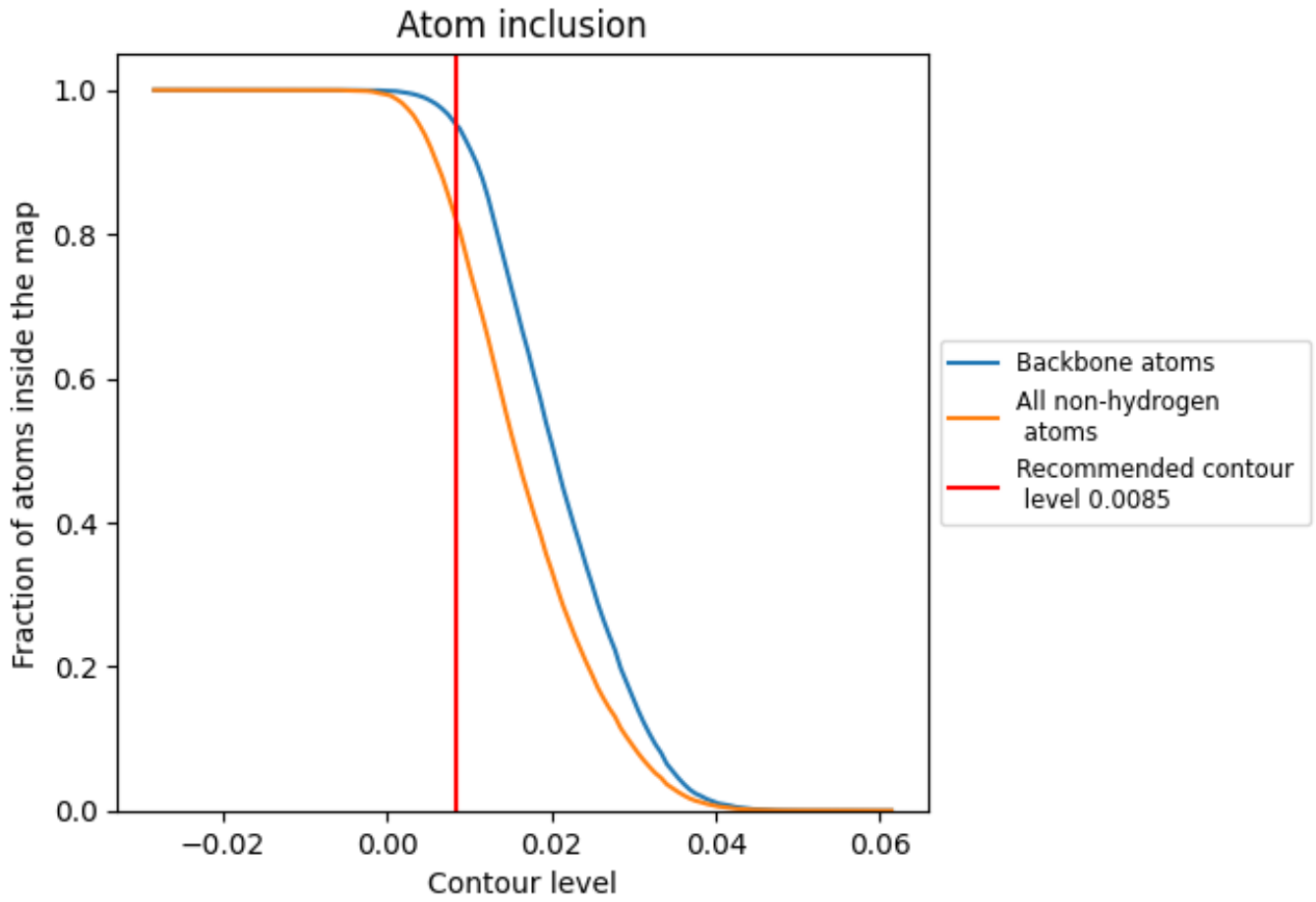
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0085).

9.4 Atom inclusion (i)



At the recommended contour level, 95% of all backbone atoms, 82% 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 (0.0085) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8170	0.4320
A	0.8170	0.4320