



wwPDB EM Validation Summary Report ⓘ

Nov 21, 2022 – 12:15 AM JST

PDB ID : 7D7D
EMDB ID : EMD-30605
Title : CryoEM structure of gp45-dependent transcription activation complex
Authors : Shi, J.; Wen, A.; Jin, S.; Feng, Y.
Deposited on : 2020-10-03
Resolution : 4.50 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.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) : 2.31.3

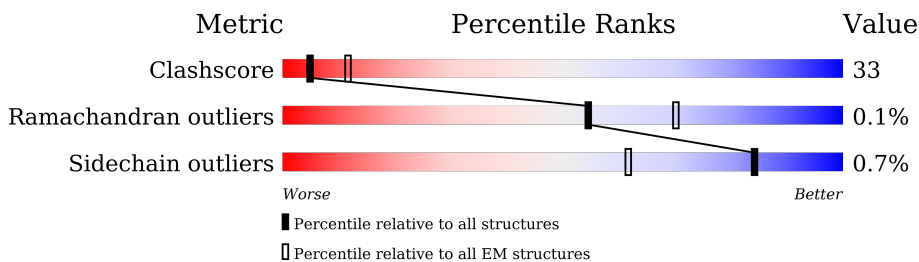
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.50 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	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 $\leq 5\%$. 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	329	
1	B	329	
2	C	1342	
3	D	1407	
4	T	59	
5	F	185	
6	N	59	
7	K	132	

Continued on next page...

Mol	Chain	Length	Quality of chain
8	E	91	 8% 33% 45% 22%
9	G	236	 8% 42% 55%
9	H	236	 8% 36% 60%
9	I	236	 22% 45% 51%

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 33755 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	A	219	Total	C	N	O	S	0	0
			1686	1056	298	326	6		
1	B	219	Total	C	N	O	S	0	0
			1689	1056	298	329	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0
			10567	6631	1841	2052	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1335	Total	C	N	O	S	0	0
			10384	6524	1851	1959	50		

- Molecule 4 is a DNA chain called DNA (template strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	T	46	Total	C	N	O	P	0	0
			939	451	158	284	46		

- Molecule 5 is a protein called gp55.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	137	Total	C	N	O	S	0	0
			1128	720	189	214	5		

- Molecule 6 is a DNA chain called DNA (nontemplate strand).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	N	48	Total	C	N	O	P	0	0
			990	471	192	279	48		

- Molecule 7 is a protein called RNA polymerase-associated protein Gp33.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	71	Total	C	N	O	S	0	0
			558	356	89	110	3		

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-19	MET	-	initiating methionine	UNP P13338
K	-18	GLY	-	expression tag	UNP P13338
K	-17	SER	-	expression tag	UNP P13338
K	-16	SER	-	expression tag	UNP P13338
K	-15	HIS	-	expression tag	UNP P13338
K	-14	HIS	-	expression tag	UNP P13338
K	-13	HIS	-	expression tag	UNP P13338
K	-12	HIS	-	expression tag	UNP P13338
K	-11	HIS	-	expression tag	UNP P13338
K	-10	HIS	-	expression tag	UNP P13338
K	-9	SER	-	expression tag	UNP P13338
K	-8	SER	-	expression tag	UNP P13338
K	-7	GLY	-	expression tag	UNP P13338
K	-6	LEU	-	expression tag	UNP P13338
K	-5	VAL	-	expression tag	UNP P13338
K	-4	PRO	-	expression tag	UNP P13338
K	-3	ARG	-	expression tag	UNP P13338
K	-2	GLY	-	expression tag	UNP P13338
K	-1	SER	-	expression tag	UNP P13338
K	0	HIS	-	expression tag	UNP P13338

- Molecule 8 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	E	71	Total	C	N	O	S	0	0
			564	345	108	110	1		

- Molecule 9 is a protein called DNA polymerase clamp.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	G	228	Total	C	N	O	S	0	0
			1749	1113	288	342	6		
9	H	228	Total	C	N	O	S	0	0
			1749	1113	288	342	6		
9	I	228	Total	C	N	O	S	0	0
			1749	1113	288	342	6		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	229	LEU	-	expression tag	UNP P04525
G	230	GLU	-	expression tag	UNP P04525
G	231	HIS	-	expression tag	UNP P04525
G	232	HIS	-	expression tag	UNP P04525
G	233	HIS	-	expression tag	UNP P04525
G	234	HIS	-	expression tag	UNP P04525
G	235	HIS	-	expression tag	UNP P04525
G	236	HIS	-	expression tag	UNP P04525
H	229	LEU	-	expression tag	UNP P04525
H	230	GLU	-	expression tag	UNP P04525
H	231	HIS	-	expression tag	UNP P04525
H	232	HIS	-	expression tag	UNP P04525
H	233	HIS	-	expression tag	UNP P04525
H	234	HIS	-	expression tag	UNP P04525
H	235	HIS	-	expression tag	UNP P04525
H	236	HIS	-	expression tag	UNP P04525
I	229	LEU	-	expression tag	UNP P04525
I	230	GLU	-	expression tag	UNP P04525
I	231	HIS	-	expression tag	UNP P04525
I	232	HIS	-	expression tag	UNP P04525
I	233	HIS	-	expression tag	UNP P04525
I	234	HIS	-	expression tag	UNP P04525
I	235	HIS	-	expression tag	UNP P04525
I	236	HIS	-	expression tag	UNP P04525

