



## wwPDB EM Validation Summary Report ⓘ

Nov 27, 2022 – 04:58 AM EST

PDB ID : 6N7X  
EMDB ID : EMD-8622  
Title : S. cerevisiae U1 snRNP  
Authors : Li, X.; Liu, S.; Jiang, J.; Zhang, L.; Espinosa, S.; Hill, R.C.; Hansen, K.C.;  
Zhou, Z.H.; Zhao, R.  
Deposited on : 2018-11-28  
Resolution : 3.60 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](mailto: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:

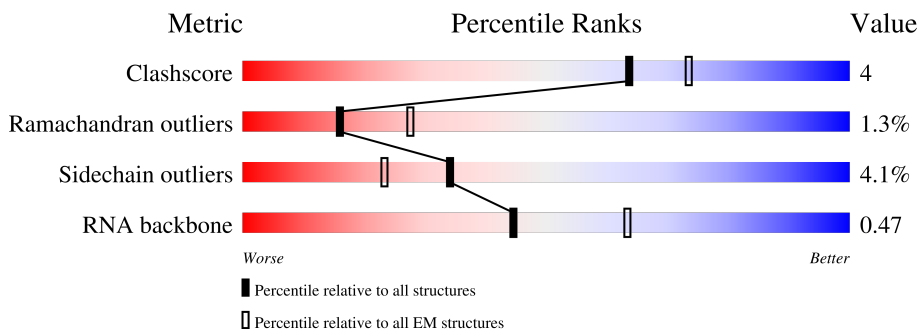
EMDB validation analysis : 0.0.1.dev43  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 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 3.60 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
RNA backbone	4643	859

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  $\geq 3$ , 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  $\leq 5\%$ . 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	300	
2	B	231	
3	C	298	
4	D	544	
5	E	629	
6	F	523	
7	G	492	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	H	620	6% 92%
9	K	196	53% 10% 37%
10	L	146	68% 13% 18%
11	M	110	82% 5% 13%
12	N	101	86% 5% 9%
13	O	94	69% 10% 21%
14	P	86	70% 9% 21%
15	Q	77	5% 92% 5%
16	R	568	7% 32% 20% 46%

## 2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 27183 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 U1 small nuclear ribonucleoprotein 70 kDa homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	186	1219	752	232	233	2	0	0

- Molecule 2 is a protein called U1 small nuclear ribonucleoprotein C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	176	1412	882	271	255	4	0	0

- Molecule 3 is a protein called U1 small nuclear ribonucleoprotein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	131	1041	663	190	185	3	0	0

- Molecule 4 is a protein called U1 small nuclear ribonucleoprotein component PRP42.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	539	4447	2918	702	813	14	0	0

- Molecule 5 is a protein called Pre-mRNA-processing factor 39.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	583	3908	2477	680	744	7	0	0

- Molecule 6 is a protein called Protein NAM8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	187	1414	889	244	271	10	0	0

- Molecule 7 is a protein called 56 kDa U1 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	229	1841	1202	300	328	11	0	0

- Molecule 8 is a protein called U1 small nuclear ribonucleoprotein component SNU71.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	52	420	266	74	79	1	0	0

- Molecule 9 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	K	124	1009	637	191	178	3	0	0

- Molecule 10 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	L	119	917	575	163	176	3	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	M	96	739	476	138	121	4	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	N	92	692	441	123	127	1	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	O	74	526	346	87	90	3	0	0

- Molecule 14 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	P	68	518	337	96	84	1	0	0

- Molecule 15 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	Q	75	543	346	94	102	1	0	0

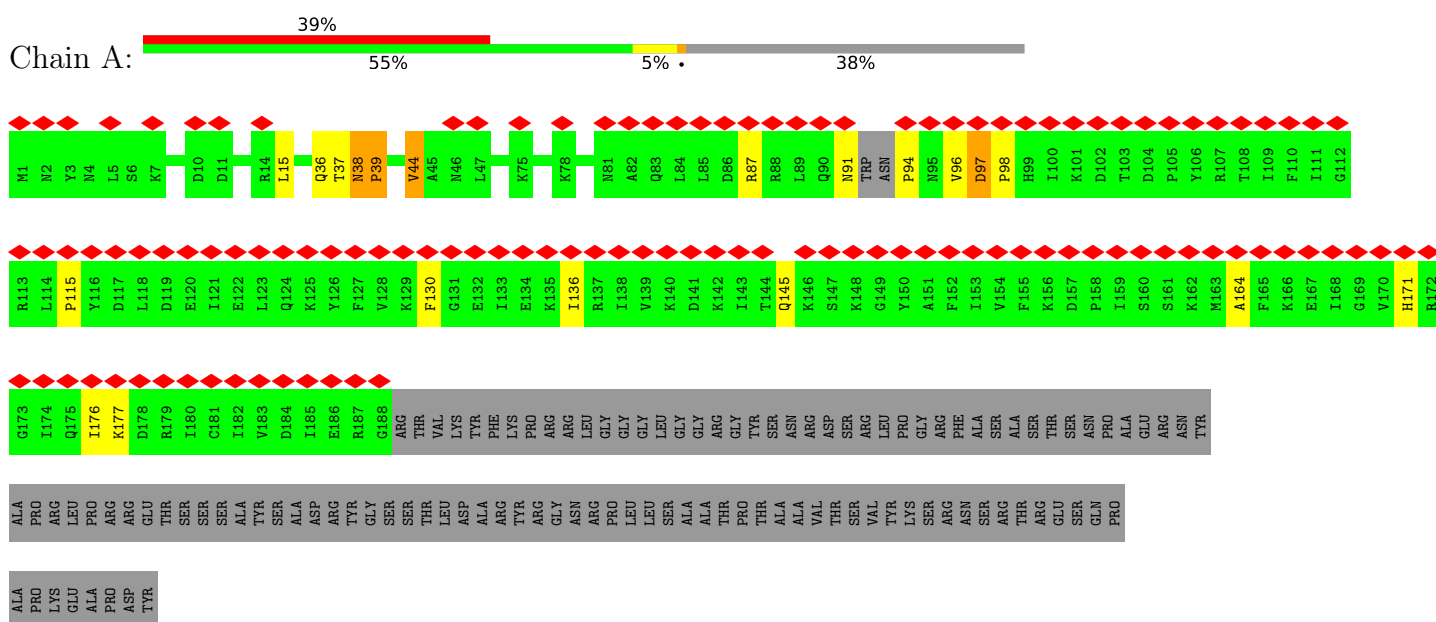
- Molecule 16 is a RNA chain called U1 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
16	R	308	6537	2923	1123	2183	308	0	0

