



## wwPDB EM Validation Summary Report ⓘ

Nov 6, 2022 – 03:20 PM EST

PDB ID : 6C9Y  
EMDB ID : EMD-7438  
Title : Cryo-EM structure of E. coli RNAP sigma70 holoenzyme  
Authors : Narayanan, A.; Vago, F.; Li, K.; Qayyum, M.Z.; Yenool, D.; Jiang, W.; Murakami, K.S.  
Deposited on : 2018-01-29  
Resolution : 4.25 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

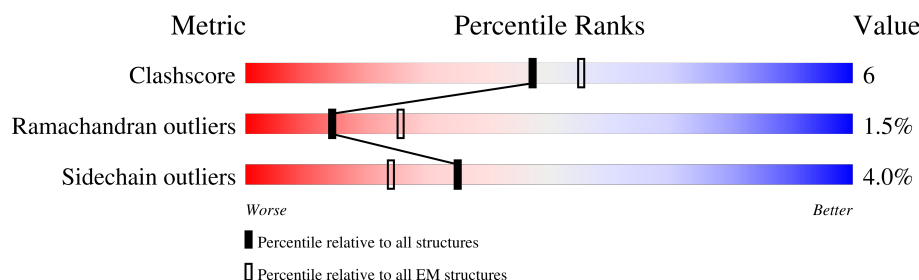
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.25 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  $\geq 3$ , 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	329	<div> <div>37%</div> <div>55%</div> <div>14%</div> <div>30%</div> </div>
1	B	329	<div> <div>34%</div> <div>52%</div> <div>14%</div> <div>33%</div> </div>
2	C	1342	<div> <div>65%</div> <div>82%</div> <div>16%</div> </div>
3	D	1407	<div> <div>63%</div> <div>73%</div> <div>21%</div> </div>
4	E	91	<div> <div>75%</div> <div>73%</div> <div>10%</div> <div>16%</div> </div>
5	F	613	<div> <div>70%</div> <div>62%</div> <div>13%</div> <div>24%</div> </div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 28920 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	230	Total	C	N	O	S	0	0
			1787	1112	317	352	6		
1	B	221	Total	C	N	O	S	0	0
			1708	1067	302	333	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
			10570	6631	1841	2055	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1350	Total	C	N	O	S	0	0
			10434	6553	1856	1976	49		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	76	Total	C	N	O	S	0	0
			605	368	115	121	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	468	Total	C	N	O	S	0	0
			3813	2389	678	723	23		

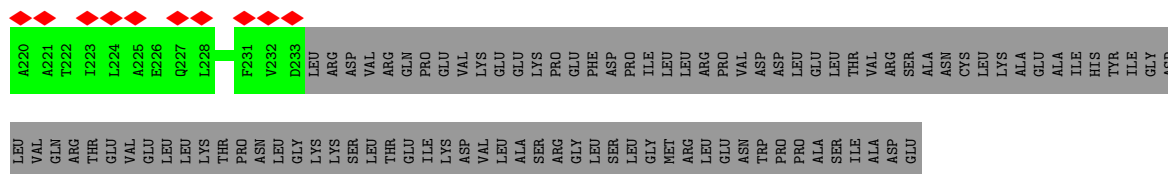
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