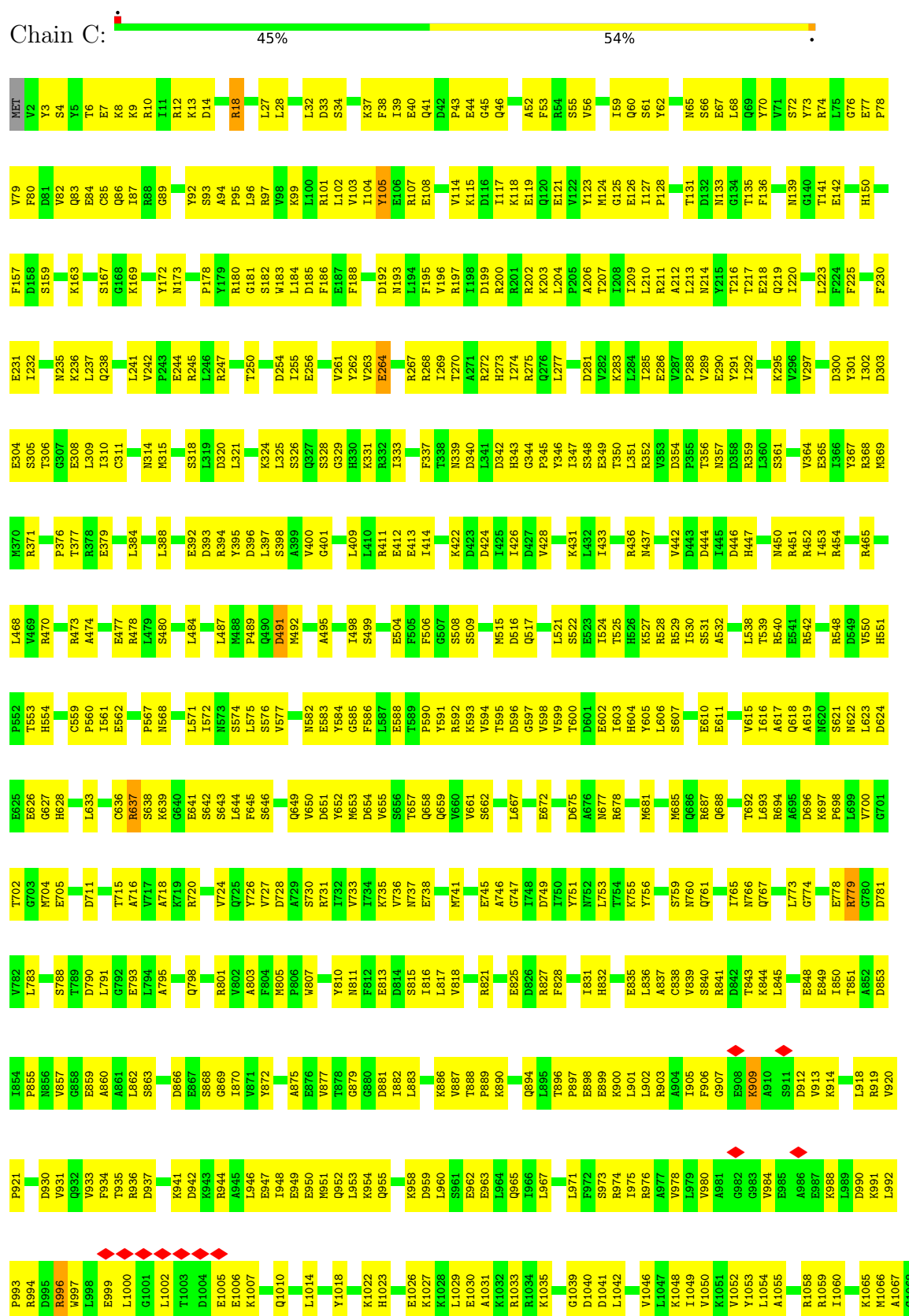
- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

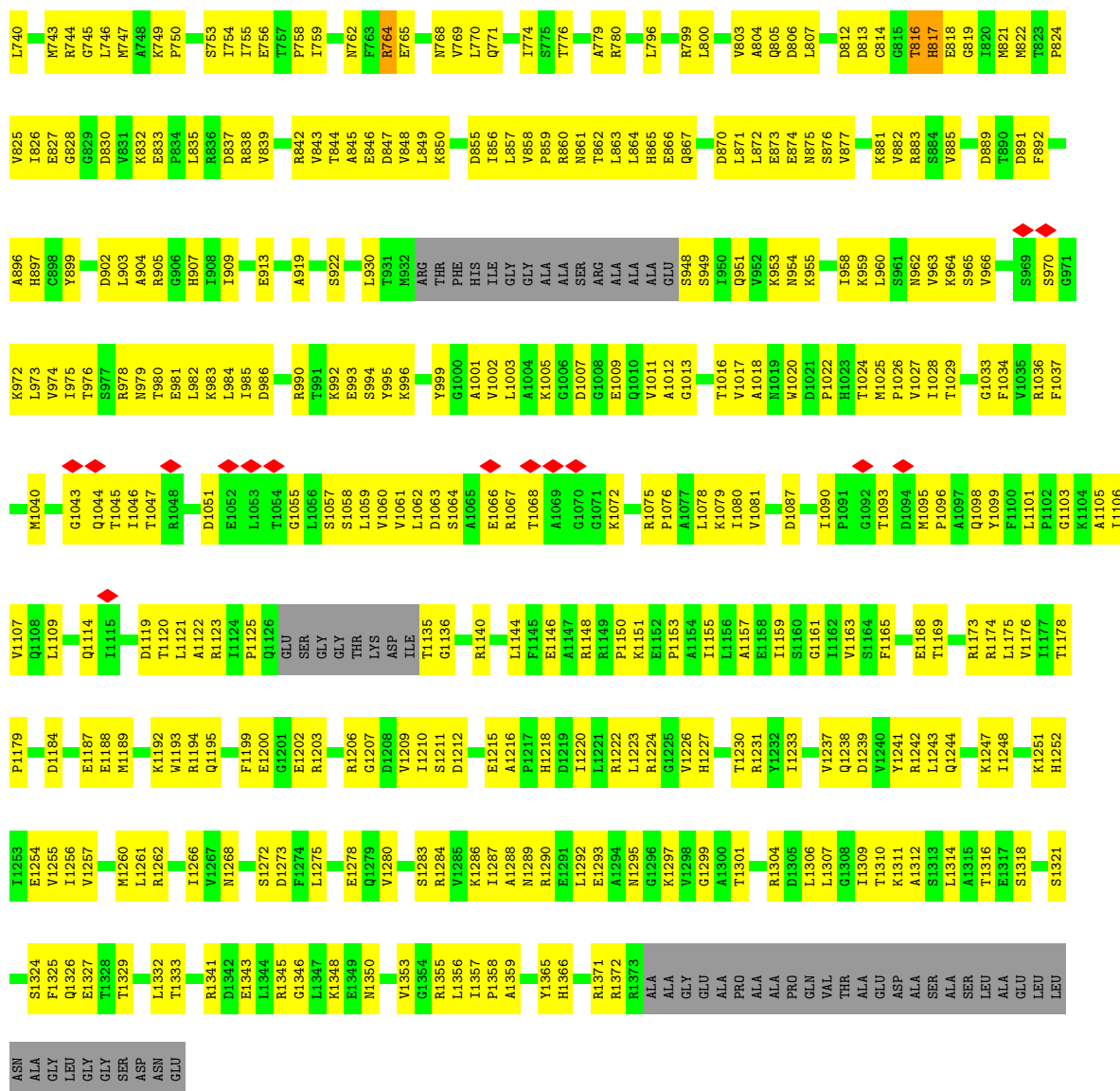
Mol	Chain	Residues	Atoms		AltConf
10	D	1	Total	Mg	0
			1	1	