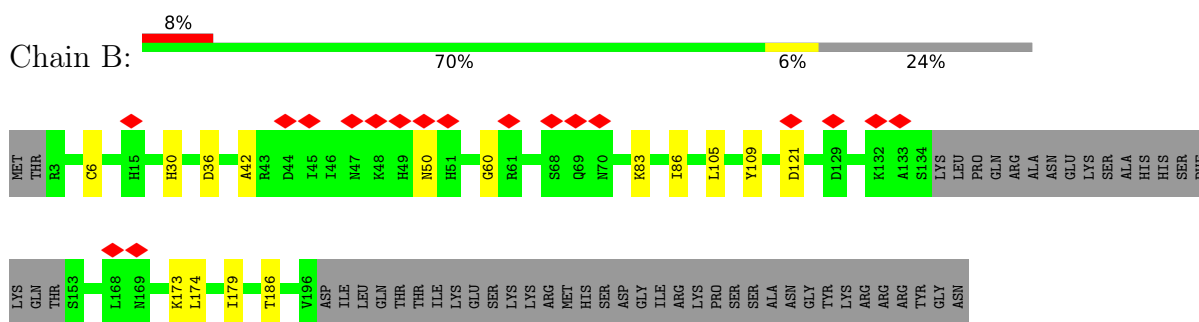
### 3 Residue-property plots [i](#)

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.

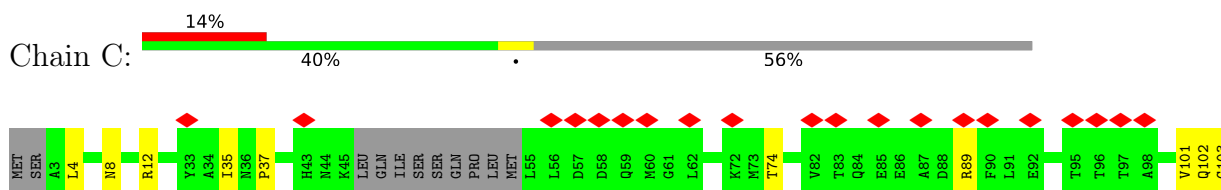
- Molecule 1: U1 small nuclear ribonucleoprotein 70 kDa homolog



