wwPDB EM Validation Summary Report ⓘ

Mar 11, 2024 – 06:43 AM EDT

PDB ID	:	6PSS
EMDB ID	:	EMD-20462
Title	:	Escherichia coli RNA polymerase promoter unwinding intermediate (TRPi1.5a) with TraR and mutant rpsT P2 promoter
Authors	:	Chen, J.; Chiu, C.E.; Campbell, E.A.; Darst, S.A.
Deposited on	:	2019-07-13
Resolution	:	3.50 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

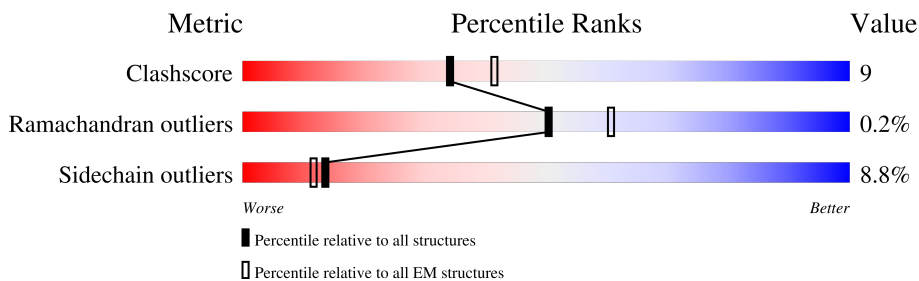|		
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)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

1   Overall quality at a glance   ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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.
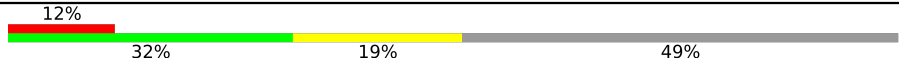
Metric	Percentile Ranks	Value
Clashscore		9
Ramachandran outliers		0.2%
Sidechain outliers		8.8%

Worse   *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all EM structures

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	G	329	54% 13% • 29%
1	H	329	5% 48% 15% • 34%
1	M	329	22% 16% 6% • 78%
2	I	1342	11% 71% 27% •
3	J	1430	20% 67% 25% • 6%
4	K	91	18% 47% 31% • 21%
5	L	616	45% 67% 22% • 9%
6	N	72	10% 72% 25% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	O	85	12% 32% 19% 49%
8	P	85	12% 36% 14% 49%

2 Entity composition ⓘ

There are 10 unique types of molecules in this entry. The entry contains 32361 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
1	G	232	Total	C	N	O	S	0	0
			1769	1106	315	342	6		
1	H	218	Total	C	N	O	S	0	0
			1669	1044	293	326	6		
1	M	73	Total	C	N	O	S	0	0
			572	362	100	108	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	I	1337	Total	C	N	O	S	0	0
			10502	6594	1830	2035	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	J	1343	Total	C	N	O	S	0	0
			10449	6567	1864	1968	50		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	1	VAL	-	expression tag	UNP P0A8T7
J	1408	LEU	-	expression tag	UNP P0A8T7
J	1409	GLU	-	expression tag	UNP P0A8T7
J	1410	LEU	-	expression tag	UNP P0A8T7
J	1411	GLU	-	expression tag	UNP P0A8T7
J	1412	VAL	-	expression tag	UNP P0A8T7
J	1413	LEU	-	expression tag	UNP P0A8T7
J	1414	PHE	-	expression tag	UNP P0A8T7
J	1415	GLN	-	expression tag	UNP P0A8T7
J	1416	GLY	-	expression tag	UNP P0A8T7
J	1417	PRO	-	expression tag	UNP P0A8T7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	1418	SER	-	expression tag	UNP P0A8T7
J	1419	SER	-	expression tag	UNP P0A8T7
J	1420	GLY	-	expression tag	UNP P0A8T7
J	1421	HIS	-	expression tag	UNP P0A8T7
J	1422	HIS	-	expression tag	UNP P0A8T7
J	1423	HIS	-	expression tag	UNP P0A8T7
J	1424	HIS	-	expression tag	UNP P0A8T7
J	1425	HIS	-	expression tag	UNP P0A8T7
J	1426	HIS	-	expression tag	UNP P0A8T7
J	1427	HIS	-	expression tag	UNP P0A8T7
J	1428	HIS	-	expression tag	UNP P0A8T7
J	1429	HIS	-	expression tag	UNP P0A8T7
J	1430	HIS	-	expression tag	UNP P0A8T7

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	K	72	577	352	110	114	1	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	L	558	4489	2801	785	876	27	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	-2	SER	-	expression tag	UNP Q0P6L9
L	-1	GLU	-	expression tag	UNP Q0P6L9
L	0	PHE	-	expression tag	UNP Q0P6L9

- Molecule 6 is a protein called Protein TraR.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	N	72	566	350	103	108	5	0	0

- Molecule 7 is a DNA chain called DNA (85-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	O	43	Total	C	N	O	P	0	0
			883	420	174	246	43		

- Molecule 8 is a DNA chain called DNA (85-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	P	43	Total	C	N	O	P	0	0
			881	422	148	268	43		

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

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

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

Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total	Zn	0
			2	2	
10	N	1	Total	Zn	0
			1	1	

3 Residue-property plots ⓘ

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 G:

• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H:

• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain M:

• Molecule 2: DNA-directed RNA polymerase subunit beta
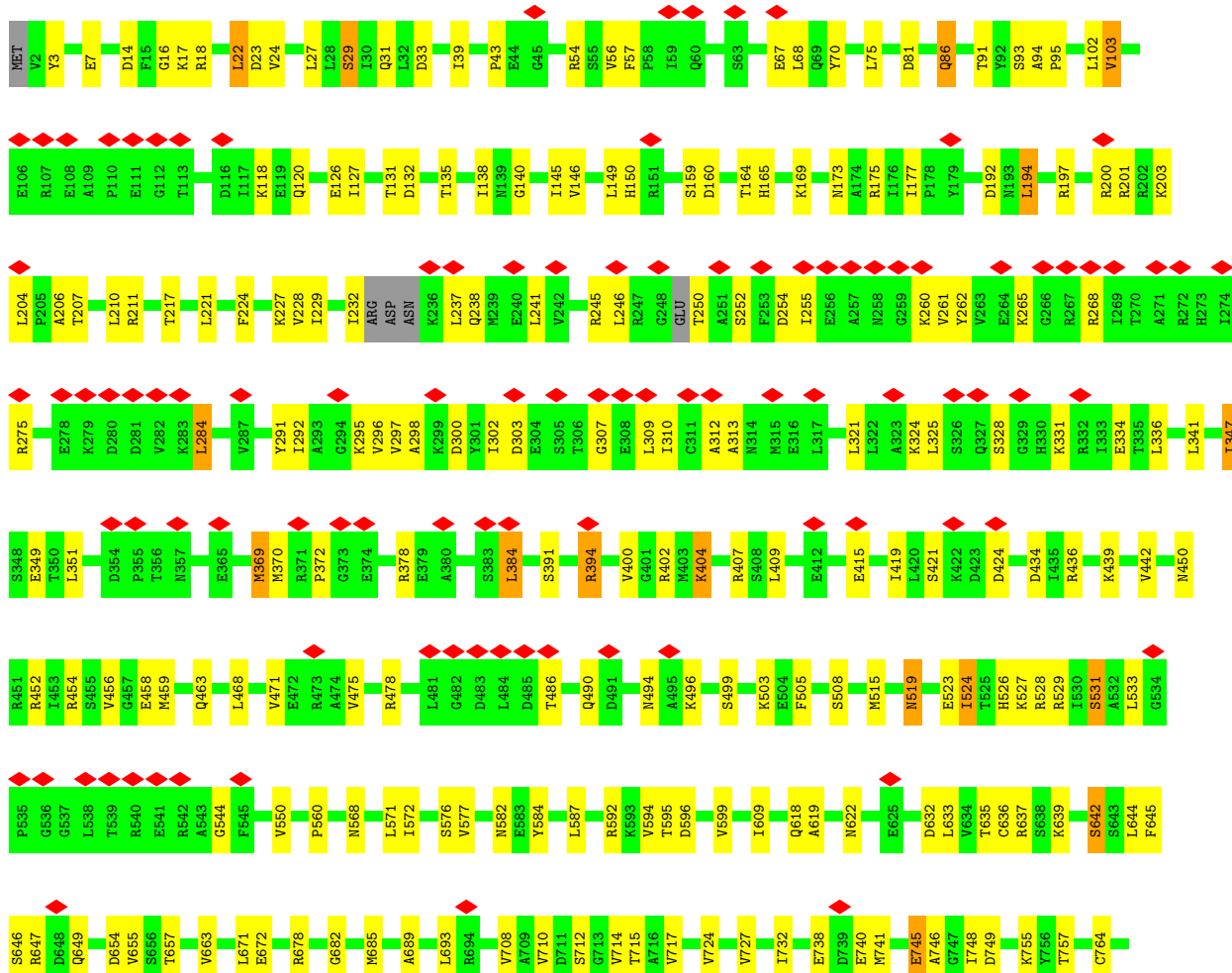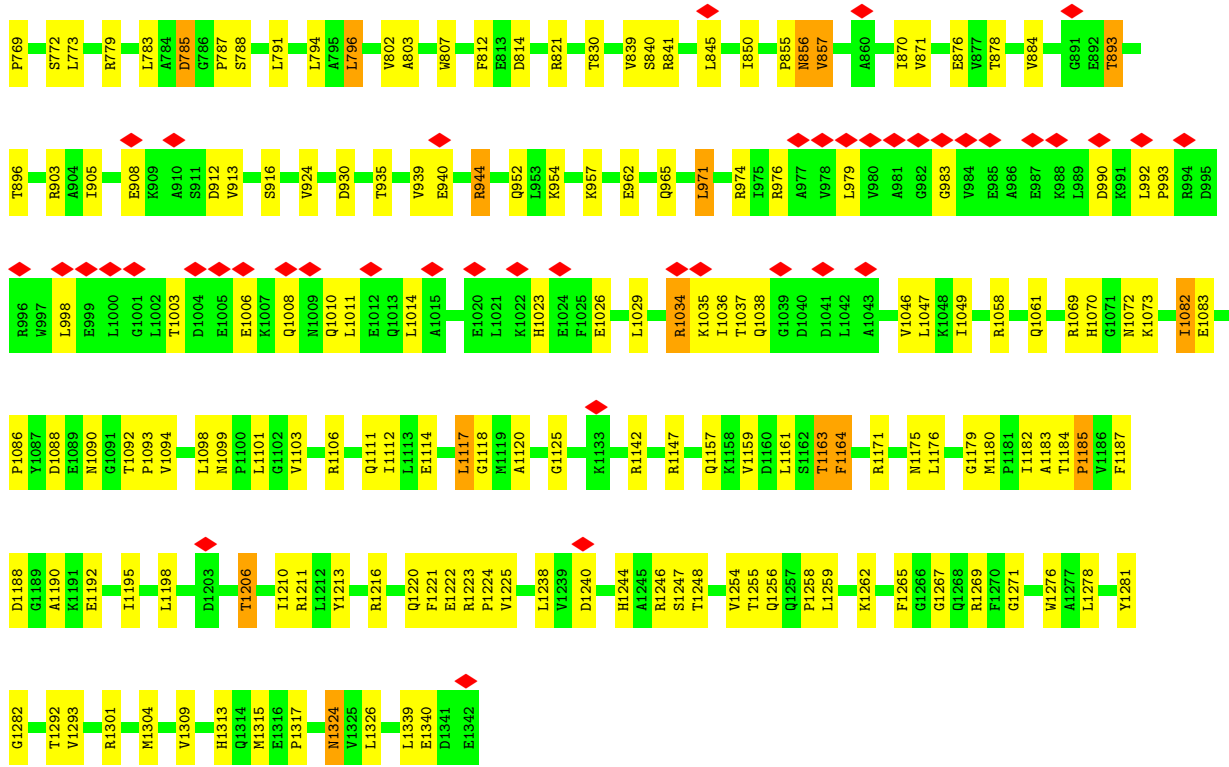
• Molecule 3: DNA-directed RNA polymerase subunit beta'
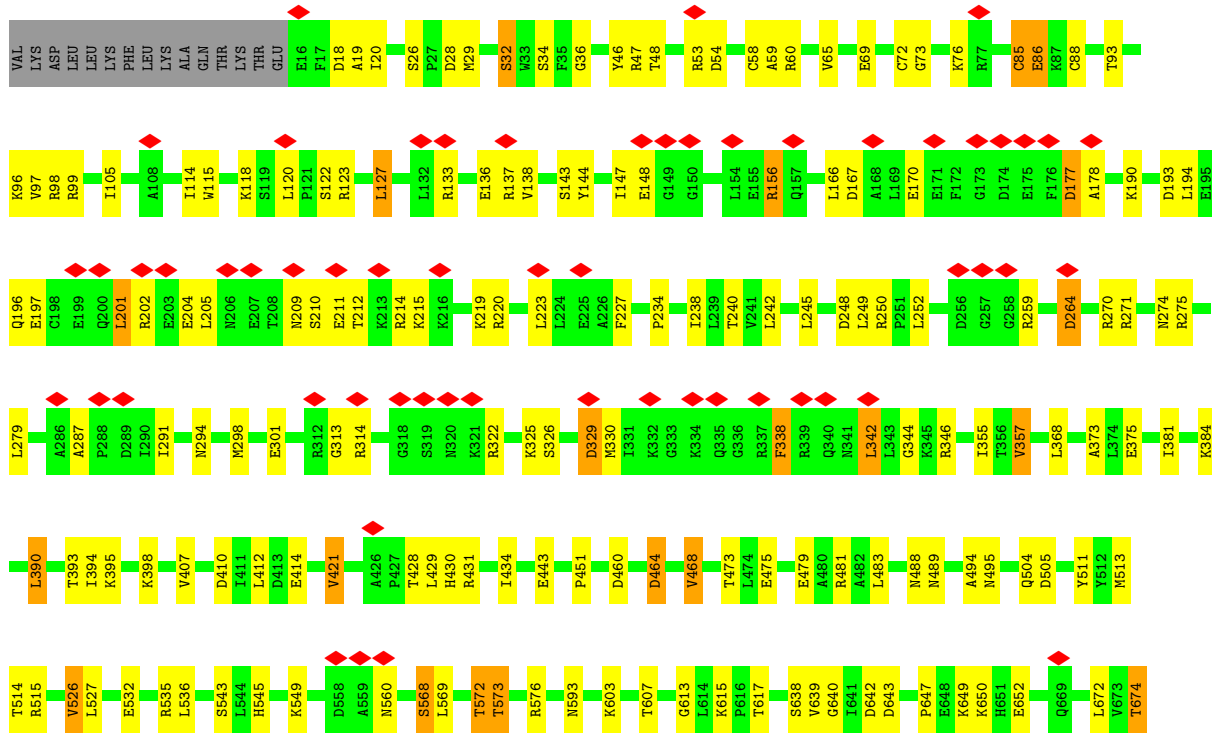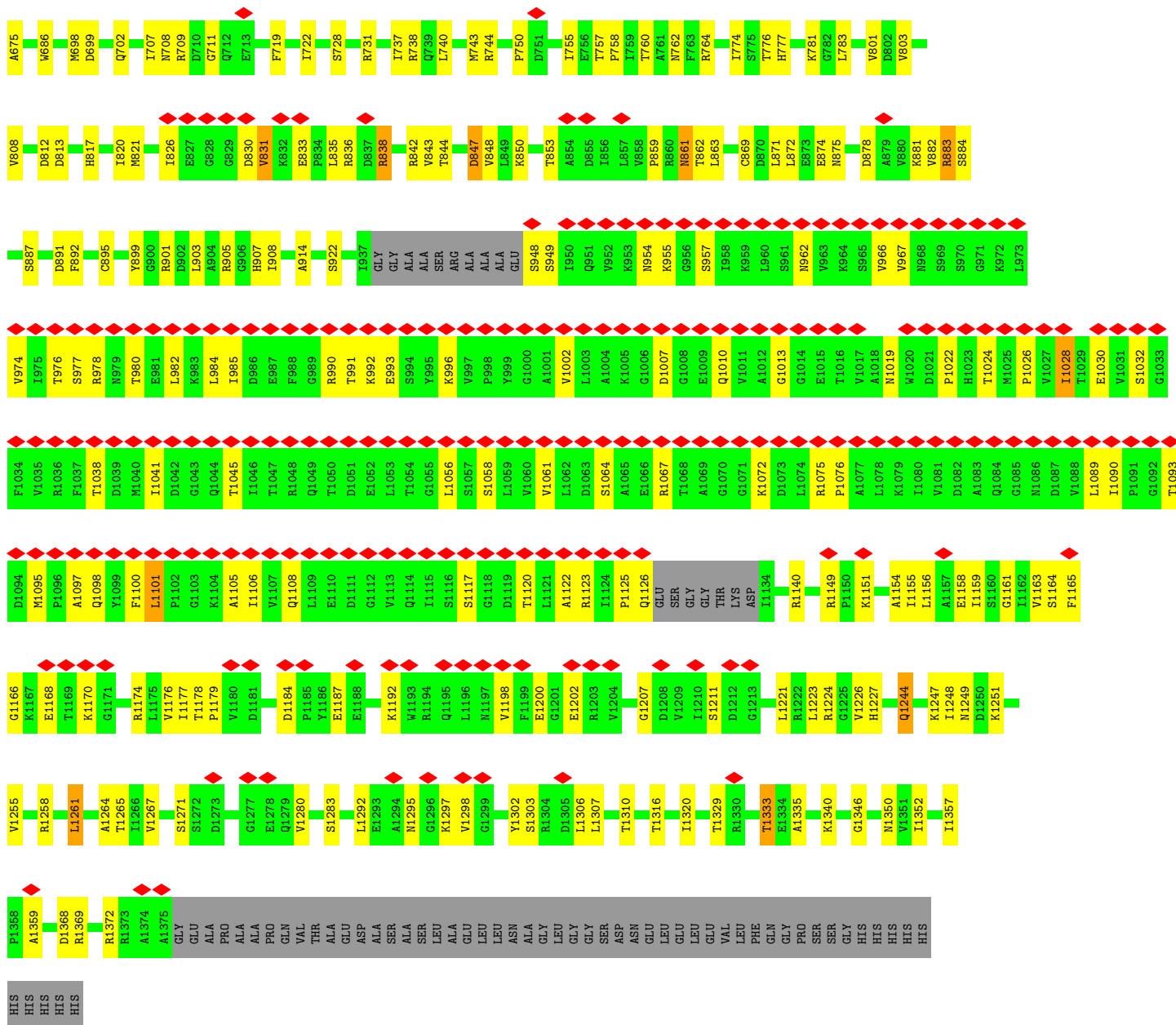
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain K:
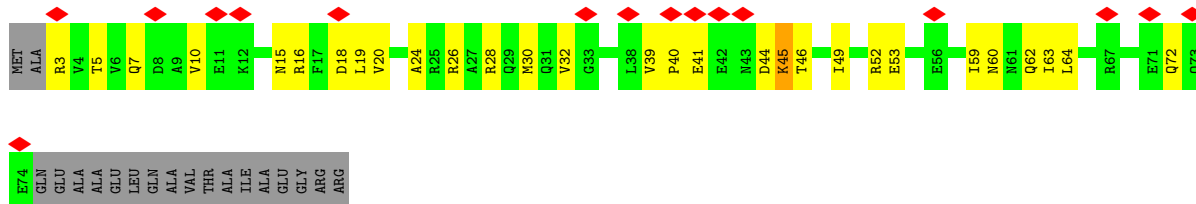
• Molecule 5: RNA polymerase sigma factor RpoD