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

• Molecule 2: DNA-directed RNA polymerase subunit beta





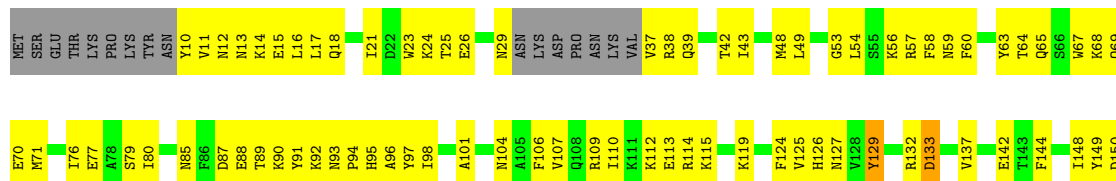
• Molecule 4: DNA (template strand)

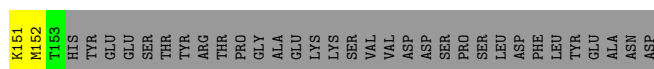
Chain T: 19% 59% 22%



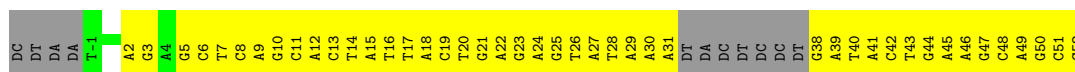
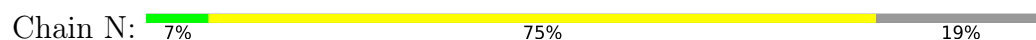
• Molecule 5: gp55

Chain F: 31% 42% 26%

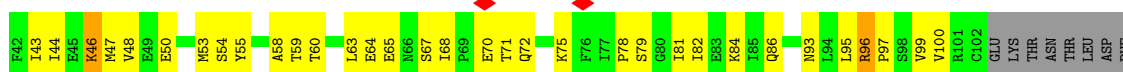
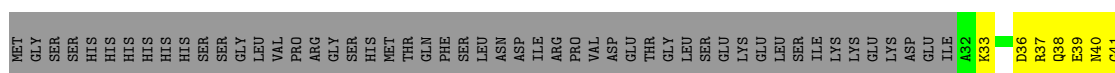
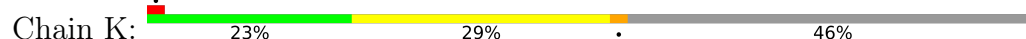




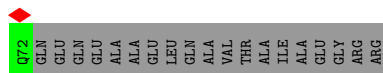
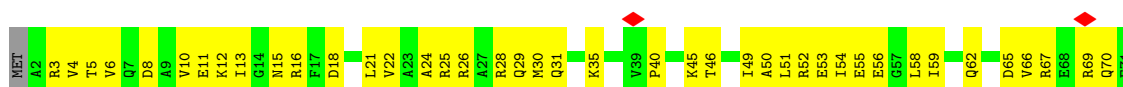
- Molecule 6: DNA (nontemplate strand)



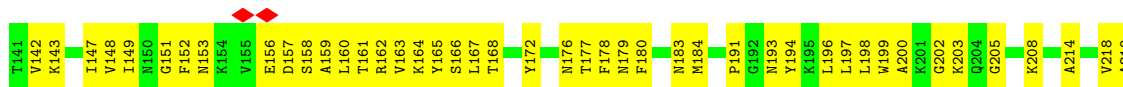
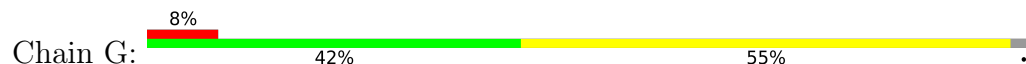
- Molecule 7: RNA polymerase-associated protein Gp33



- Molecule 8: DNA-directed RNA polymerase subunit omega

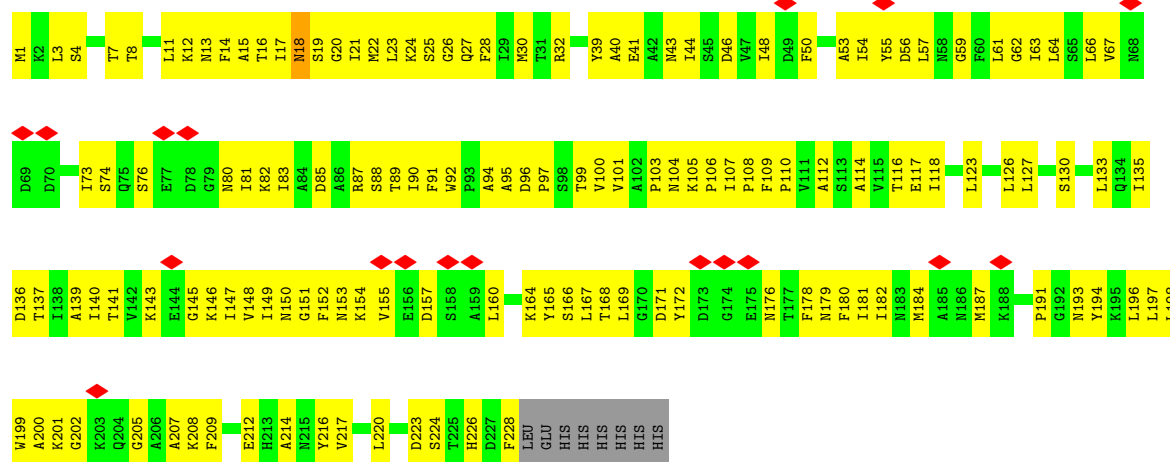


- Molecule 9: DNA polymerase clamp

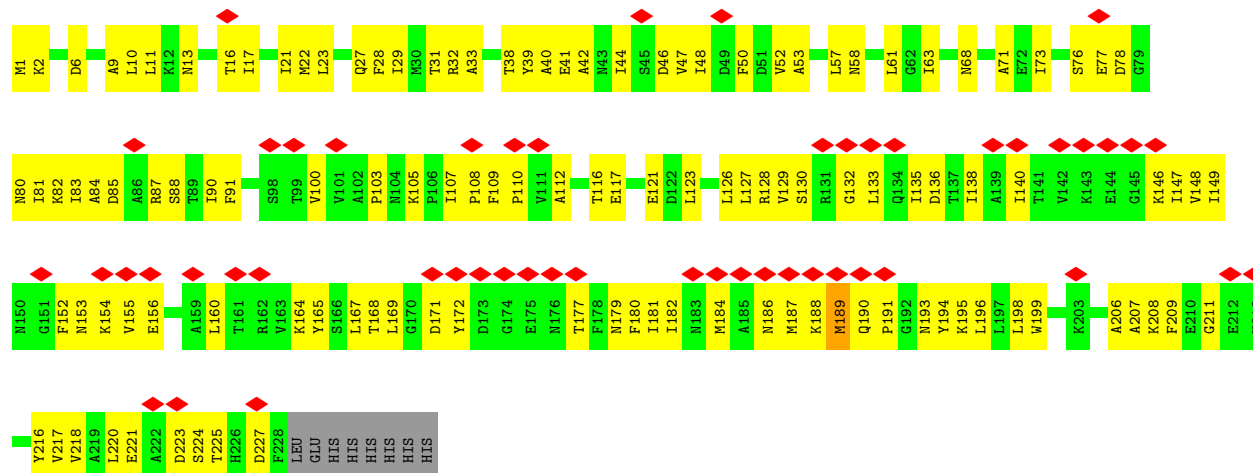
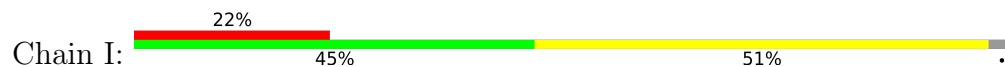