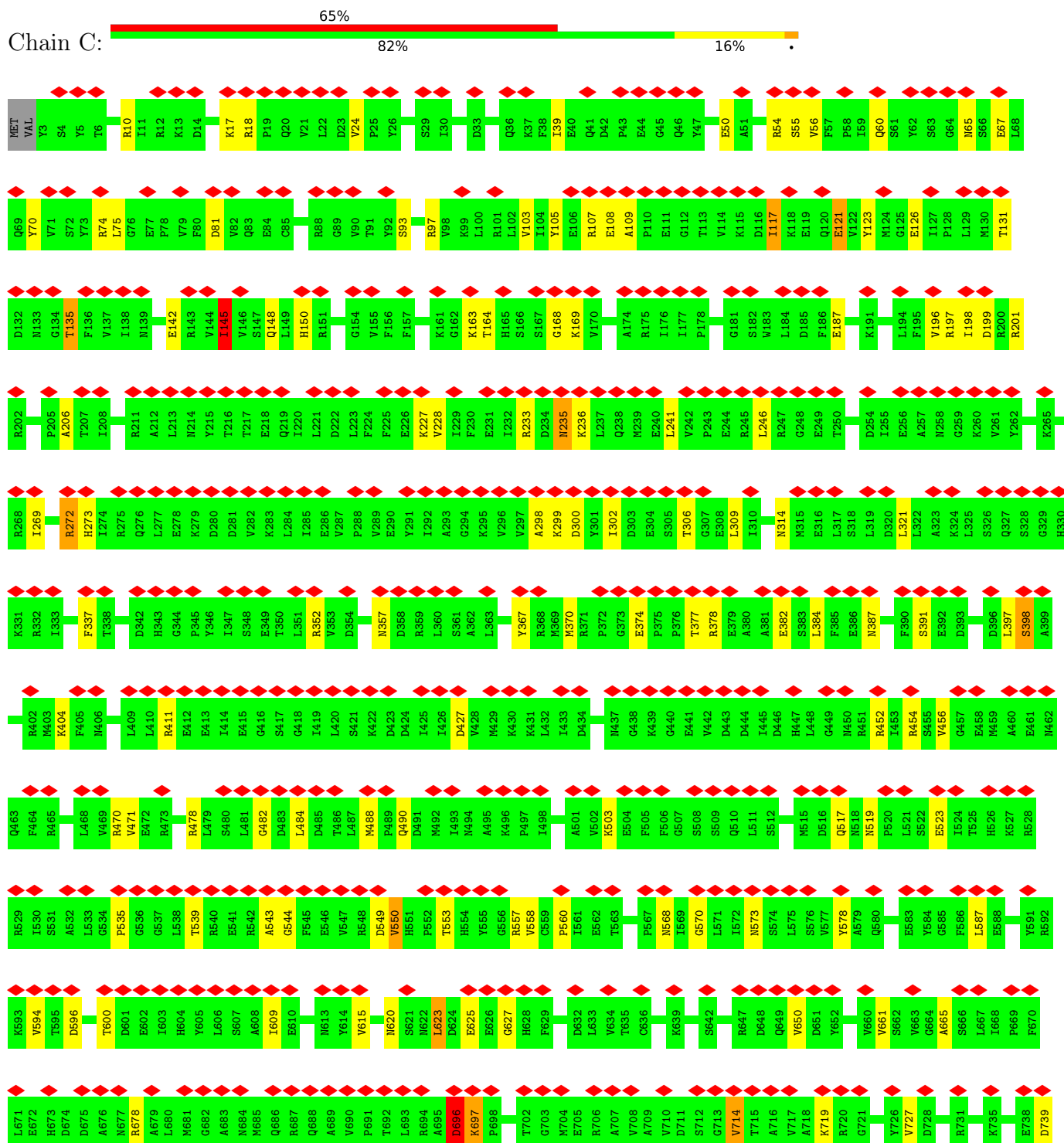
- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