Chain L:

45%

67%   22%   9%
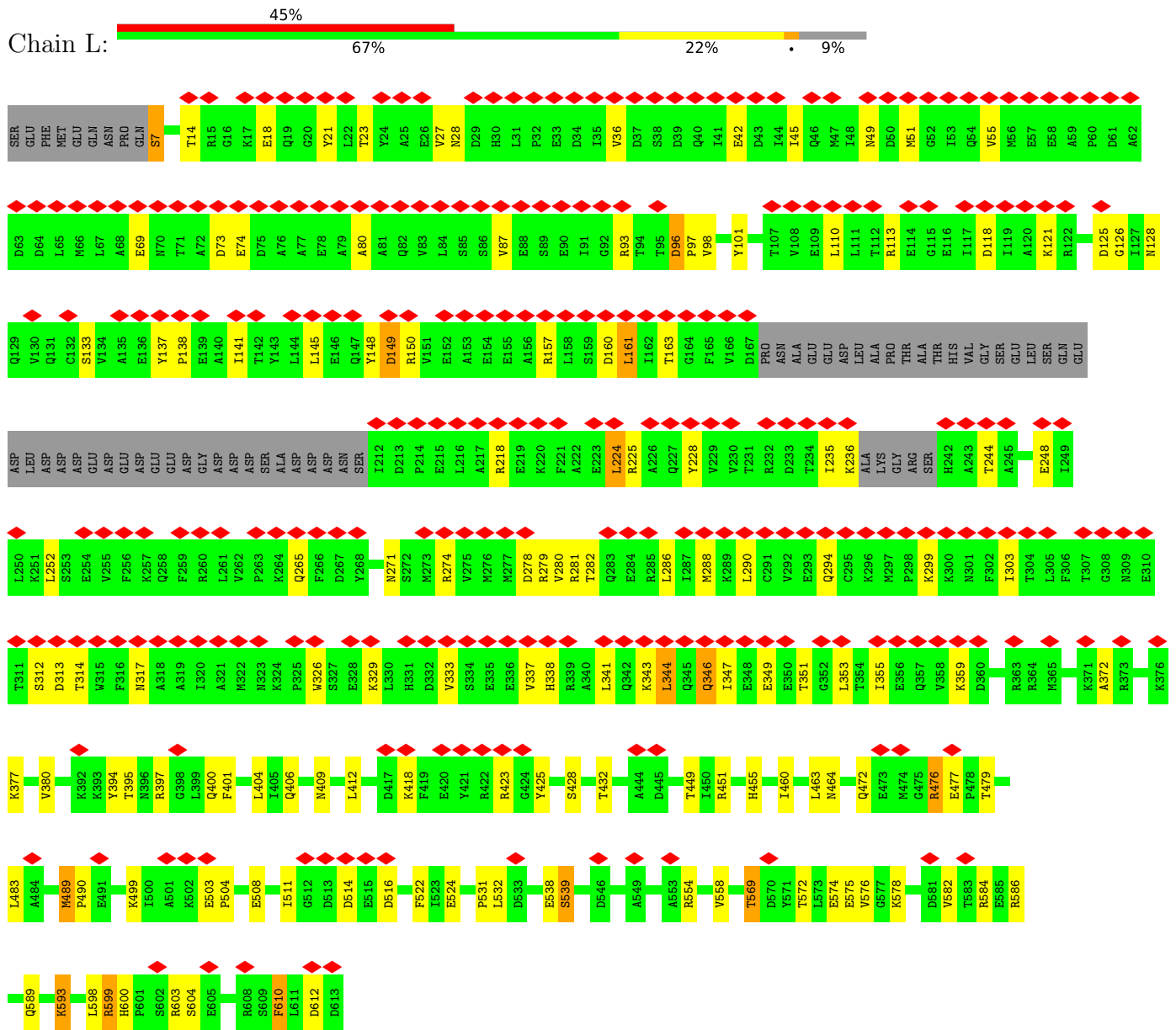
• Molecule 6: Protein TraR

Chain N:

10%

72%   25%
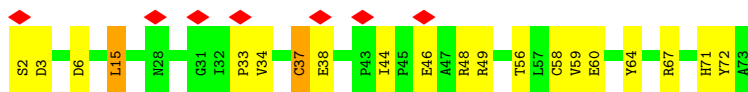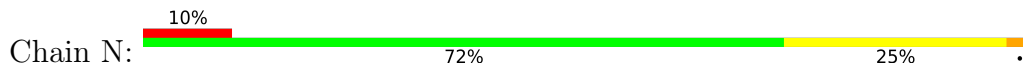
• Molecule 7: DNA (85-MER)

Chain O:

12%

32%   19%   49%
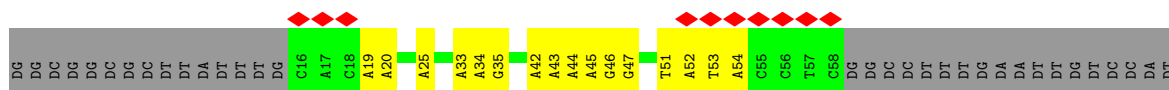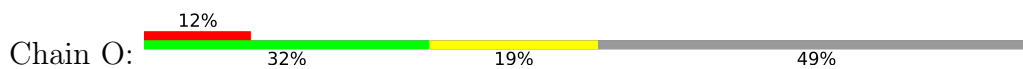
DA DT DA DG DA DC DG DC

• Molecule 8: DNA (85-MER)

Chain P:

12%

36% 14% 49%

DG DC DG DT DT DC DT DA DT DA DT DG DG DA DC DA DA DT DT DC DA DA DA DG DG DC G28 A29 G30 G31 A32 A33 T34 G35 C36 G37 T41 T44 A45 T52 G56 A59 G62 G63 T67 G68 T69 G70 DC DA DA DA
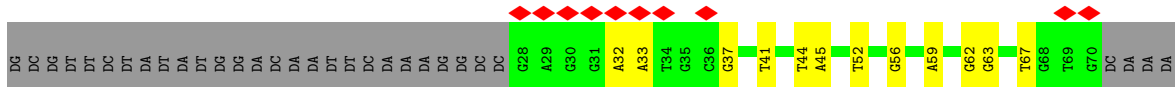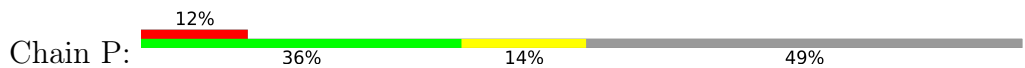
DT DA DA DG DC DG DC DG DC DC

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	56721	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 $(e^-/\text{Å}^2)$	50	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	0.123	Depositor
Minimum map value	-0.079	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.02	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 ⓘ

5.1   Standard geometry ⓘ

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	G	0.43	0/1791	0.60	0/2431
1	H	0.39	0/1688	0.59	0/2289
1	M	0.35	0/579	0.66	0/784
2	I	0.45	0/10667	0.60	3/14393 (0.0%)
3	J	0.43	0/10608	0.60	1/14323 (0.0%)
4	K	0.36	0/579	0.58	0/779
5	L	0.31	0/4546	0.58	4/6123 (0.1%)
6	N	0.42	0/575	0.63	0/777
7	O	0.71	0/994	1.00	1/1530 (0.1%)
8	P	0.69	0/984	1.08	2/1518 (0.1%)
All	All	0.44	0/33011	0.64	11/44947 (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	H	0	1
2	I	0	4
5	L	0	1
6	N	0	1
All	All	0	7

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	550	VAL	C-N-CA	6.74	138.56	121.70
2	I	998	LEU	CA-CB-CG	6.17	129.50	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	L	569	THR	C-N-CA	6.05	136.83	121.70
7	O	47	DG	O5'-P-OP1	-6.01	100.29	105.70
5	L	344	LEU	CA-CB-CG	5.76	128.55	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	78	ILE	Peptide
2	I	1185	PRO	Peptide
2	I	261	VAL	Peptide
2	I	857	VAL	Peptide
2	I	893	THR	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	G	1769	0	1789	34	0
1	H	1669	0	1698	39	0
1	M	572	0	602	11	0
2	I	10502	0	10510	204	0
3	J	10449	0	10673	213	0
4	K	577	0	588	19	0
5	L	4489	0	4506	90	0
6	N	566	0	553	11	0
7	O	883	0	481	11	0
8	P	881	0	491	8	0
9	J	1	0	0	0	0
10	J	2	0	0	0	0
10	N	1	0	0	0	0
All	All	32361	0	31891	570	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 9.

The worst 5 of 570 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)
5:L:149:ASP:OD1	5:L:225:ARG:NH2	1.66	1.26
5:L:149:ASP:CG	5:L:225:ARG:HH22	1.58	1.07
5:L:149:ASP:CG	5:L:225:ARG:NH2	2.15	0.95
5:L:45:ILE:O	5:L:49:ASN:HB2	1.75	0.87
2:I:29:SER:O	2:I:33:ASP:HB3	1.88	0.72

There are no symmetry-related clashes.

5.3   Torsion angles   ⓘ

5.3.1   Protein backbone   ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	230/329 (70%)	213 (93%)	17 (7%)	0	100	100
1	H	214/329 (65%)	193 (90%)	21 (10%)	0	100	100
1	M	71/329 (22%)	59 (83%)	12 (17%)	0	100	100
2	I	1331/1342 (99%)	1221 (92%)	105 (8%)	5 (0%)	34	72
3	J	1337/1430 (94%)	1240 (93%)	96 (7%)	1 (0%)	51	84
4	K	70/91 (77%)	69 (99%)	1 (1%)	0	100	100
5	L	552/616 (90%)	514 (93%)	37 (7%)	1 (0%)	47	81
6	N	70/72 (97%)	61 (87%)	8 (11%)	1 (1%)	11	46
All	All	3875/4538 (85%)	3570 (92%)	297 (8%)	8 (0%)	50	81

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	N	34	VAL
2	I	856	ASN
5	L	504	PRO
2	I	254	ASP
3	J	859	PRO

5.3.2   Protein sidechains   ⓘ

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	G	192/286 (67%)	175 (91%)	17 (9%)	9	37
1	H	184/286 (64%)	166 (90%)	18 (10%)	8	33
1	M	65/286 (23%)	57 (88%)	8 (12%)	4	23
2	I	1142/1157 (99%)	1039 (91%)	103 (9%)	9	37
3	J	1126/1189 (95%)	1026 (91%)	100 (9%)	9	37
4	K	63/75 (84%)	57 (90%)	6 (10%)	8	34
5	L	490/543 (90%)	456 (93%)	34 (7%)	15	47
6	N	60/61 (98%)	53 (88%)	7 (12%)	5	26
All	All	3322/3883 (86%)	3029 (91%)	293 (9%)	13	38

5 of 293 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	1267	VAL
6	N	15	LEU
4	K	5	THR
5	L	346	GLN
2	I	791	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
4	K	72	GLN
5	L	8	GLN
6	N	68	GLN
2	I	1116	HIS
2	I	1111	GLN

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 4 ligands modelled in this entry, 4 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 ⓘ