• Molecule 9: DNA polymerase clamp



• Molecule 9: DNA polymerase clamp



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	8981	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 ($e^-/\text{\AA}^2$)	59	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	0.287	Depositor
Minimum map value	-0.180	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	334.592, 334.592, 334.592	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.307, 1.307, 1.307	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	0.45	0/1706	0.58	0/2312
1	B	0.41	0/1708	0.56	0/2315
2	C	0.47	0/10736	0.55	0/14487
3	D	0.44	0/10541	0.55	0/14232
4	T	0.94	0/1048	1.10	0/1613
5	F	0.47	0/1149	0.54	0/1546
6	N	0.89	0/1113	1.01	0/1713
7	K	0.34	0/566	0.48	0/763
8	E	0.32	0/566	0.50	0/762
9	G	0.31	0/1778	0.51	0/2410
9	H	0.34	0/1778	0.51	0/2410
9	I	0.31	0/1778	0.52	0/2410
All	All	0.48	0/34467	0.59	0/46973

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1686	0	1726	112	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1689	0	1725	142	0
2	C	10567	0	10585	720	0
3	D	10384	0	10607	677	0
4	T	939	0	526	54	0
5	F	1128	0	1105	80	0
6	N	990	0	540	75	0
7	K	558	0	550	45	0
8	E	564	0	579	35	0
9	G	1749	0	1755	138	0
9	H	1749	0	1755	171	0
9	I	1749	0	1755	102	0
10	D	1	0	0	0	0
11	D	2	0	0	0	0
All	All	33755	0	33208	2193	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 33.

The worst 5 of 2193 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)
9:H:141:THR:CG2	9:H:179:ASN:HA	1.61	1.29
3:D:1161:GLY:HA3	3:D:1178:THR:O	1.37	1.21
9:H:141:THR:HG22	9:H:179:ASN:HA	1.33	1.09
1:B:33:ARG:NH1	2:C:1081:PRO:HG3	1.68	1.08
5:F:10:TYR:HD1	5:F:49:LEU:HD11	1.19	1.06

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	A	215/329 (65%)	175 (81%)	40 (19%)	0	100	100
1	B	215/329 (65%)	188 (87%)	27 (13%)	0	100	100
2	C	1338/1342 (100%)	1096 (82%)	241 (18%)	1 (0%)	51	85
3	D	1329/1407 (94%)	1093 (82%)	235 (18%)	1 (0%)	51	85
5	F	133/185 (72%)	113 (85%)	19 (14%)	1 (1%)	19	60
7	K	69/132 (52%)	58 (84%)	11 (16%)	0	100	100
8	E	69/91 (76%)	59 (86%)	10 (14%)	0	100	100
9	G	226/236 (96%)	209 (92%)	17 (8%)	0	100	100
9	H	226/236 (96%)	211 (93%)	15 (7%)	0	100	100
9	I	226/236 (96%)	209 (92%)	17 (8%)	0	100	100
All	All	4046/4523 (90%)	3411 (84%)	632 (16%)	3 (0%)	54	85

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	F	129	TYR
3	D	764	ARG
2	C	45	GLY

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	A	186/286 (65%)	183 (98%)	3 (2%)	62	79
1	B	187/286 (65%)	186 (100%)	1 (0%)	88	93
2	C	1155/1157 (100%)	1144 (99%)	11 (1%)	76	86
3	D	1120/1168 (96%)	1116 (100%)	4 (0%)	91	94
5	F	121/166 (73%)	119 (98%)	2 (2%)	60	78
7	K	61/120 (51%)	59 (97%)	2 (3%)	38	61
8	E	61/75 (81%)	61 (100%)	0	100	100
9	G	189/197 (96%)	189 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	H	189/197 (96%)	188 (100%)	1 (0%)	88	93
9	I	189/197 (96%)	188 (100%)	1 (0%)	88	93
All	All	3458/3849 (90%)	3433 (99%)	25 (1%)	84	90

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1291	LEU
3	D	816	THR
9	I	189	MET
3	D	250	ARG
3	D	817	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 69 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	127	ASN
7	K	66	ASN
9	H	213	HIS
2	C	1134	GLN
2	C	1116	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

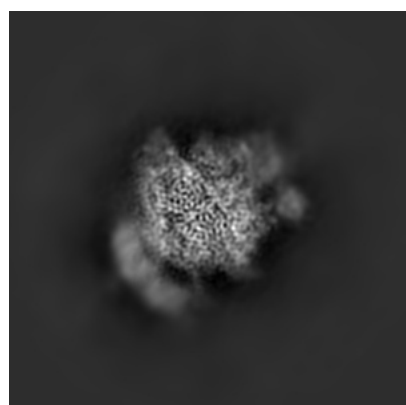
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-30605. These allow visual inspection of the internal detail of the map and identification of artifacts.