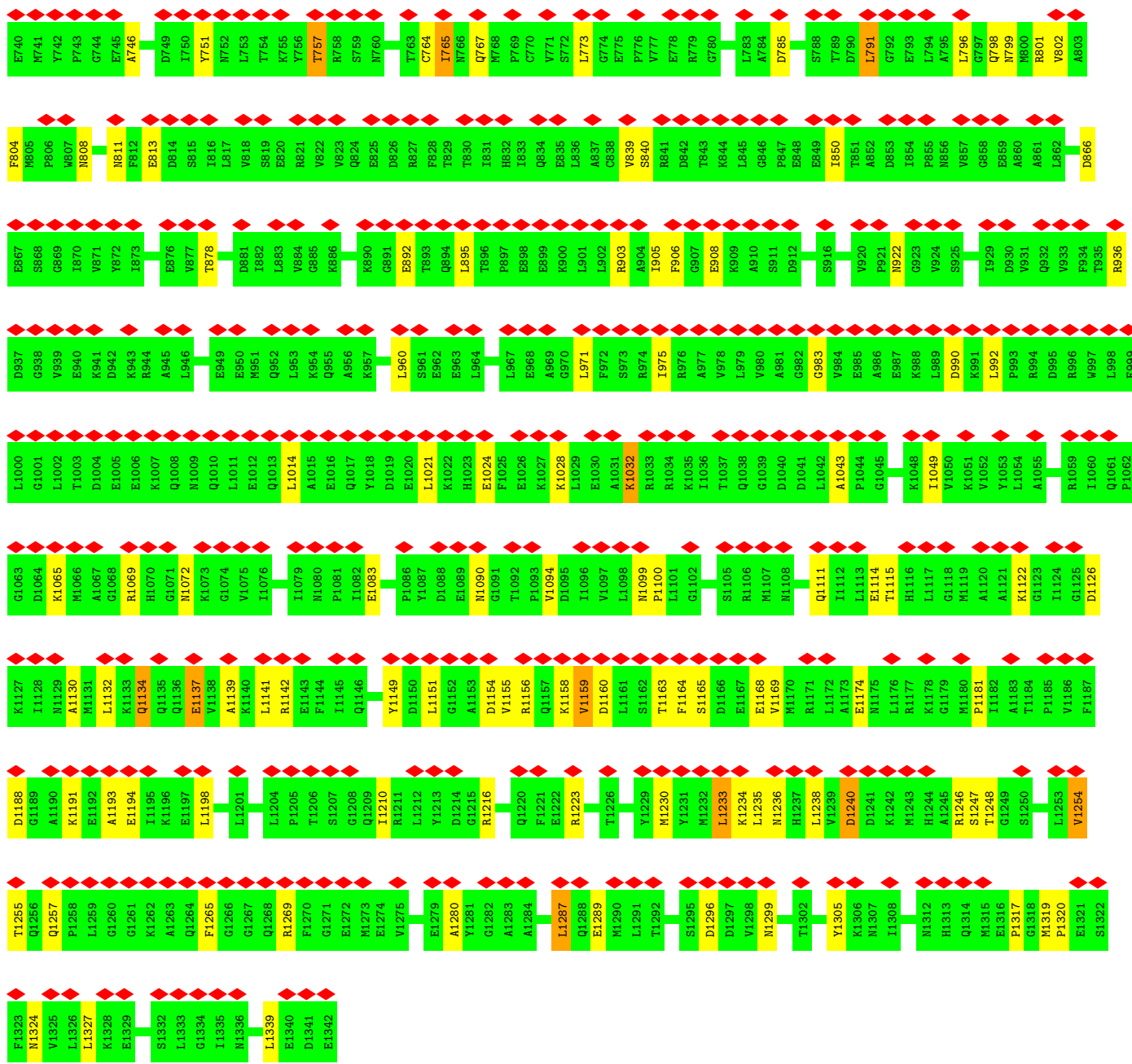
Mol	Chain	Residues	Atoms		AltConf
7	D	2	Total 2	Zn 2	0



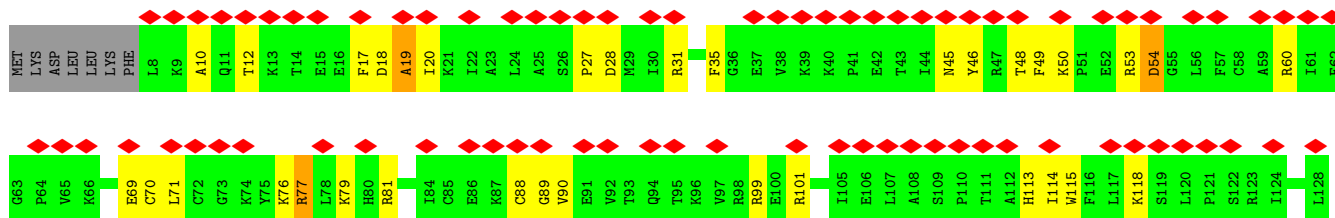
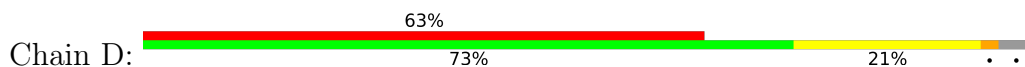