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-20462. 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
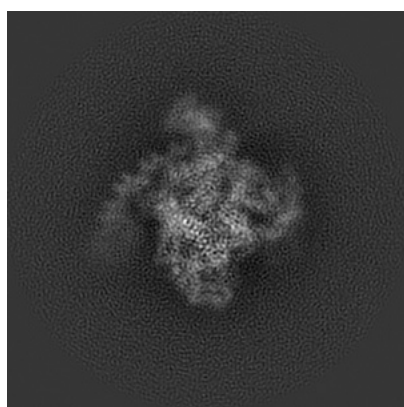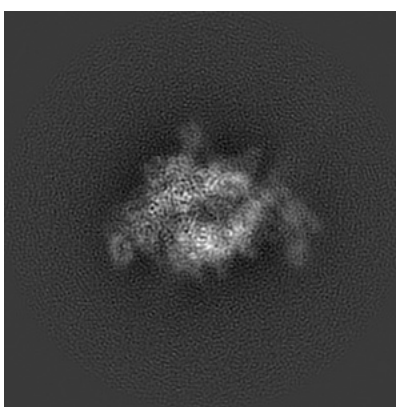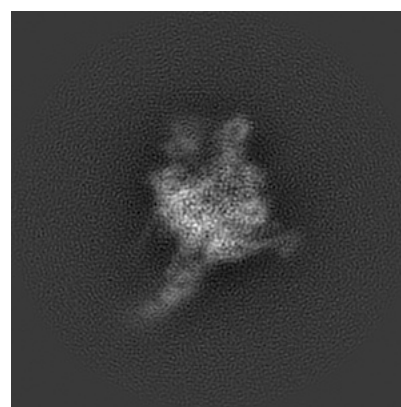
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ
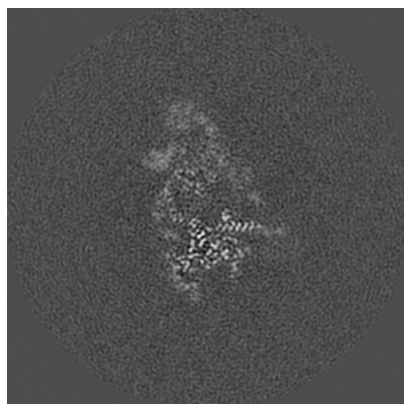
6.2.1   Primary map
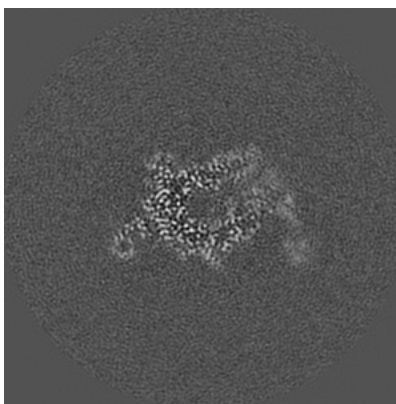
X Index: 128   Y Index: 128   Z Index: 128
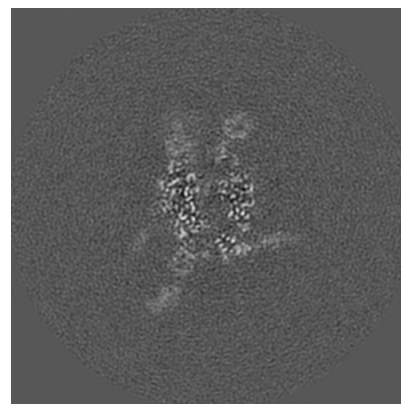
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map
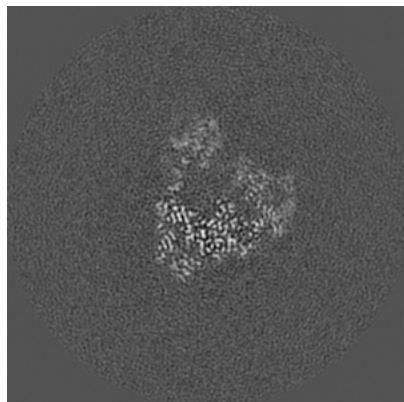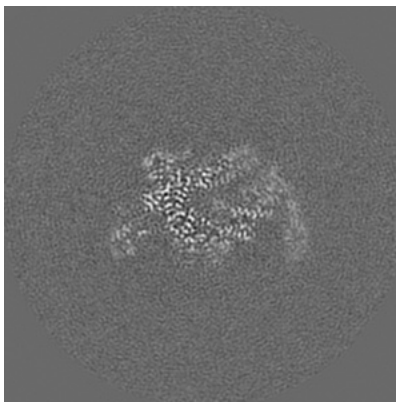
| X Index: 137 | Y Index: 124 | Z Index: 115 |
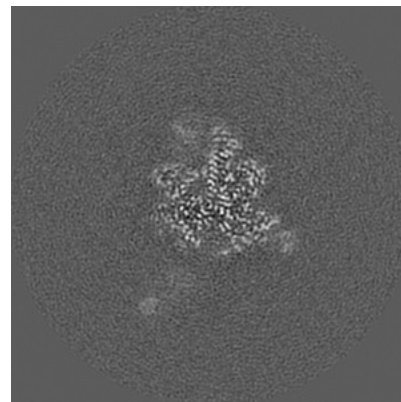
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