- Molecule 2: U1 small nuclear ribonucleoprotein C

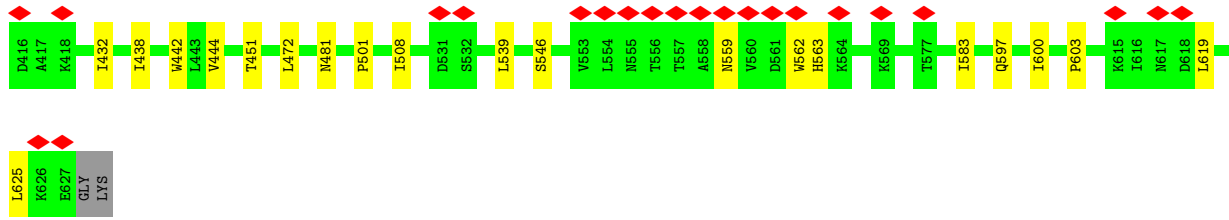


- Molecule 3: U1 small nuclear ribonucleoprotein A

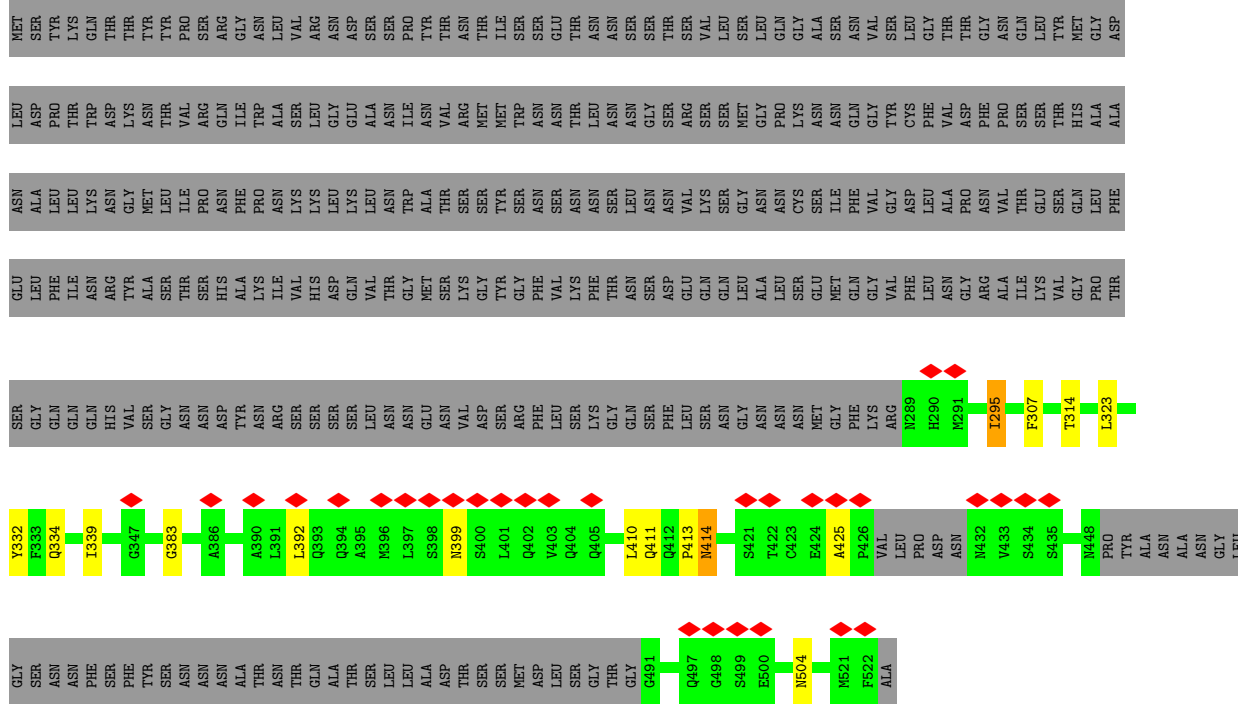




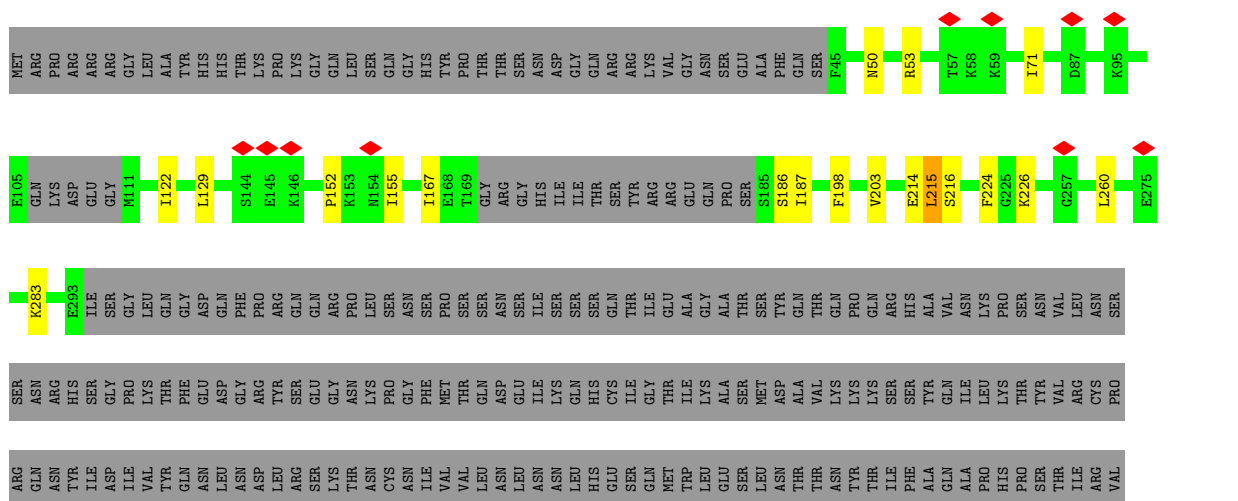




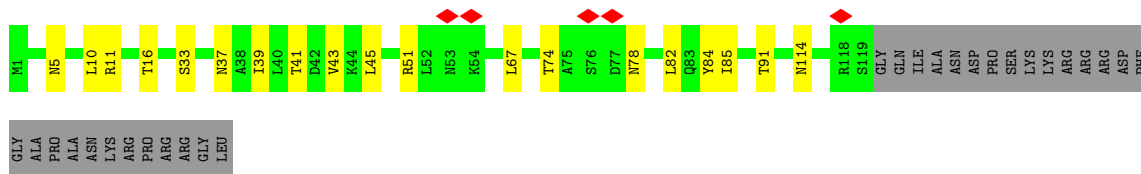
• Molecule 6: Protein NAM8



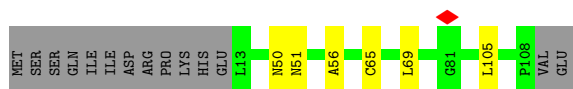
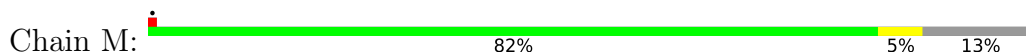
• Molecule 7: 56 kDa U1 small nuclear ribonucleoprotein component



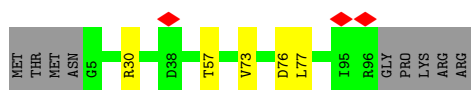
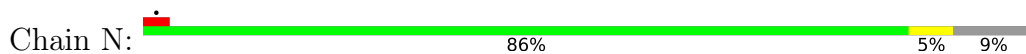




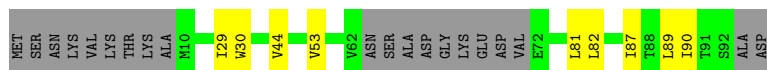
• Molecule 11: Small nuclear ribonucleoprotein Sm D2



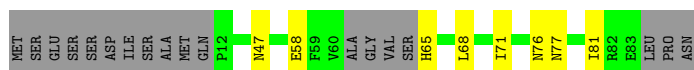
• Molecule 12: Small nuclear ribonucleoprotein Sm D3



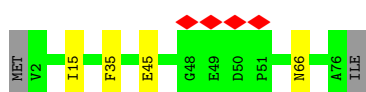
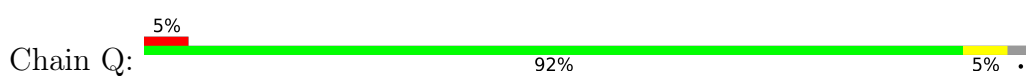
• Molecule 13: Small nuclear ribonucleoprotein E



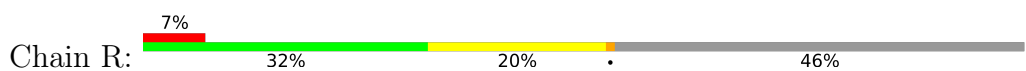
• Molecule 14: Small nuclear ribonucleoprotein F



• Molecule 15: Small nuclear ribonucleoprotein G



• Molecule 16: U1 snRNA





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	352900	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 ( $e^-/\text{\AA}^2$ )	51.9	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	0.343	Depositor
Minimum map value	-0.213	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.06	Depositor
Map size (Å)	435.2, 435.2, 435.2	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.36, 1.36, 1.36	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.46	0/1237	0.79	2/1691 (0.1%)
2	B	0.42	0/1441	0.66	0/1941
3	C	0.42	0/1055	0.63	0/1417
4	D	0.45	0/4561	0.64	1/6186 (0.0%)
5	E	0.45	0/3971	0.68	6/5462 (0.1%)
6	F	0.43	0/1441	0.63	0/1957
7	G	0.44	0/1882	0.60	0/2537
8	H	0.55	0/428	0.56	0/575
9	K	0.46	0/1016	0.77	1/1355 (0.1%)
10	L	0.44	0/926	0.68	0/1257
11	M	0.39	0/751	0.66	0/1014
12	N	0.46	0/704	0.69	0/957
13	O	0.44	0/535	0.56	0/730
14	P	0.44	0/529	0.66	0/715
15	Q	0.43	0/548	0.70	0/745
16	R	0.31	0/7295	0.77	2/11341 (0.0%)
All	All	0.41	0/28320	0.70	12/39880 (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
1	A	0	5
5	E	0	11
All	All	0	16

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	115	PRO	N-CA-CB	6.29	110.85	103.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	216	PRO	N-CA-CB	5.97	110.46	103.30
5	E	187	PRO	N-CA-CB	5.85	110.32	103.30
16	R	152	G	C2'-C3'-O3'	5.79	122.97	113.70
5	E	90	PRO	N-CA-CB	5.71	110.15	103.30

There are no chirality outliers.