• Molecule 2: DNA-directed RNA polymerase subunit beta





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



E981	L982	K983	L984	I985	F988	G989	P990	T991	K992	E993	S994	Y995	K996	P997	P998	Y999	G1000	A1001	V1002	L1003	A1004	K1005	G1006	E947	S948	S949	I950	Q1010	V1011	A1012	G1013	G1014	E1015	T1016	V1017	A1018	N1019	V1020	P1021	P1022	K964	H1023	T1024	P1025	P1026	V1027	E1030	V1031	S1032	G1033	F1034	V1035	R1036	F1037	T1038	D1039	M1040	I1041	D1042	
Q921	S922	I923	G924	E925	P926	G927	T928	Q929	L930	T931	MET	ARG	THR	PHE	HIS	ILE	GLY	GLY	ALA	SER	ARG	ALA	A945	A946	E947	S948	S949	I950	Q951	V952	K953	G954	G1014	G956	S957	I958	K959	L960	S961	N962	N963	P1022	P1023	P1024	G965	G906	H907	I908	S970	G971	K972	L973	G974	I975	T976	S977	R978	N979	T980	
T786	A787	L788	K789	T790	A791	T792	S793	G794	Y795	L796	T797	R798	L871	L872	E873	S876	V877	D878	A879	R883	S884	G885	V886	S887	C888	D889	M821	M822	T823	P824	G825	I826	C895	E827	H897	C898	G829	D830	V831	G900	K832	E833	P834	L835	R836	D837	R838	V839	L840	G841	R842	V843	E846	D847	V848	L849	K850	P851	G852	T853
S721	I722	Y723	M724	K725	A726	G729	A730	R731	A734	A735	T736	T737	R738	A741	G742	H743	R744	G745	L746	M747	A748	K749	P750	D751	G752	S753	T754	I755	E756	T757	P758	T759	T760	A761	M762	F763	R764	E765	N768	V769	L770	Q771	T772	F773	I774	S775	T776	H777	R780	K781	G782	L783	A784	D785						
E652	I653	I654	S655	E656	A657	E658	A659	E660	V661	A662	E663	L664	Q665	Q667	F668	Q669	S670	G671	L672	V673	T674	A675	G676	E677	R678	G679	L614	K615	P616	T617	V618	I619	F620	A621	D622	Q623	S694	R695	A696	M697	N700	Q702	T703	E704	T705	V706	I707	N708	R709	D710	Q712	E713	E714	K715	M720					
P584	D516	C517	V518	N519	A520	K521	L527	T528	G529	P530	K531	E532	A533	E534	R535	L536	Y537	R538	S539	G540	L541	A542	S543	L544	H545	A546	R547	V548	K549	V550	E554	Y555	E556	K557	D558	A559	N560	G561	E562	K566	T567	S568	L569	K570	D571	T572	E573	R576	A577	I578	L579	M580	M581	T582	V583					
L452	V453	C454	A455	A456	Y457	N458	A459	D460	F461	D462	G463	D464	Q465	A466	A467	V468	H469	V470	P471	L472	T473	L474	E475	A476	Q477	L478	E479	A480	L483	M484	M485	S486	N489	S492	P493	A494	N495	G496	E497	P498	I499	I500	V501	P502	S503	Q504	D505	V506	V507	L508	G509	L510	Y511	M512	T514					
R388	G389	L390	A391	T392	T393	I394	K395	K398	K399	M400	V401	A402	R403	E404	V408	N409	D410	I411	L412	D413	V414	V415	A416	R417	E418	H419	L422	L423	N424	R425	A426	I427	T428	L429	H430	R431	L432	G433	I434	Q435	A436	F437	E438	P439	V440	L441	I442	E443	G444	K445	A446	I447	Q448	H449	A450	P451				
K321	R322	P323	L324	K325	S326	D329	M330	I331	K332	G333	K334	Q335	G336	R337	F338	N341	L342	L343	G344	K345	R346	V347	D348	Y349	S350	G351	R352	S353	V354	I355	T356	V357	L361	R362	L363	C366	G367	L368	P369	K370	K371	A372	M372	S373	L374	E375	G376	R377	K378	P379	Y382	L385	E386	L387						
K190	S191	M192	D193	L194	E195	Q196	E197	C198	E199	Q200	L201	R202	E203	E204	E207	T208	N209	S210	E211	T212	K213	K214	K215	L216	L217	T218	K219	R220	I221	D222	L223	L224	E225	A226	F227	V228	Q229	S230	P234	M237	I238	L239	T240	V241	L242	P243	V244	R245	L246	P246	T247	D248	L249	R250	P251	L252				
D256	Q257	G258	R259	F260	A261	T262	S263	D264	L265	N266	D267	L268	Y269	R270	R271	N274	R275	R278	L279	K280	R281	L282	L283	D284	L285	A286	A287	P288	D289	I290	V292	R293	N294	R297	N298	E301	A302	V303	D304	A305	D308	N309	G310	R311	R312	G313	R314	A315	I316	T317	G318	S319	N320							
D129	M130	P131	L132	R133	D134	I135	E136	R137	V138	L139	F141	E142	S143	Y144	V145	V146	I147	E148	G149	G150	M151	T152	M153	L154	E155	R156	Q157	Q158	I159	L160	T161	E162	E163	Y164	L165	L166	D167	A168	L169	E170	E171	F172	G173	D174	E175	L176	D177	A178	K179	M180	E183	A184	I185	Q186	A187	L188	L189			







## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	96067	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 ( $e^-/\text{\AA}^2$ )	45	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	0.159	Depositor
Minimum map value	-0.076	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.032	Depositor
Map size ( $\text{\AA}$ )	315.84, 315.84, 315.84	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.316, 1.316, 1.316	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	0.40	0/1809	0.82	3/2451 (0.1%)
1	B	0.40	0/1728	0.80	4/2341 (0.2%)
2	C	0.41	1/10739 (0.0%)	0.75	14/14489 (0.1%)
3	D	0.39	0/10591	0.75	11/14307 (0.1%)
4	E	0.34	0/607	0.66	0/817
5	F	0.34	0/3864	0.76	3/5194 (0.1%)
All	All	0.39	1/29338 (0.0%)	0.76	35/39599 (0.1%)

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
2	C	0	4
3	D	0	2
5	F	0	3
All	All	0	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	550	VAL	C-N	-5.44	1.21	1.34

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	70	TYR	CB-CG-CD1	-9.71	115.17	121.00
2	C	70	TYR	CB-CG-CD2	8.85	126.31	121.00
2	C	1021	LEU	CA-CB-CG	7.09	131.60	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	91	ARG	CG-CD-NE	-6.80	97.52	111.80
3	D	1221	LEU	CA-CB-CG	6.78	130.89	115.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	109	ALA	Peptide
2	C	198	ILE	Peptide
2	C	236	LYS	Peptide
2	C	397	LEU	Peptide
3	D	901	ARG	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	1787	0	1810	32	0
1	B	1708	0	1741	28	0
2	C	10570	0	10581	115	0
3	D	10434	0	10599	174	0
4	E	605	0	612	6	0
5	F	3813	0	3880	49	0
6	D	1	0	0	0	0
7	D	2	0	0	0	0
All	All	28920	0	29223	373	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 6.

The worst 5 of 373 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)
3:D:1140:ARG:HH12	3:D:1144:LEU:HG	1.32	0.94
1:A:233:ASP:HA	1:B:218:ARG:HH12	1.40	0.87

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:74:ARG:NH1	2:C:121:GLU:OE2	2.06	0.87
2:C:299:LYS:NZ	2:C:300:ASP:O	2.11	0.81
2:C:719:LYS:HZ3	2:C:751:TYR:HE1	1.28	0.80

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	228/329 (69%)	204 (90%)	24 (10%)	0	100	100
1	B	217/329 (66%)	196 (90%)	18 (8%)	3 (1%)	11	47
2	C	1338/1342 (100%)	1209 (90%)	109 (8%)	20 (2%)	10	46
3	D	1344/1407 (96%)	1212 (90%)	105 (8%)	27 (2%)	7	40
4	E	74/91 (81%)	70 (95%)	3 (4%)	1 (1%)	11	47
5	F	462/613 (75%)	430 (93%)	27 (6%)	5 (1%)	14	52
All	All	3663/4111 (89%)	3321 (91%)	286 (8%)	56 (2%)	14	46

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	21	SER
1	B	193	GLU
2	C	398	SER
2	C	484	LEU
2	C	697	LYS

### 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	198/286 (69%)	195 (98%)	3 (2%)	65	80
1	B	189/286 (66%)	182 (96%)	7 (4%)	34	59
2	C	1155/1157 (100%)	1104 (96%)	51 (4%)	28	54
3	D	1115/1168 (96%)	1067 (96%)	48 (4%)	29	55
4	E	65/75 (87%)	61 (94%)	4 (6%)	18	46
5	F	417/540 (77%)	406 (97%)	11 (3%)	46	67
All	All	3139/3512 (89%)	3015 (96%)	124 (4%)	35	57

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1296	ASP
4	E	36	ASP
3	D	386	GLU
4	E	5	THR
5	F	491	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	45	ASN
3	D	430	HIS
3	D	320	ASN
3	D	435	GLN
2	C	330	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 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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.



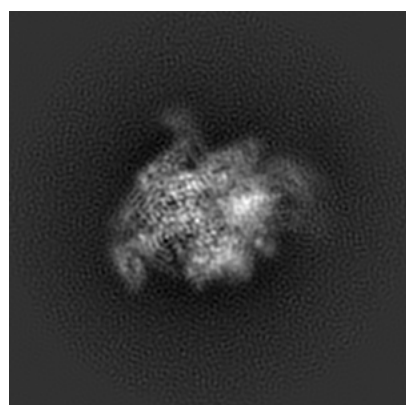
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-7438. These allow visual inspection of the internal detail of the map and identification of artifacts.