| X | Y | Z |
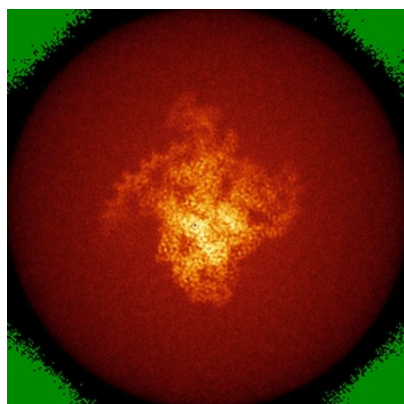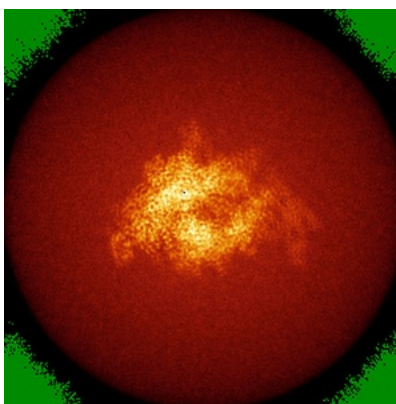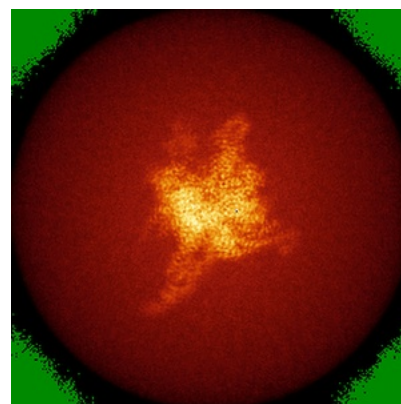
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map
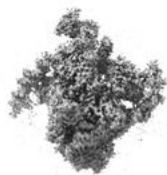
X   Y   Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. 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 ⓘ
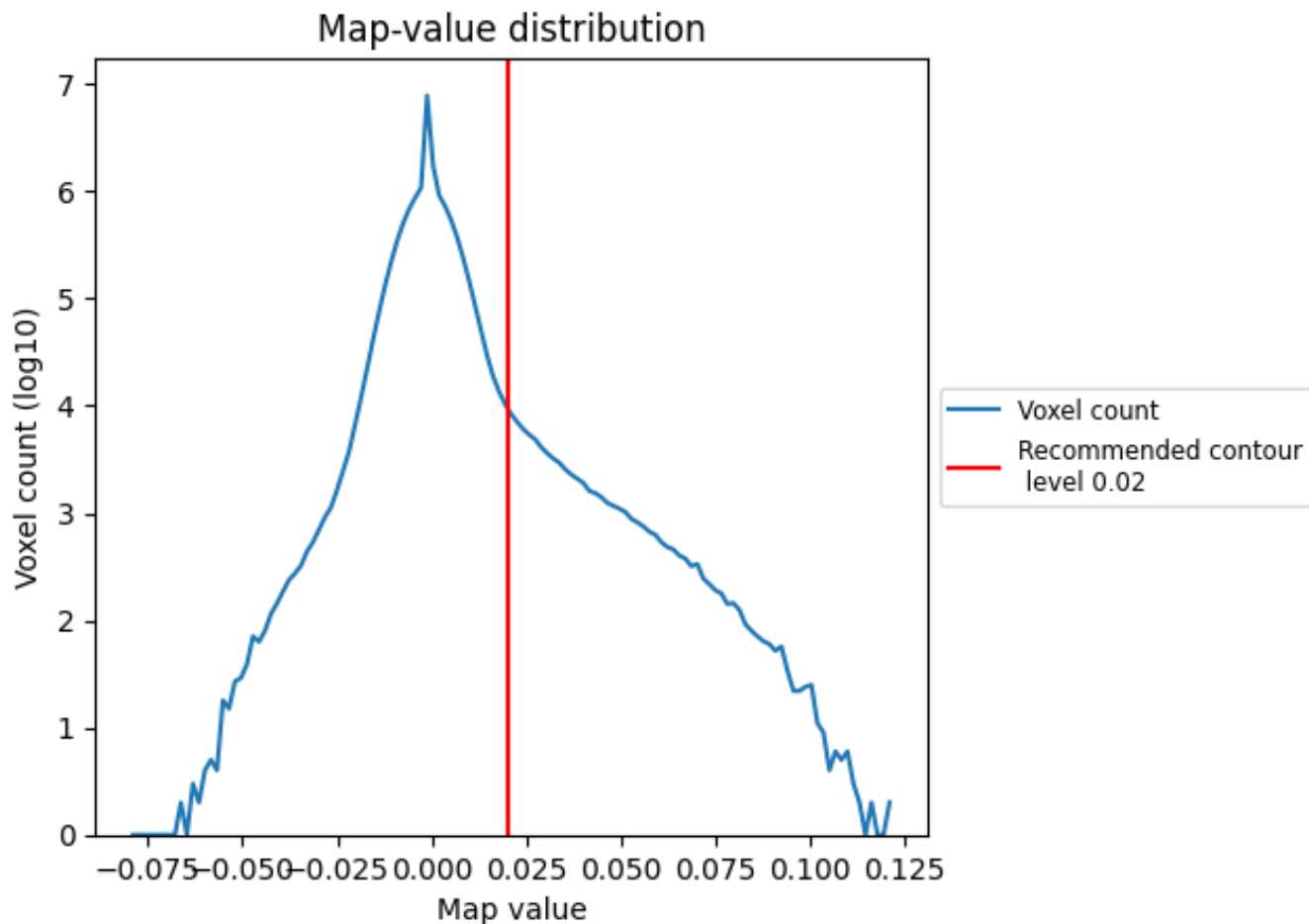
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ
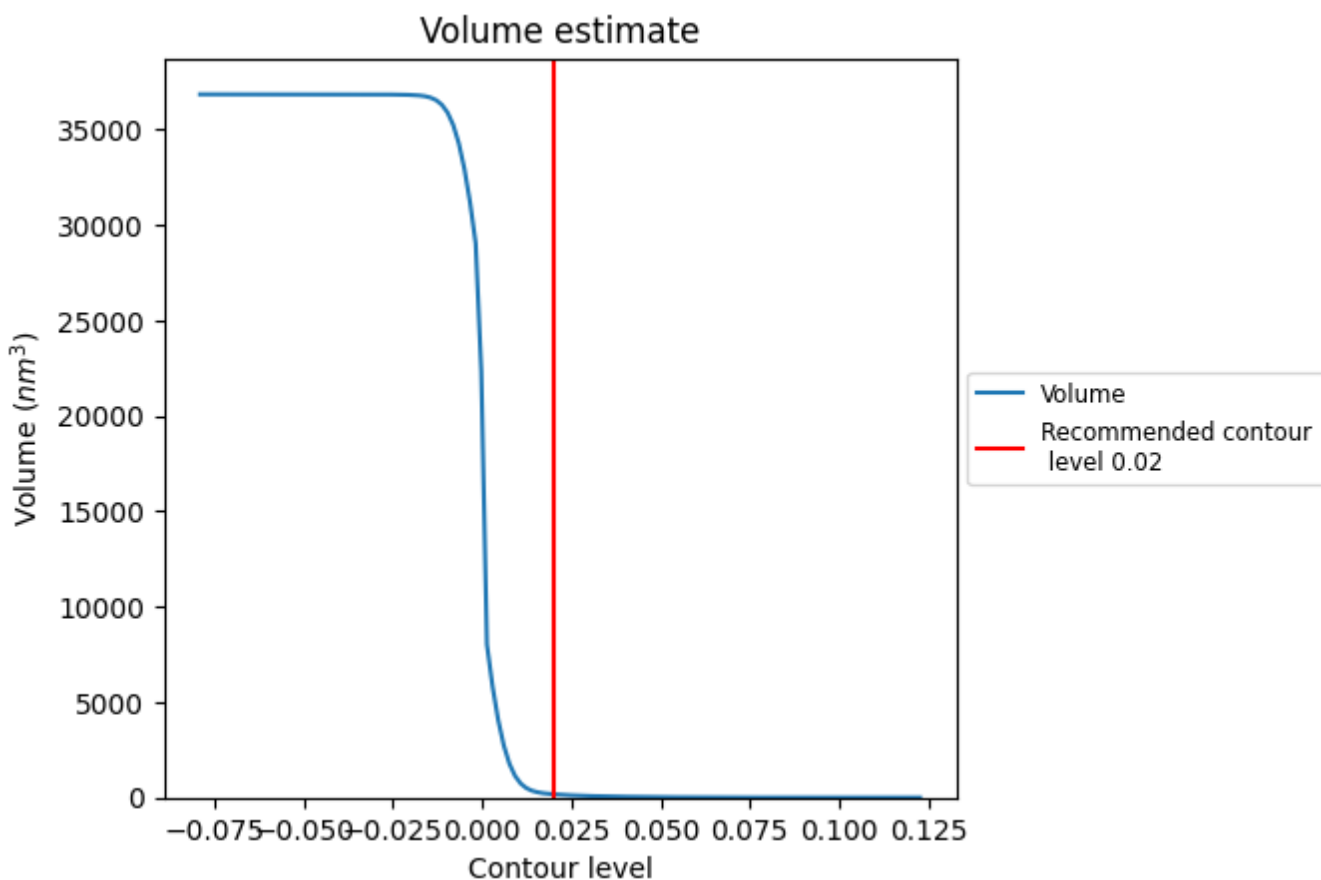
The volume at the recommended contour level is 173 nm^3; this corresponds to an approximate mass of 156 kDa.

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 ⓘ
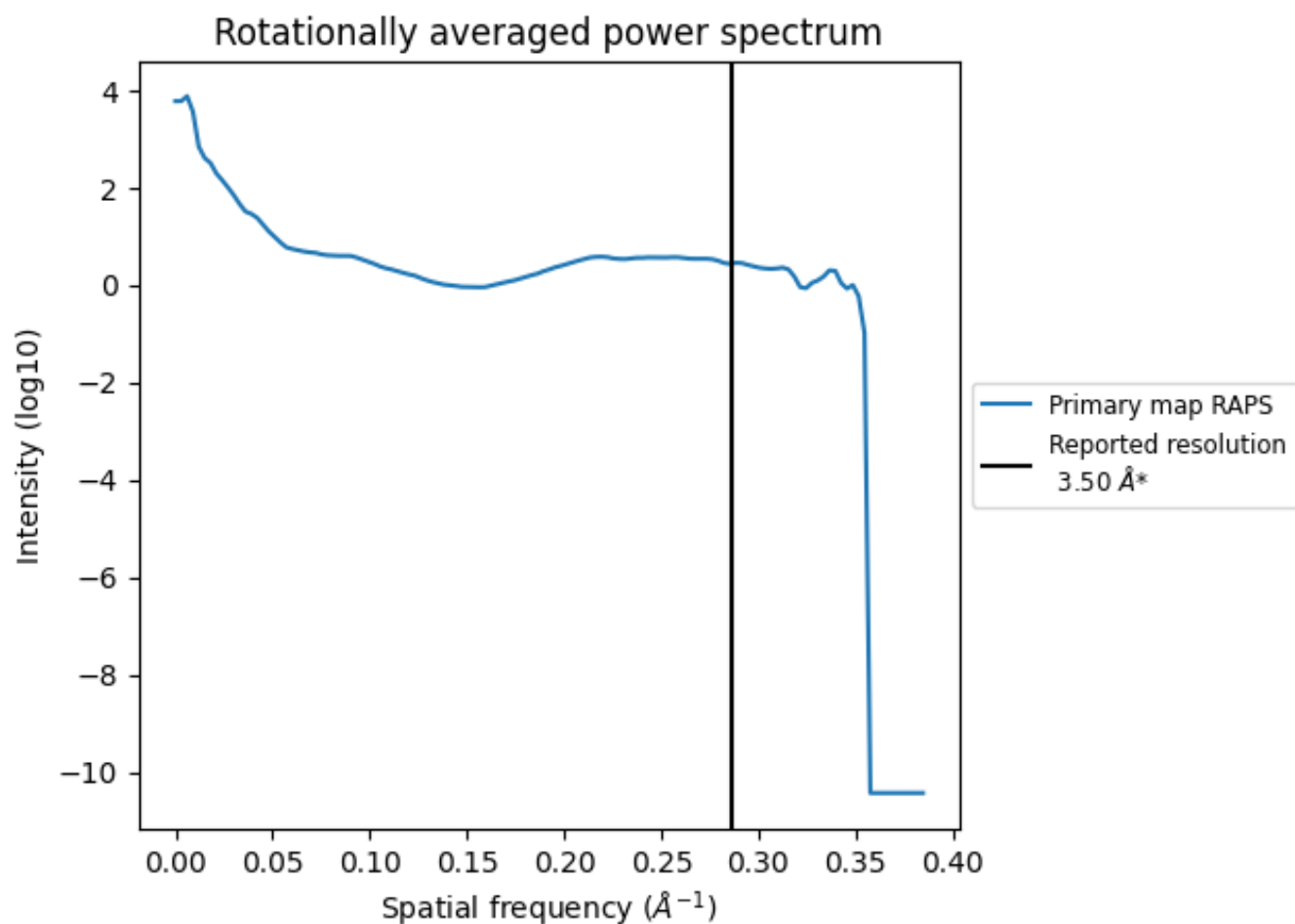
*Reported resolution corresponds to spatial frequency of 0.286 Å^{-1}