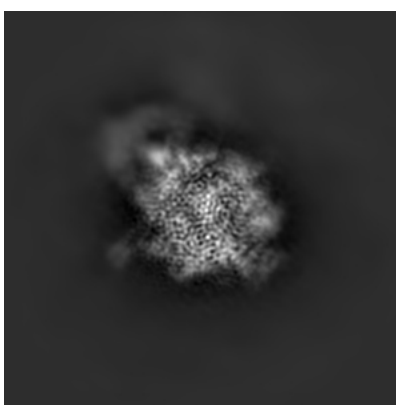
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

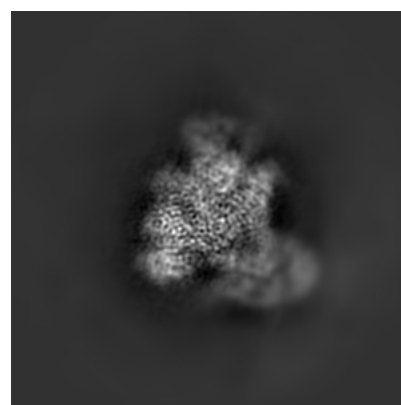
6.1.1 Primary map



X



Y

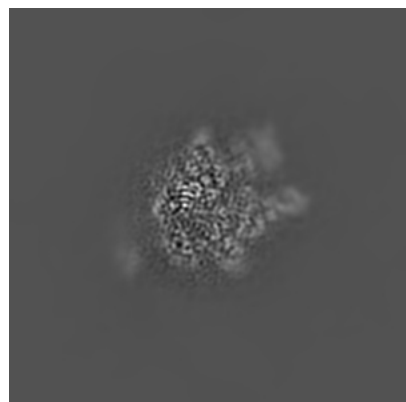


Z

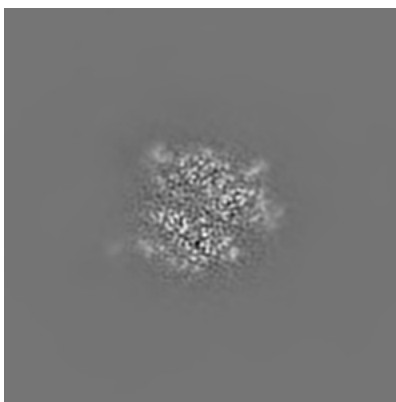
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

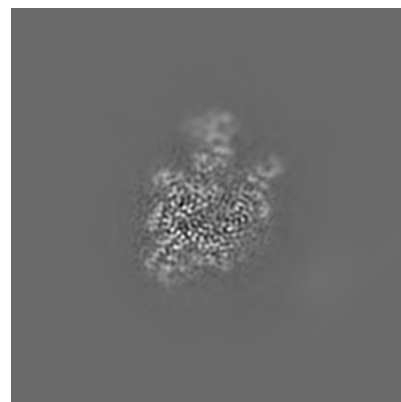
6.2.1 Primary map



X Index: 128



Y Index: 128

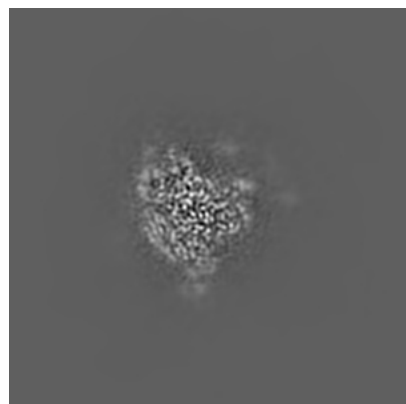


Z Index: 128

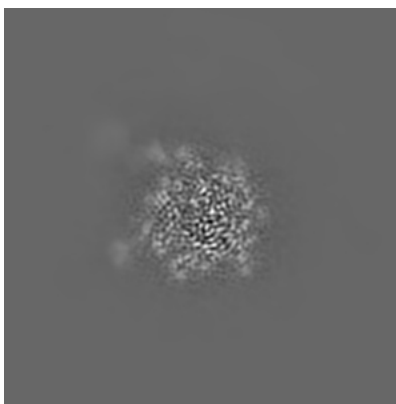
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

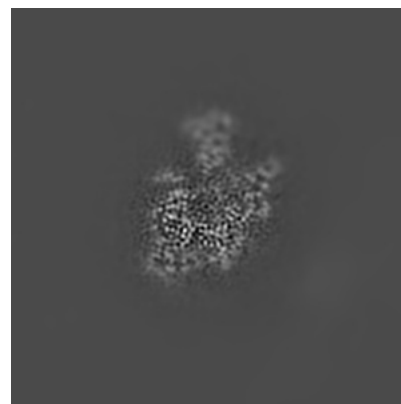
6.3.1 Primary map



X Index: 107



Y Index: 114

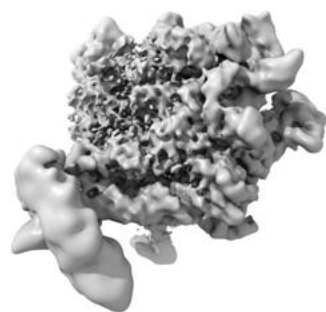


Z Index: 130

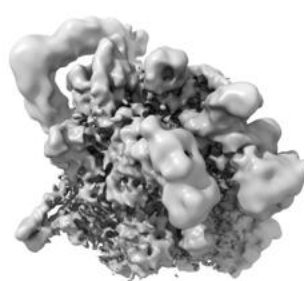
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views [i](#)

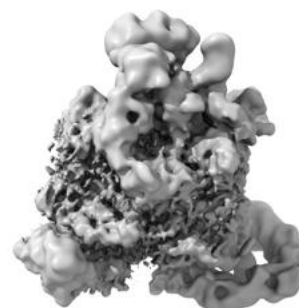
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