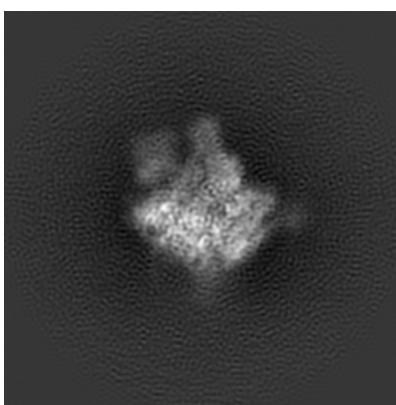
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

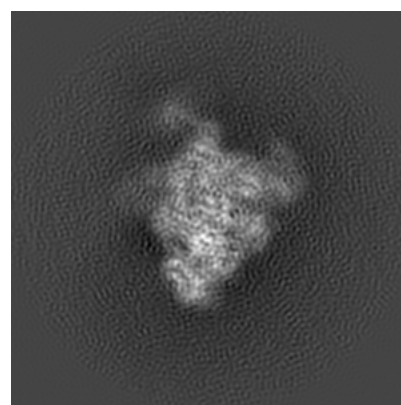
#### 6.1.1 Primary map



X



Y

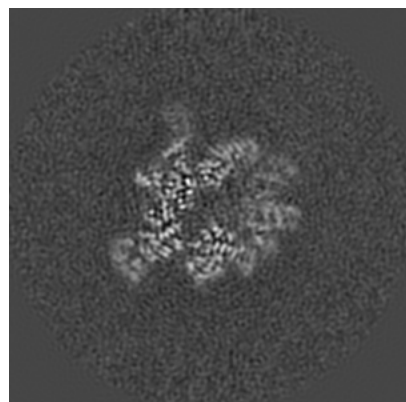


Z

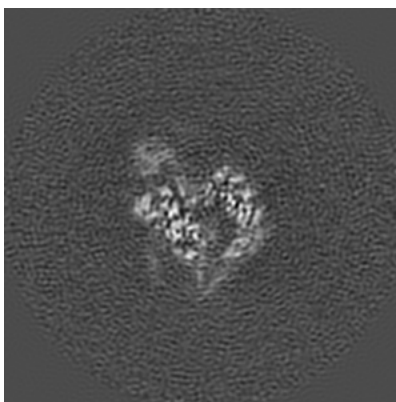
The images above show the map projected in three orthogonal directions.

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

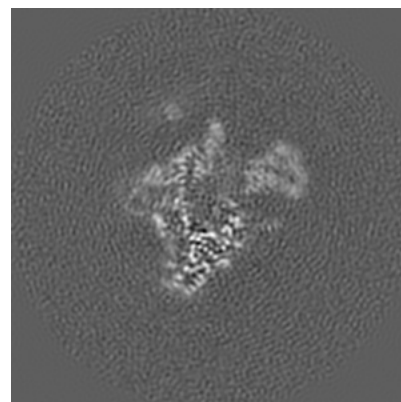
#### 6.2.1 Primary map



X Index: 120



Y Index: 120

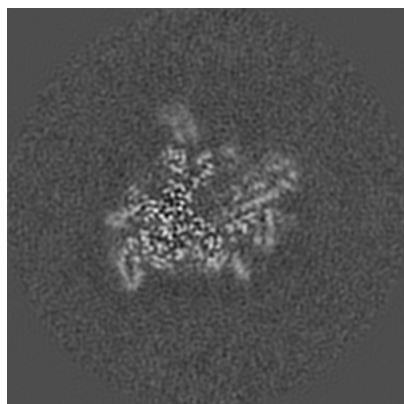


Z Index: 120

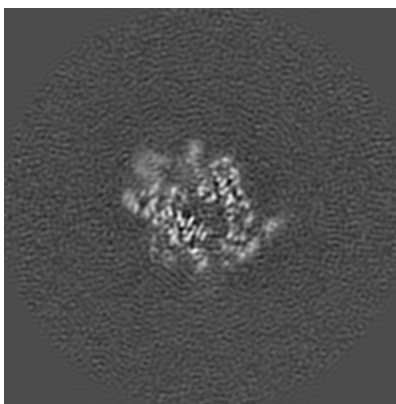
The images above show central slices of the map in three orthogonal directions.