8   Fourier-Shell correlation ⓘ

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1   FSC ⓘ
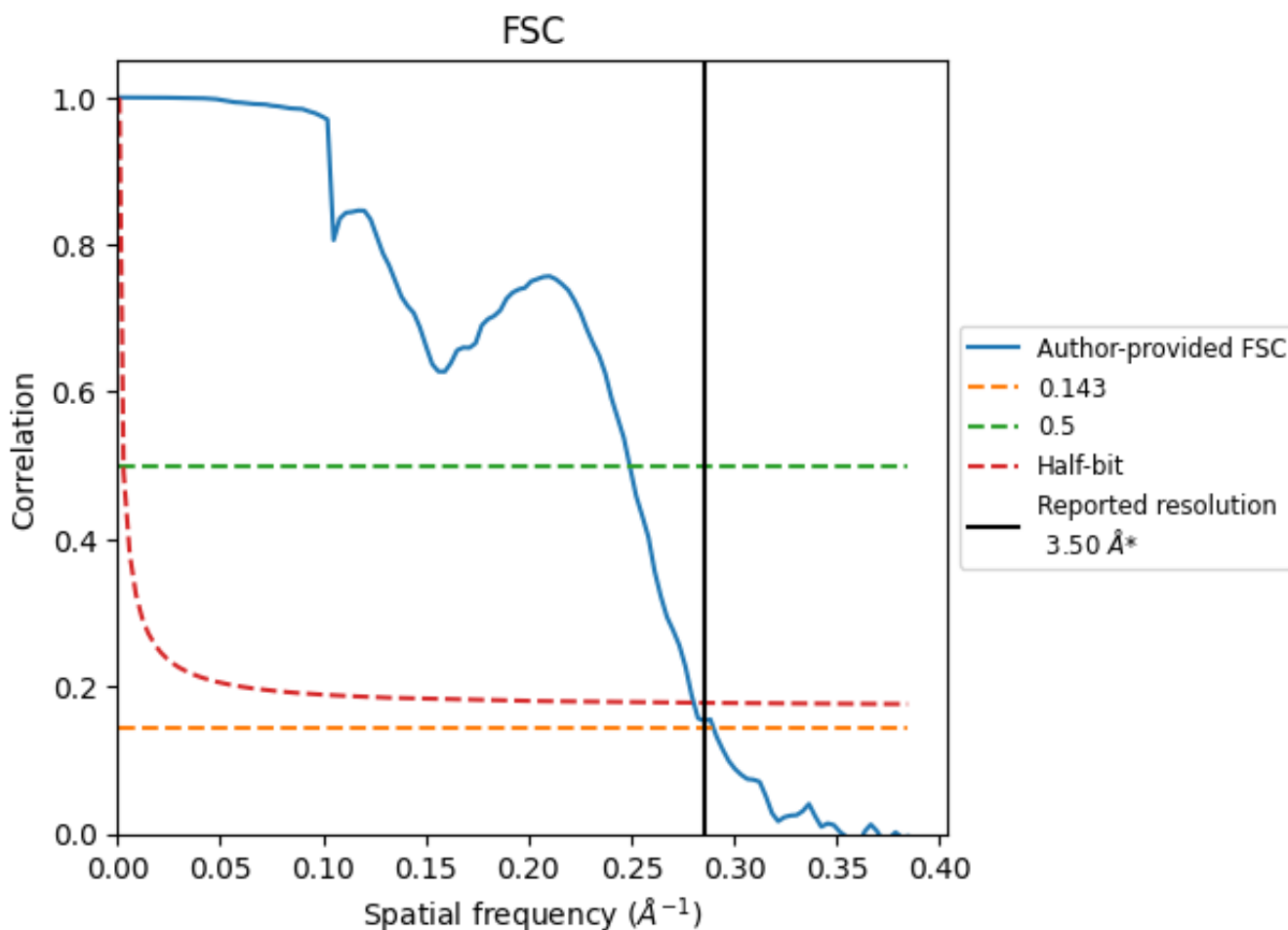
*Reported resolution corresponds to spatial frequency of 0.286 Å^{-1}

8.2   Resolution estimates ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.50	-	-
Author-provided FSC curve	3.45	4.01	3.57
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-20462 and PDB model 6PSS. Per-residue inclusion information can be found in section 3 on page 7.