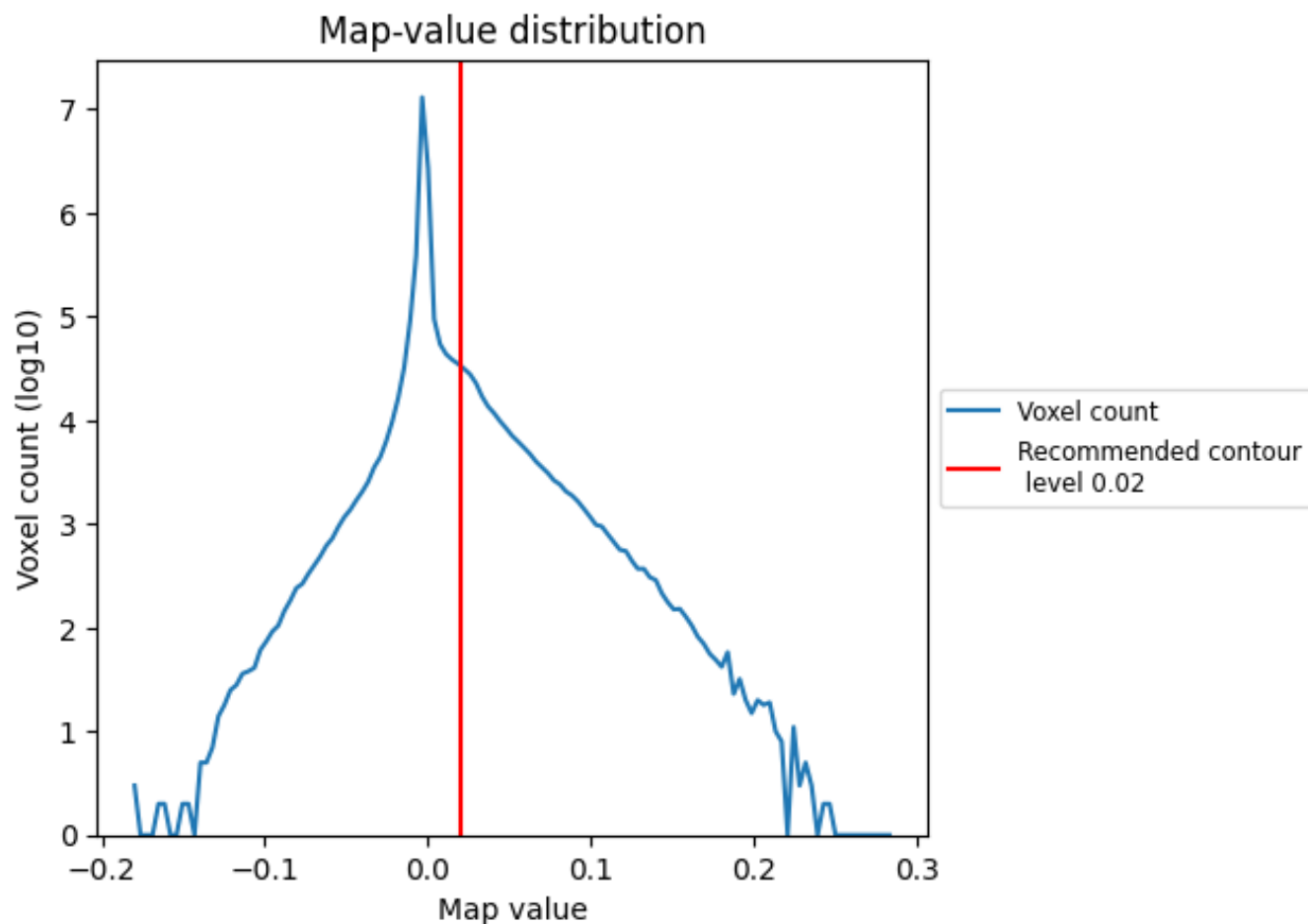
6.5 Mask visualisation

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

7 Map analysis [i](#)

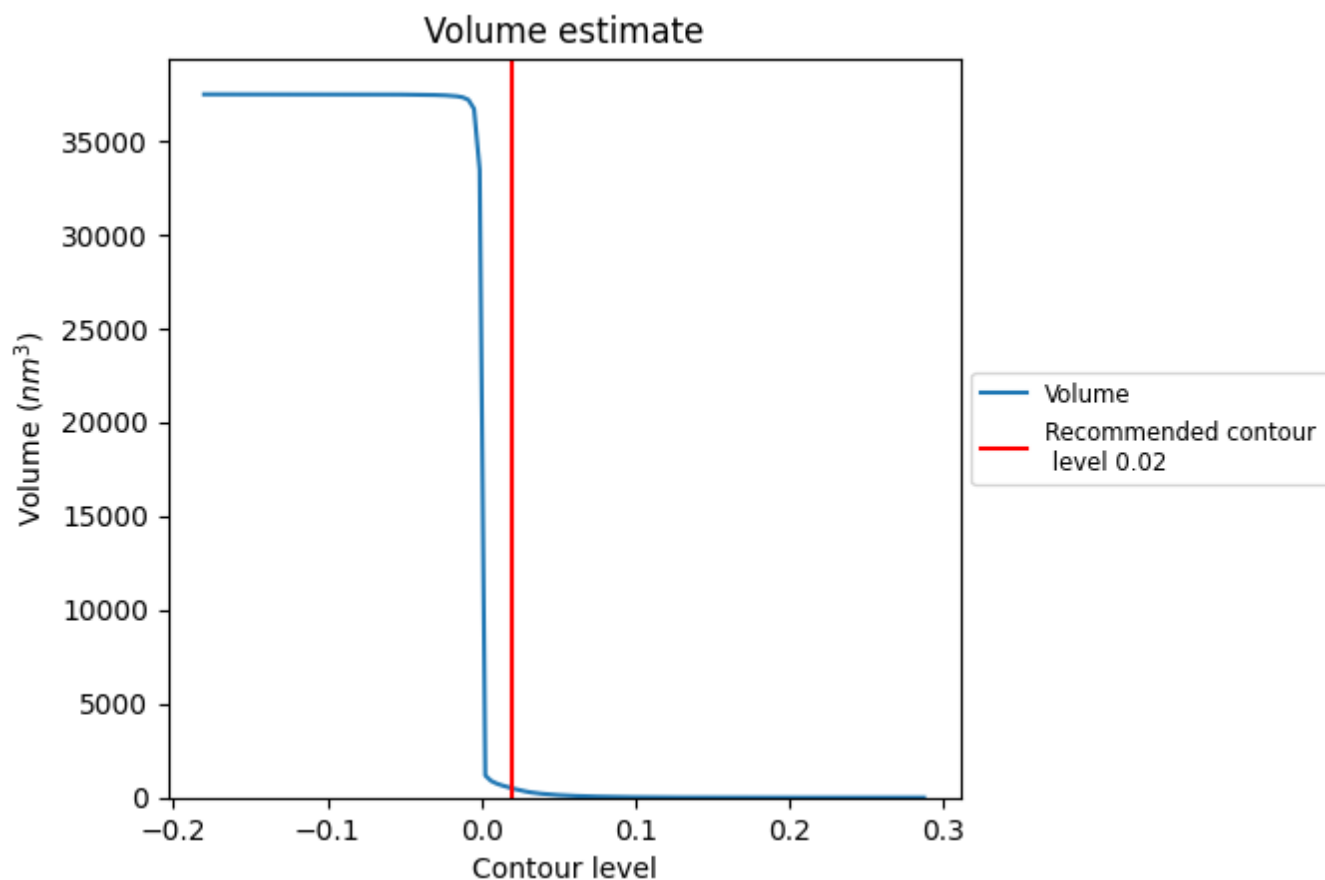
This section contains the results of statistical analysis of the map.

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.