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

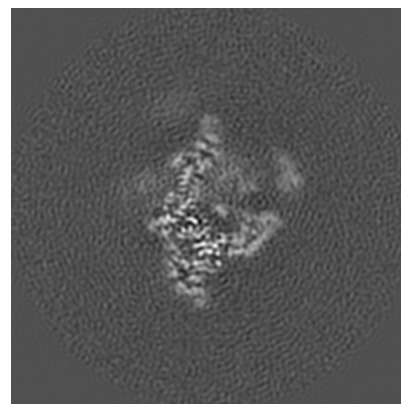
### 6.3.1 Primary map



X Index: 114



Y Index: 114



Z Index: 114

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

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

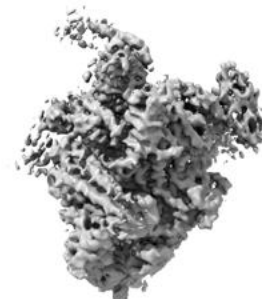
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.032. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

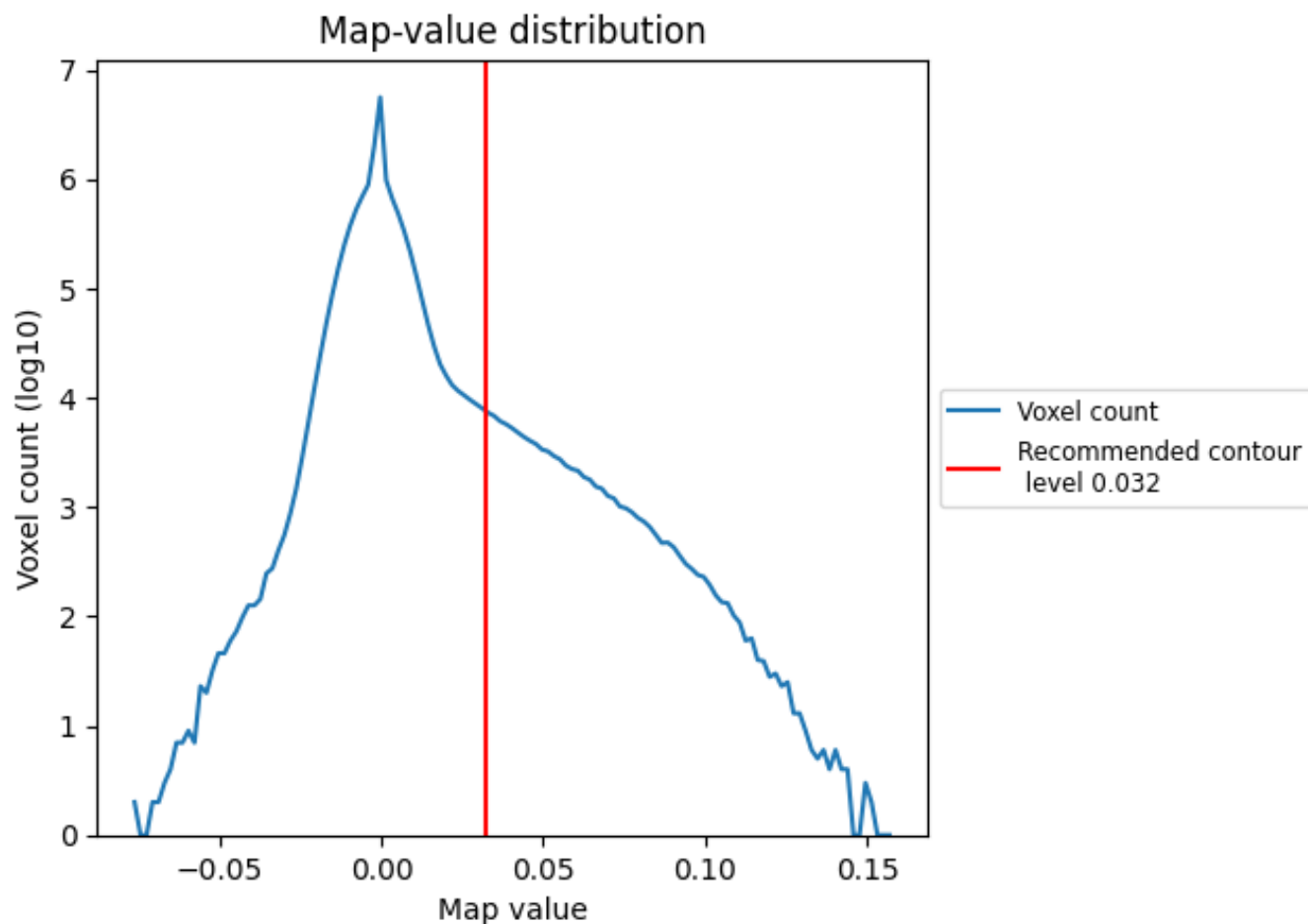
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