5 of 16 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	136	ILE	Peptide
1	A	145	GLN	Peptide
1	A	176	ILE	Peptide
1	A	96	VAL	Peptide
1	A	97	ASP	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	1219	0	943	5	0
2	B	1412	0	1394	10	0
3	C	1041	0	1087	3	0
4	D	4447	0	4337	54	0
5	E	3908	0	3102	37	0
6	F	1414	0	1363	10	0
7	G	1841	0	1813	8	0
8	H	420	0	421	32	0
9	K	1009	0	1100	14	0
10	L	917	0	961	8	0
11	M	739	0	756	2	0
12	N	692	0	710	1	0
13	O	526	0	515	6	0
14	P	518	0	502	5	0
15	Q	543	0	538	1	0
16	R	6537	0	3294	60	0
All	All	27183	0	22836	213	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 4.

The worst 5 of 213 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)
8:H:45:HIS:CD2	8:H:46:ILE:HG13	1.50	1.43
8:H:45:HIS:NE2	8:H:46:ILE:HG13	1.57	1.19
4:D:376:SER:HB2	8:H:50:ARG:NH1	1.62	1.15
4:D:54:LYS:HE3	8:H:1:MET:CE	1.82	1.09
4:D:54:LYS:HE3	8:H:1:MET:HE2	1.25	1.09

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	182/300 (61%)	155 (85%)	21 (12%)	6 (3%)	4	31
2	B	172/231 (74%)	153 (89%)	18 (10%)	1 (1%)	25	64
3	C	125/298 (42%)	112 (90%)	9 (7%)	4 (3%)	4	31
4	D	537/544 (99%)	494 (92%)	39 (7%)	4 (1%)	22	61
5	E	579/629 (92%)	517 (89%)	46 (8%)	16 (3%)	5	34
6	F	181/523 (35%)	168 (93%)	11 (6%)	2 (1%)	14	53
7	G	223/492 (45%)	208 (93%)	15 (7%)	0	100	100
8	H	50/620 (8%)	48 (96%)	2 (4%)	0	100	100
9	K	120/196 (61%)	107 (89%)	13 (11%)	0	100	100
10	L	117/146 (80%)	103 (88%)	12 (10%)	2 (2%)	9	45
11	M	94/110 (86%)	90 (96%)	4 (4%)	0	100	100
12	N	90/101 (89%)	87 (97%)	3 (3%)	0	100	100

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	O	70/94 (74%)	69 (99%)	1 (1%)	0	100	100
14	P	64/86 (74%)	59 (92%)	5 (8%)	0	100	100
15	Q	73/77 (95%)	64 (88%)	9 (12%)	0	100	100
All	All	2677/4447 (60%)	2434 (91%)	208 (8%)	35 (1%)	16	50

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	ASP
1	A	98	PRO
3	C	102	GLN
4	D	413	ASN
5	E	86	LYS

### 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	84/265 (32%)	80 (95%)	4 (5%)	25	60
2	B	154/214 (72%)	149 (97%)	5 (3%)	39	70
3	C	113/273 (41%)	109 (96%)	4 (4%)	36	68
4	D	483/519 (93%)	464 (96%)	19 (4%)	32	65
5	E	291/603 (48%)	285 (98%)	6 (2%)	53	78
6	F	154/451 (34%)	149 (97%)	5 (3%)	39	70
7	G	199/448 (44%)	194 (98%)	5 (2%)	47	75
8	H	48/568 (8%)	40 (83%)	8 (17%)	2	14
9	K	113/176 (64%)	107 (95%)	6 (5%)	22	58
10	L	106/129 (82%)	96 (91%)	10 (9%)	8	38
11	M	77/103 (75%)	73 (95%)	4 (5%)	23	58
12	N	76/89 (85%)	72 (95%)	4 (5%)	22	58
13	O	52/83 (63%)	52 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
14	P	51/77 (66%)	50 (98%)	1 (2%)	55 79
15	Q	54/66 (82%)	51 (94%)	3 (6%)	21 56
All	All	2055/4064 (51%)	1971 (96%)	84 (4%)	34 64

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	K	43	LEU
10	L	114	ASN
9	K	56	THR
10	L	16	THR
11	M	105	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
6	F	432	ASN
8	H	37	GLN
7	G	207	HIS
8	H	51	ASN
4	D	65	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
16	R	299/568 (52%)	68 (22%)	17 (5%)

5 of 68 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
16	R	17	A
16	R	51	A
16	R	54	C
16	R	55	G
16	R	56	C

5 of 17 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
16	R	279	U
16	R	564	A
16	R	113	G
16	R	152	G
16	R	186	U

#### 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](#)

There are no ligands in this entry.

#### 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.

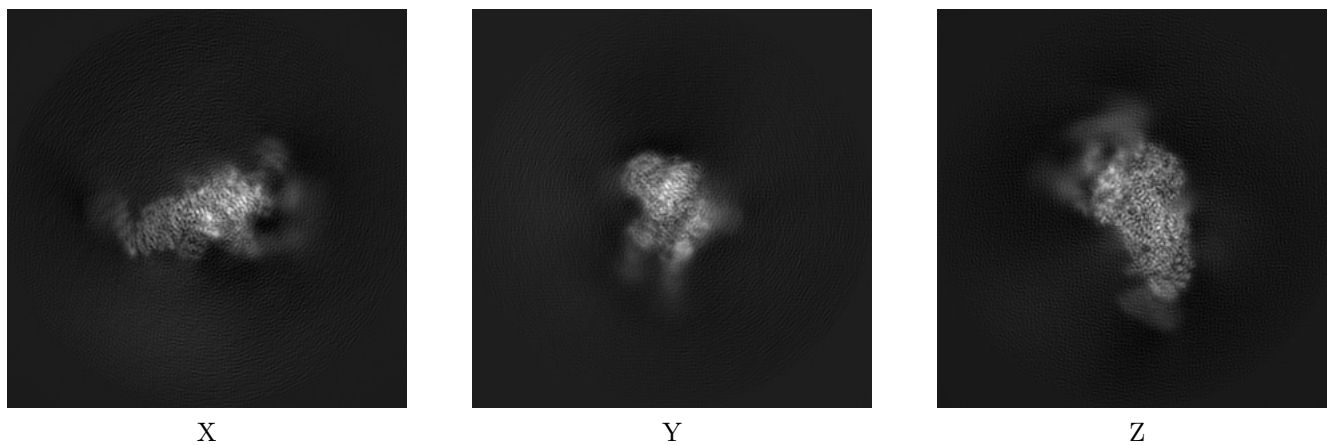
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-8622. 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](#)