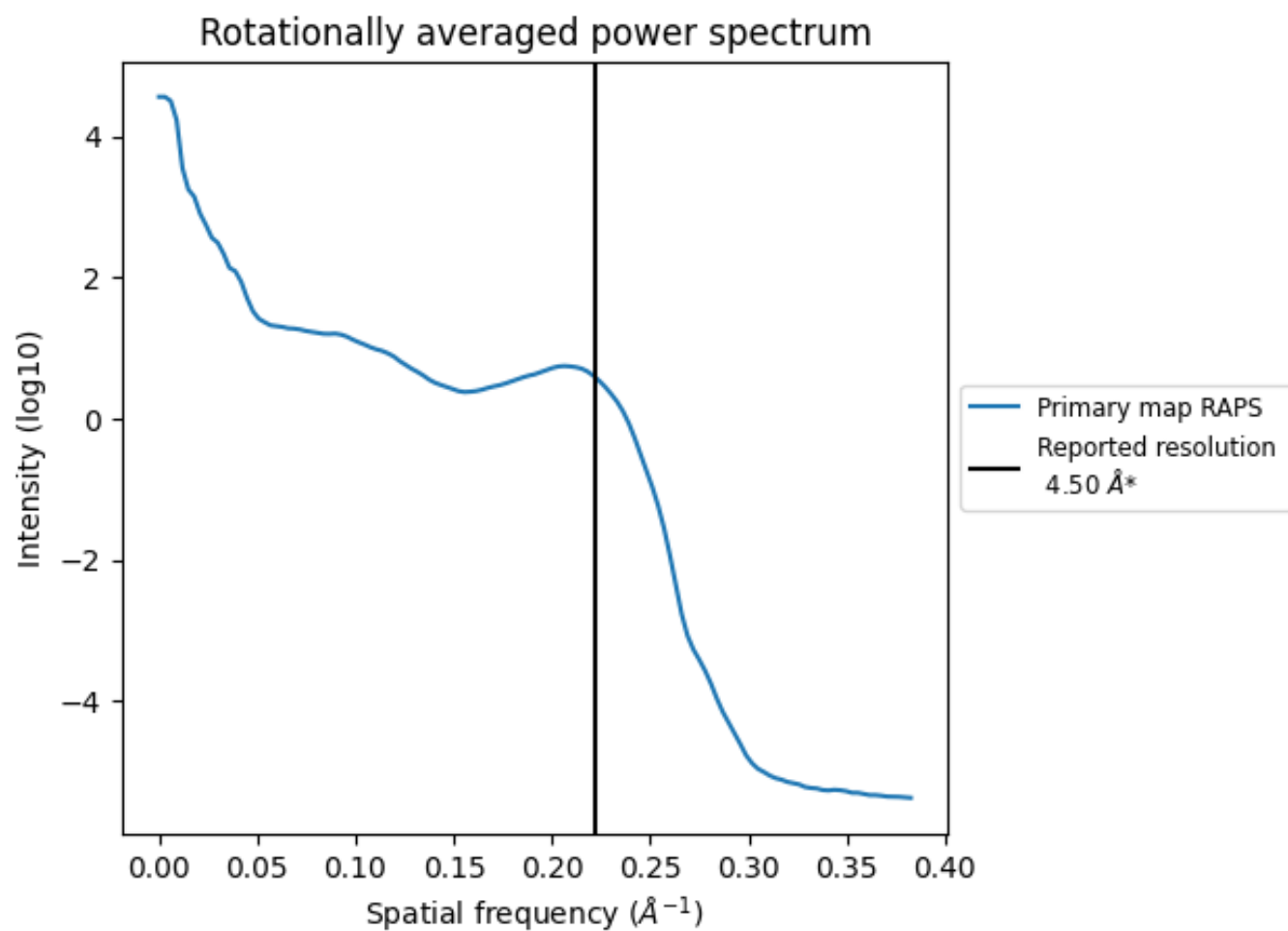
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 491 nm³; this corresponds to an approximate mass of 443 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 0.222 Å⁻¹

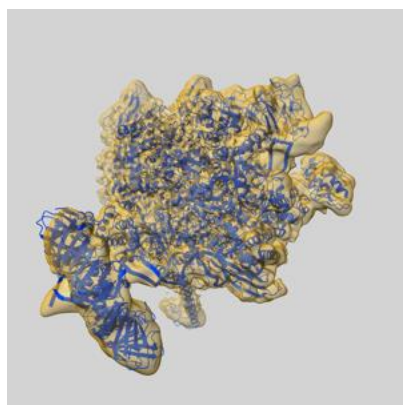
8 Fourier-Shell correlation

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

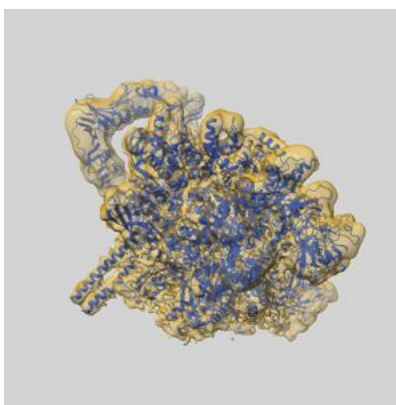
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-30605 and PDB model 7D7D. Per-residue inclusion information can be found in section [3](#) on page [8](#).

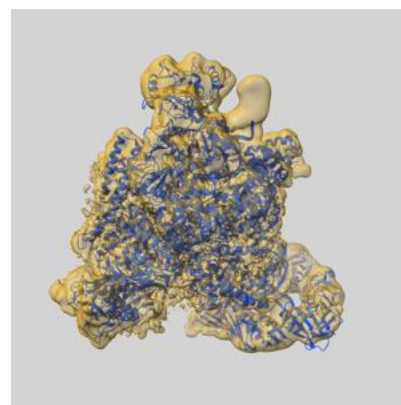
9.1 Map-model overlay [i](#)