9.1 Map-model overlay (i)

| X | Y | Z |
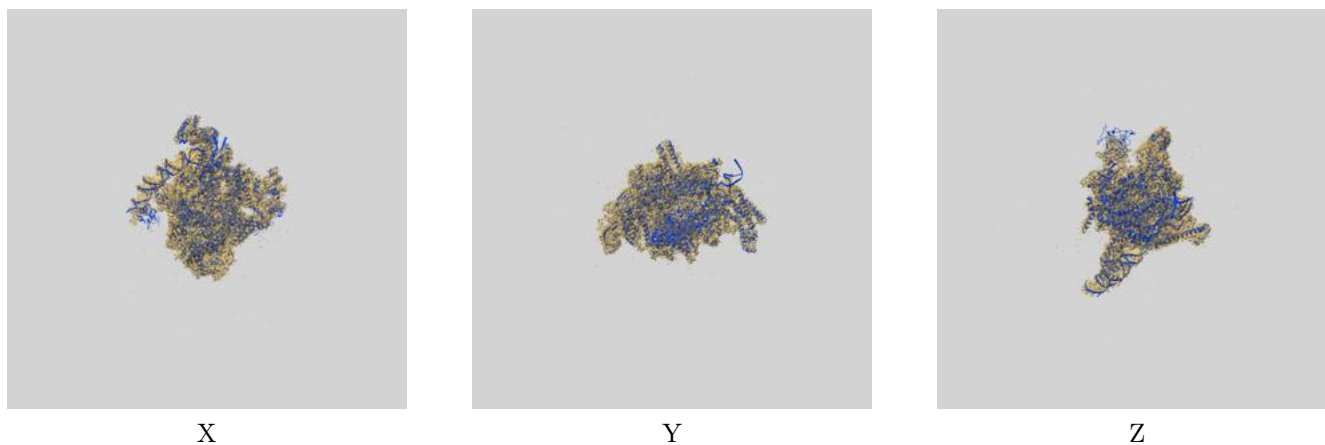
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)
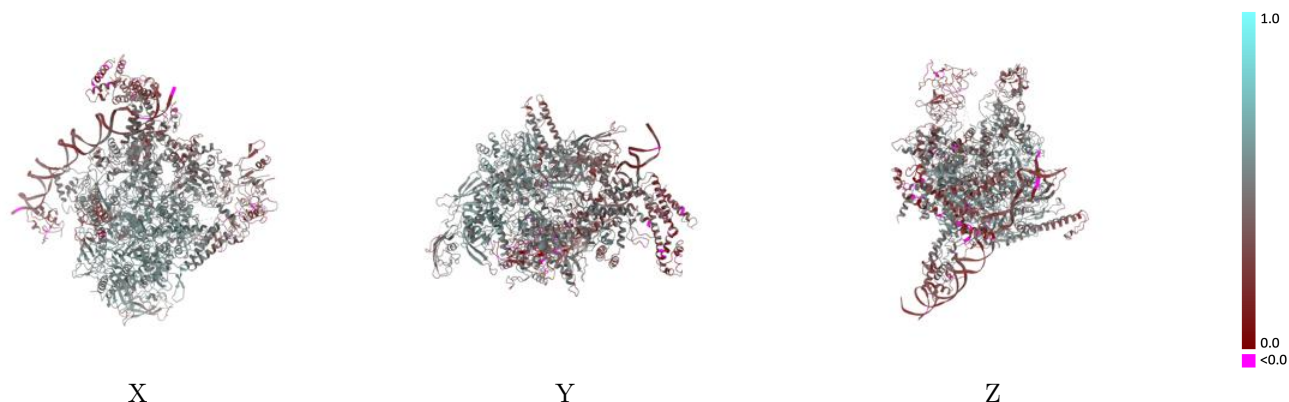
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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)
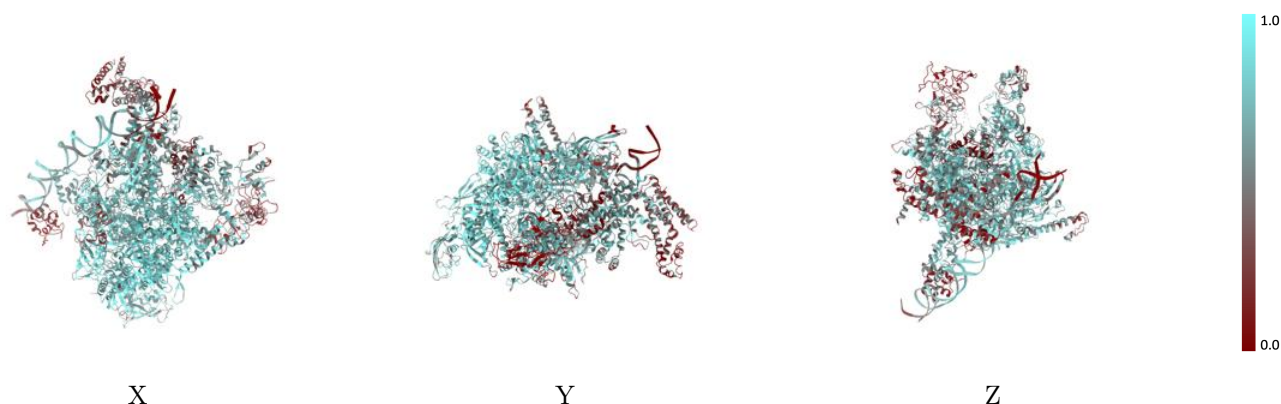
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.02).

9.4　Atom inclusion ⓘ
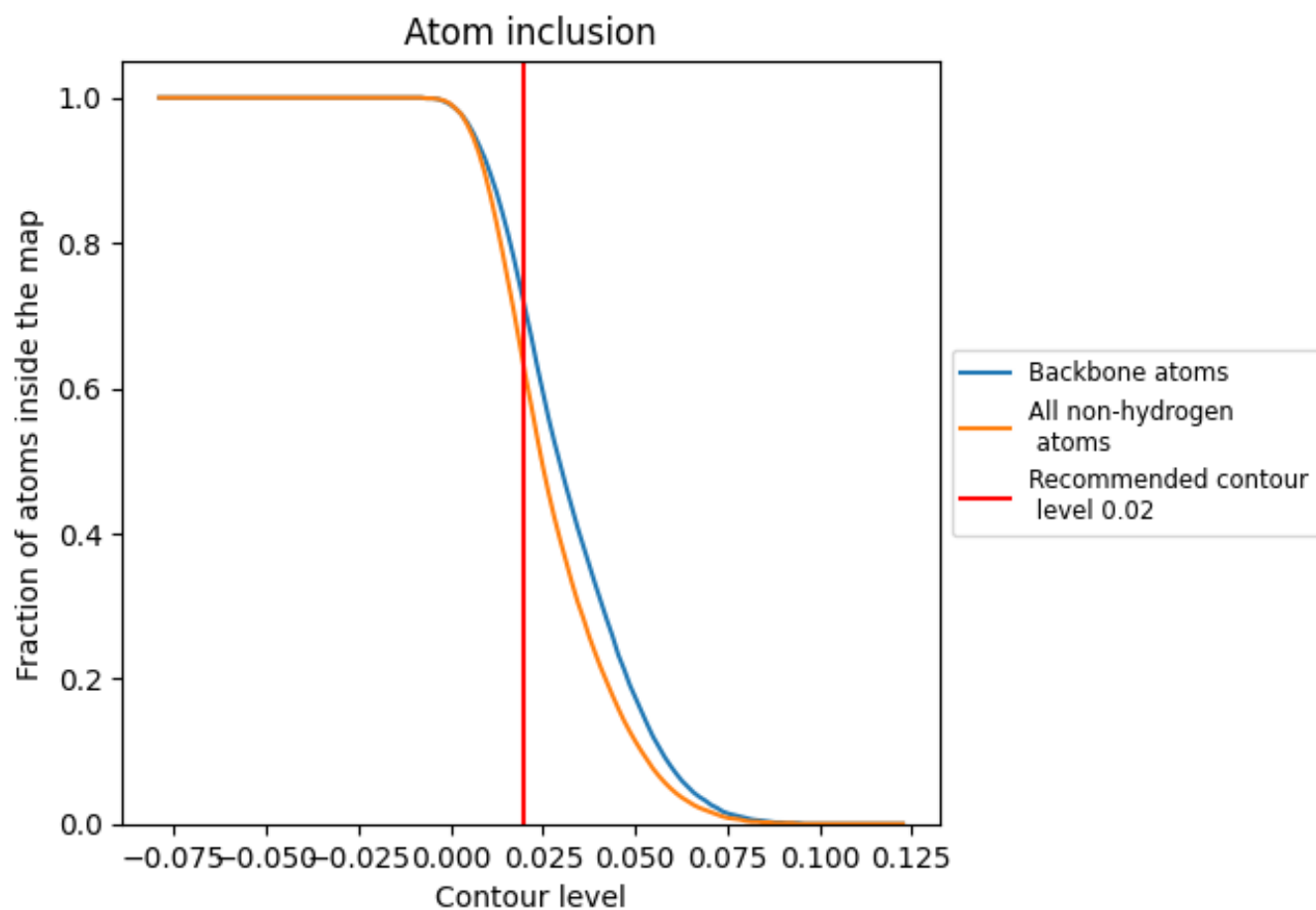
At the recommended contour level, 71% of all backbone atoms, 63% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6260	0.4460
G	0.7660	0.5190
H	0.7060	0.4790
I	0.7180	0.4930
J	0.6360	0.4620
K	0.5940	0.4850
L	0.4020	0.3450
M	0.0610	0.2650
N	0.7040	0.4730
O	0.5720	0.2720
P	0.5050	0.2640

1.0

0.0
<0.0