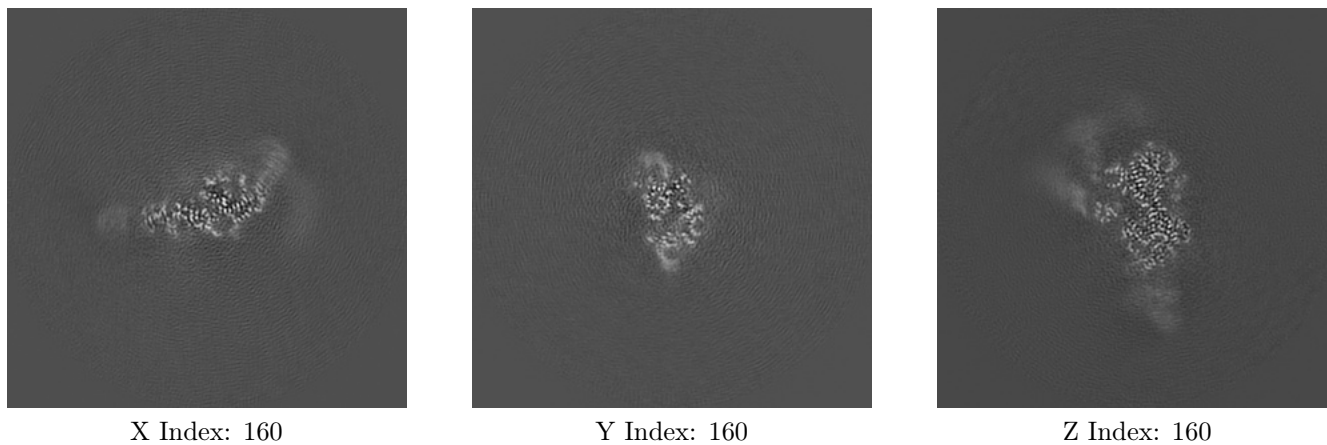
#### 6.1.1 Primary map



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

### 6.2 Central slices [i](#)

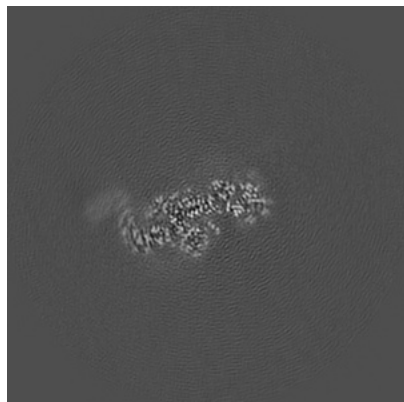
#### 6.2.1 Primary map



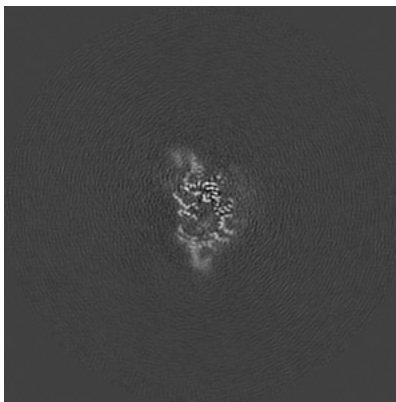
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

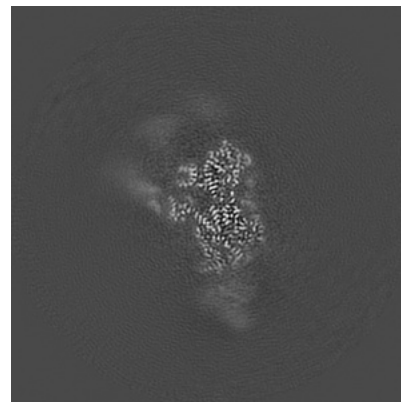
### 6.3.1 Primary map



X Index: 178



Y Index: 162



Z Index: 158

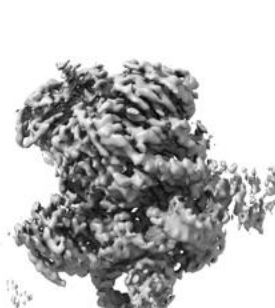
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal surface views [i](#)

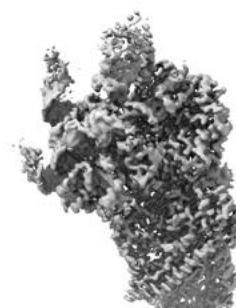
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.06. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

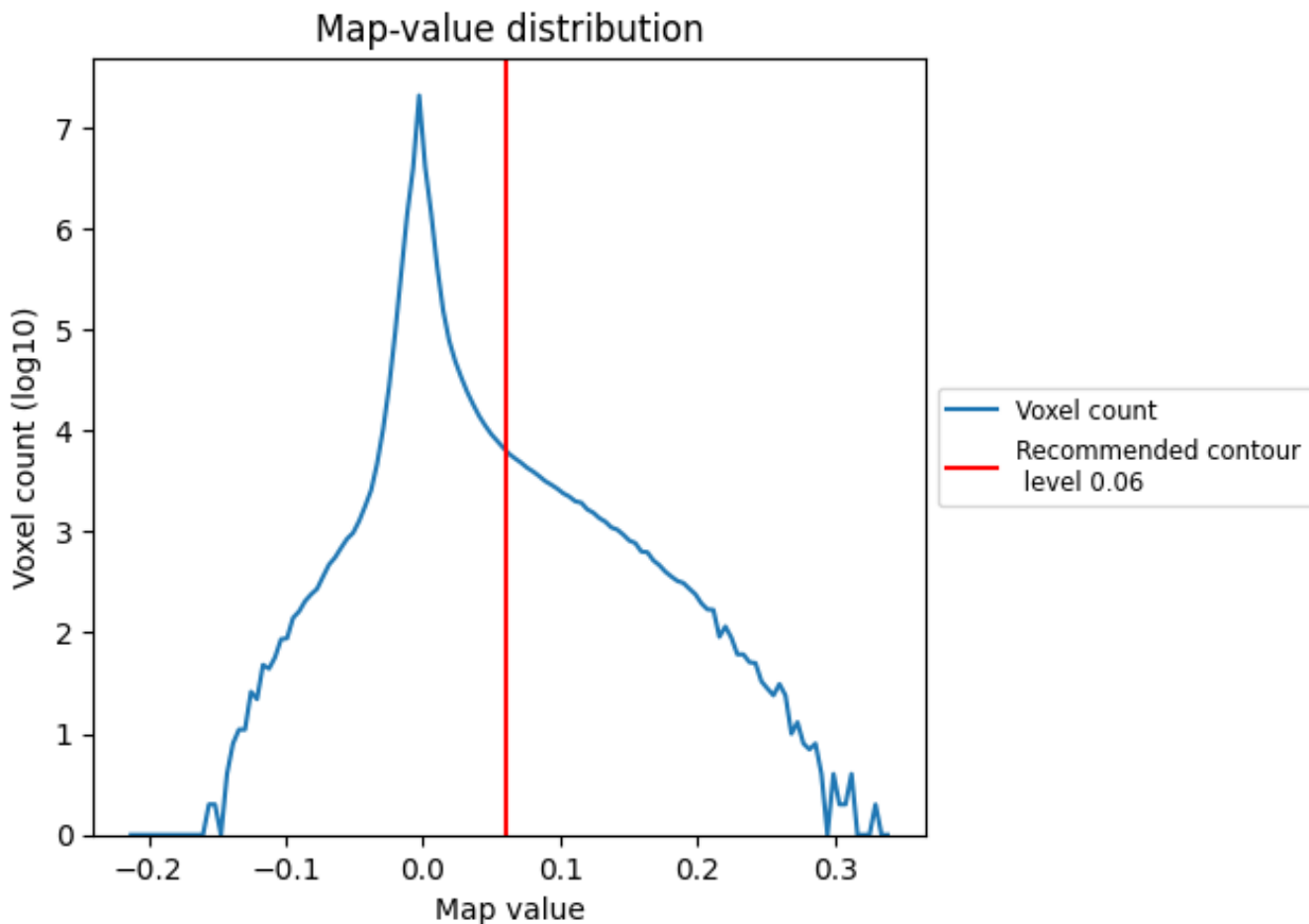
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