X



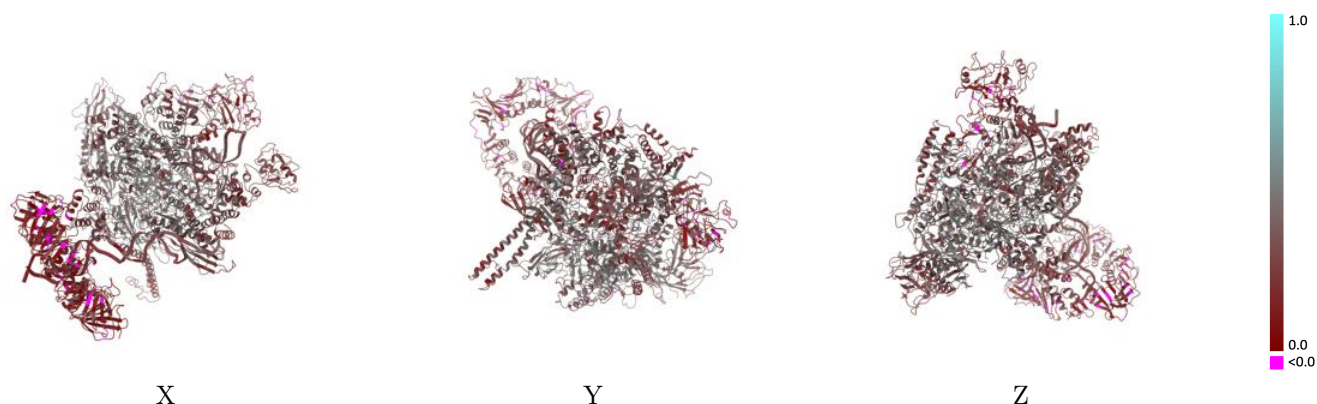
Y



Z

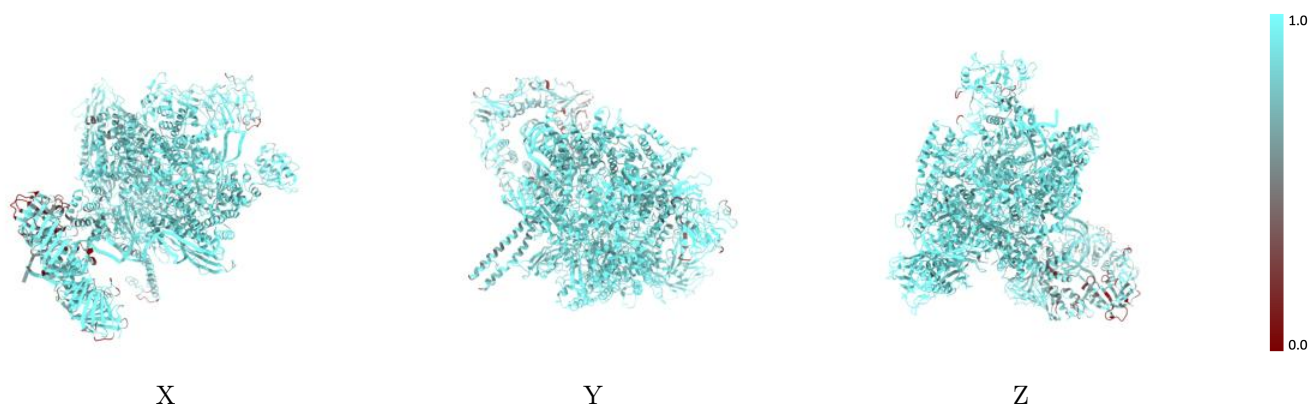
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



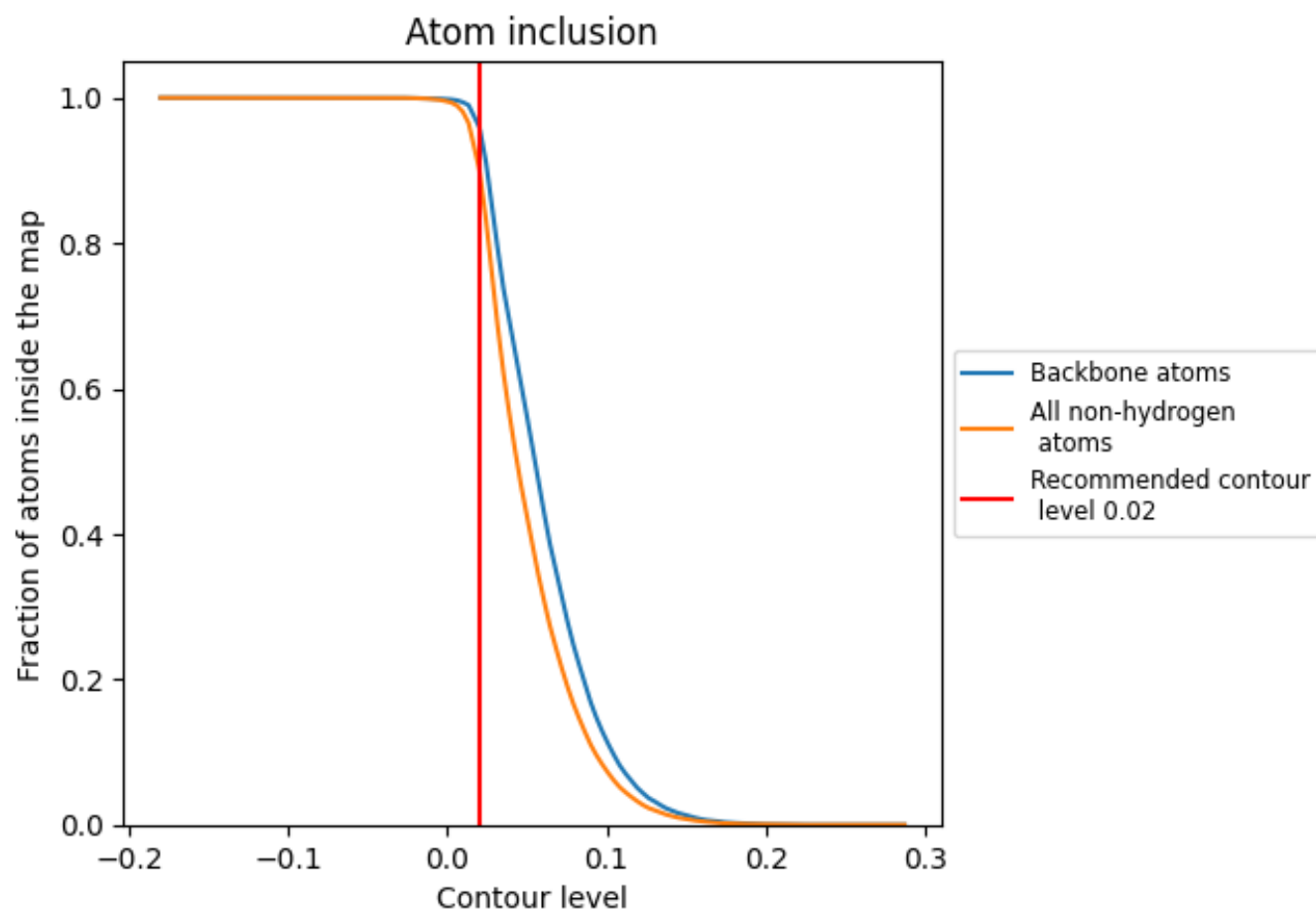
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.02).

9.4 Atom inclusion [i](#)



At the recommended contour level, 96% of all backbone atoms, 90% 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	<div></div> 0.9031	<div></div> 0.3160
A	<div></div> 0.9504	<div></div> 0.3770
B	<div></div> 0.9486	<div></div> 0.3360
C	<div></div> 0.9262	<div></div> 0.3610
D	<div></div> 0.9289	<div></div> 0.3470
E	<div></div> 0.7482	<div></div> 0.3240
F	<div></div> 0.9233	<div></div> 0.3320
G	<div></div> 0.8203	<div></div> 0.1850
H	<div></div> 0.8544	<div></div> 0.1440
I	<div></div> 0.6880	<div></div> 0.1290
K	<div></div> 0.8094	<div></div> 0.2660
N	<div></div> 0.9212	<div></div> 0.2720
T	<div></div> 0.9468	<div></div> 0.2870

1.0

0.0

<0.0