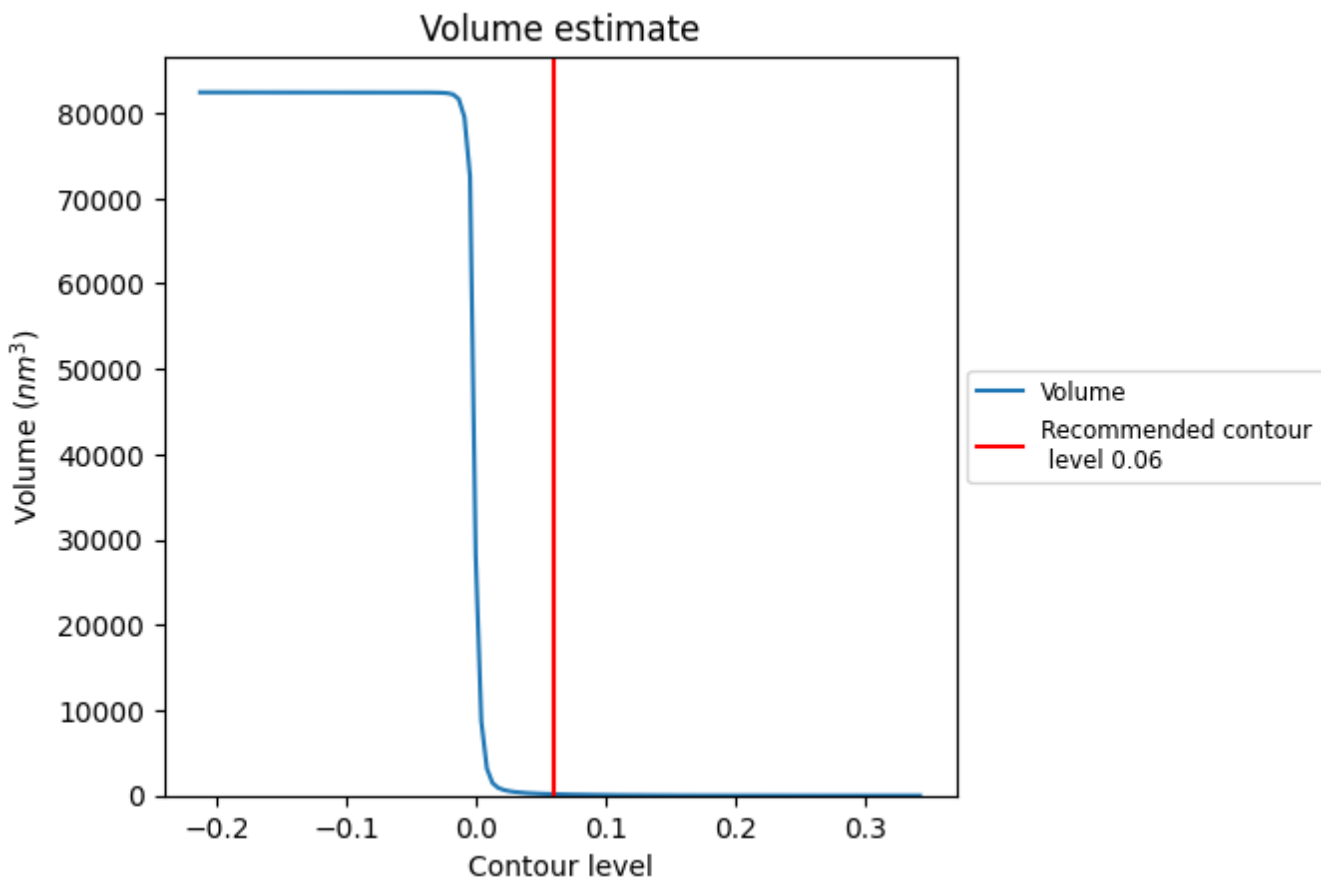
This section contains the results of statistical analysis of the map.

### 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](#)

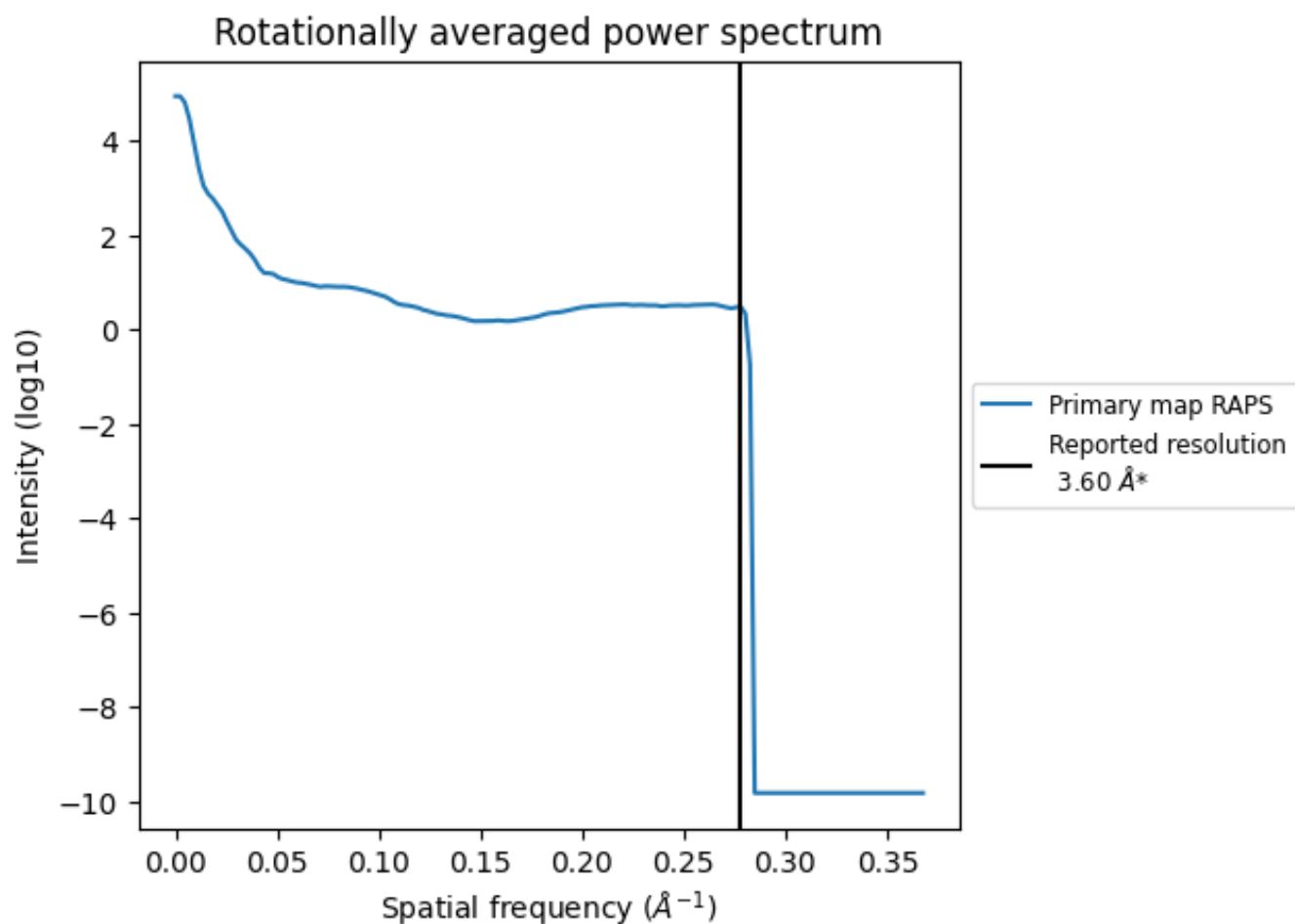


The volume at the recommended contour level is 164 nm<sup>3</sup>; this corresponds to an approximate mass of 148 kDa.

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 0.278 Å<sup>-1</sup>

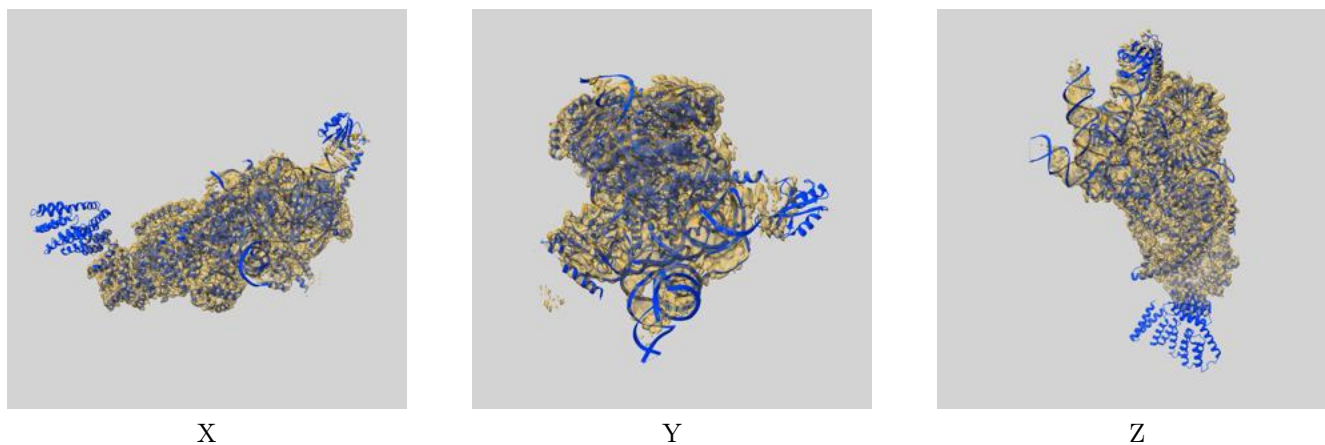
## 8 Fourier-Shell correlation

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

## 9 Map-model fit [i](#)

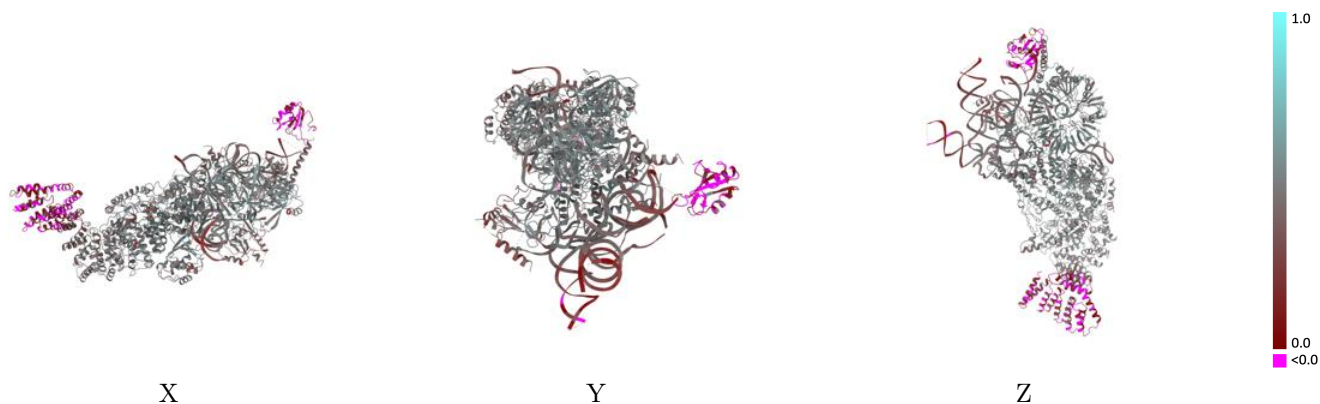
This section contains information regarding the fit between EMDB map EMD-8622 and PDB model 6N7X. Per-residue inclusion information can be found in section 3 on page 7.

### 9.1 Map-model overlay [i](#)



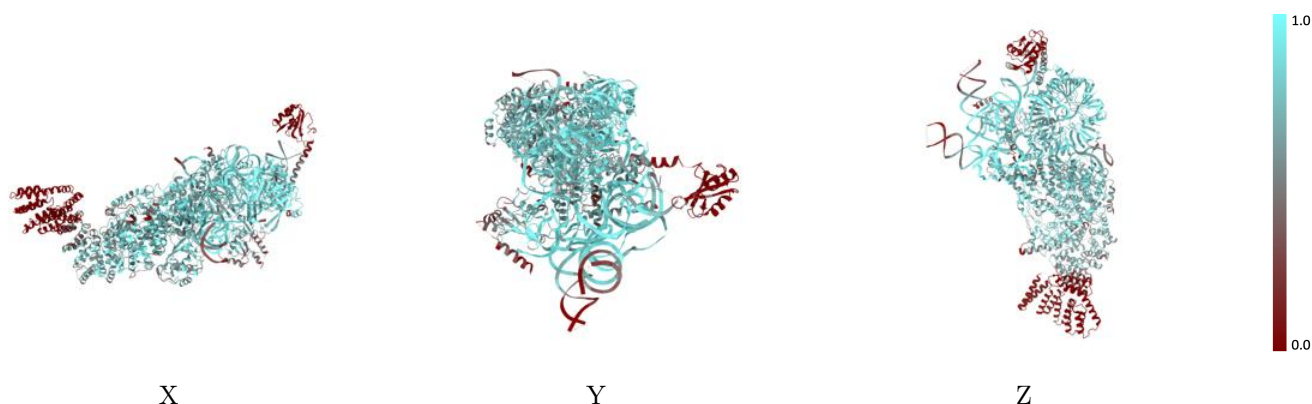
The images above show the 3D surface view of the map at the recommended contour level 0.06 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [\(i\)](#)



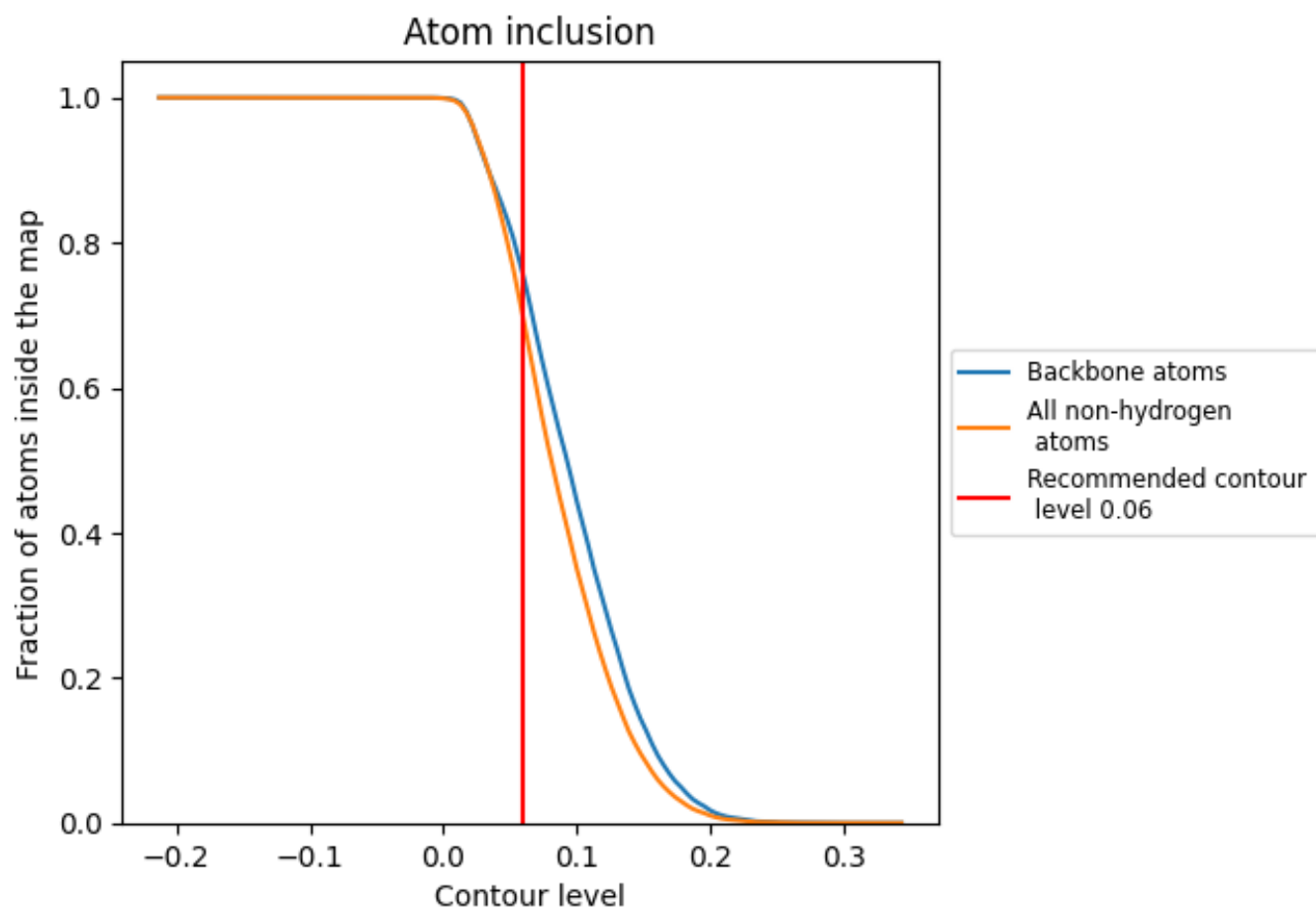
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 [\(i\)](#)



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 (0.06).



































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 75% of all backbone atoms, 69% 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 (0.06) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6931	 0.4280
A	 0.4165	 0.2810
B	 0.7176	 0.4670
C	 0.5511	 0.3960
D	 0.7961	 0.4950
E	 0.4314	 0.3250
F	 0.6700	 0.4750
G	 0.7141	 0.4750
H	 0.5704	 0.4320
K	 0.7563	 0.5070
L	 0.7456	 0.4920
M	 0.8122	 0.4890
N	 0.8272	 0.5200
O	 0.8571	 0.4920
P	 0.8817	 0.5010
Q	 0.7784	 0.5020
R	 0.7750	 0.3910