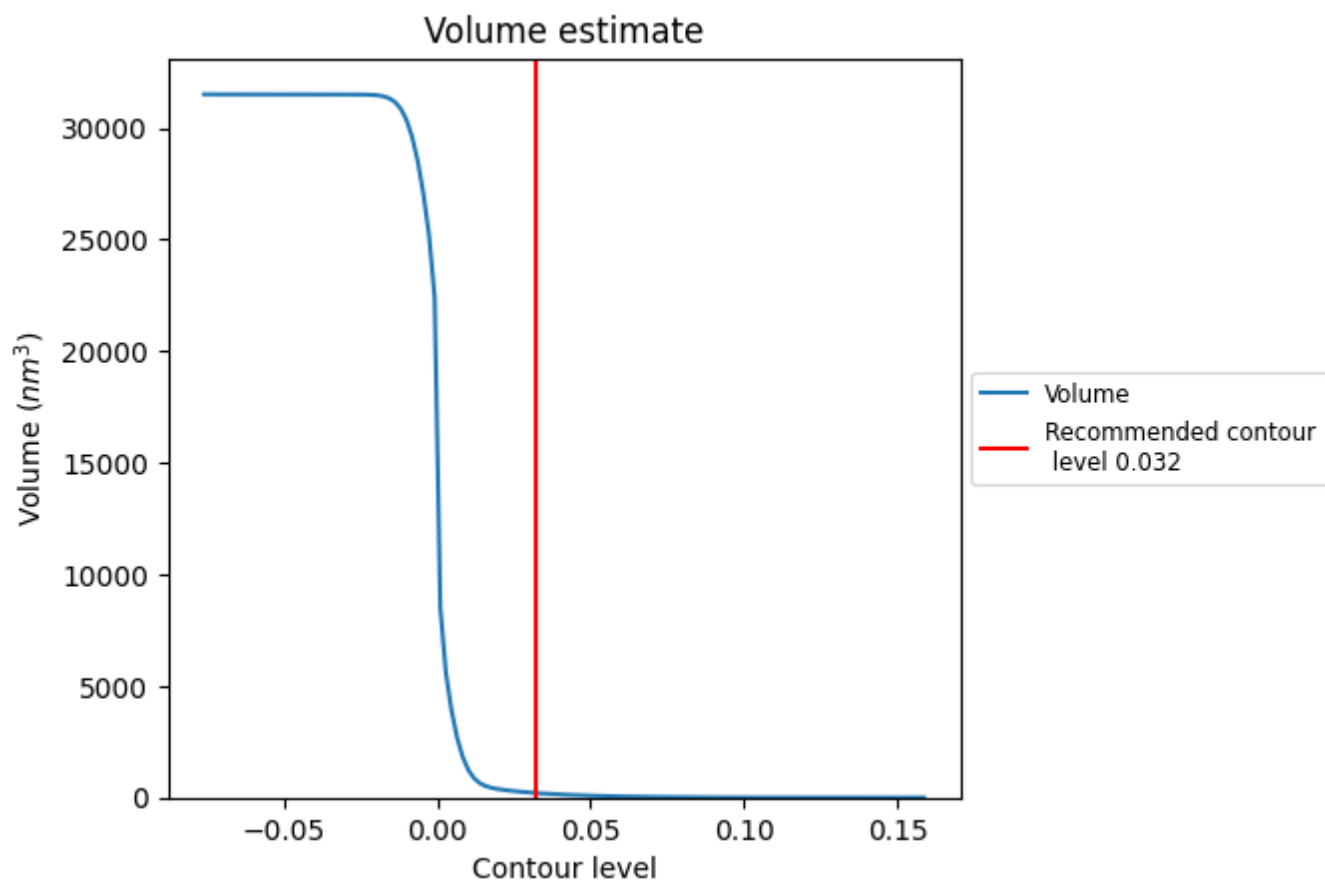
This section contains the results of statistical analysis of the map.

### 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.

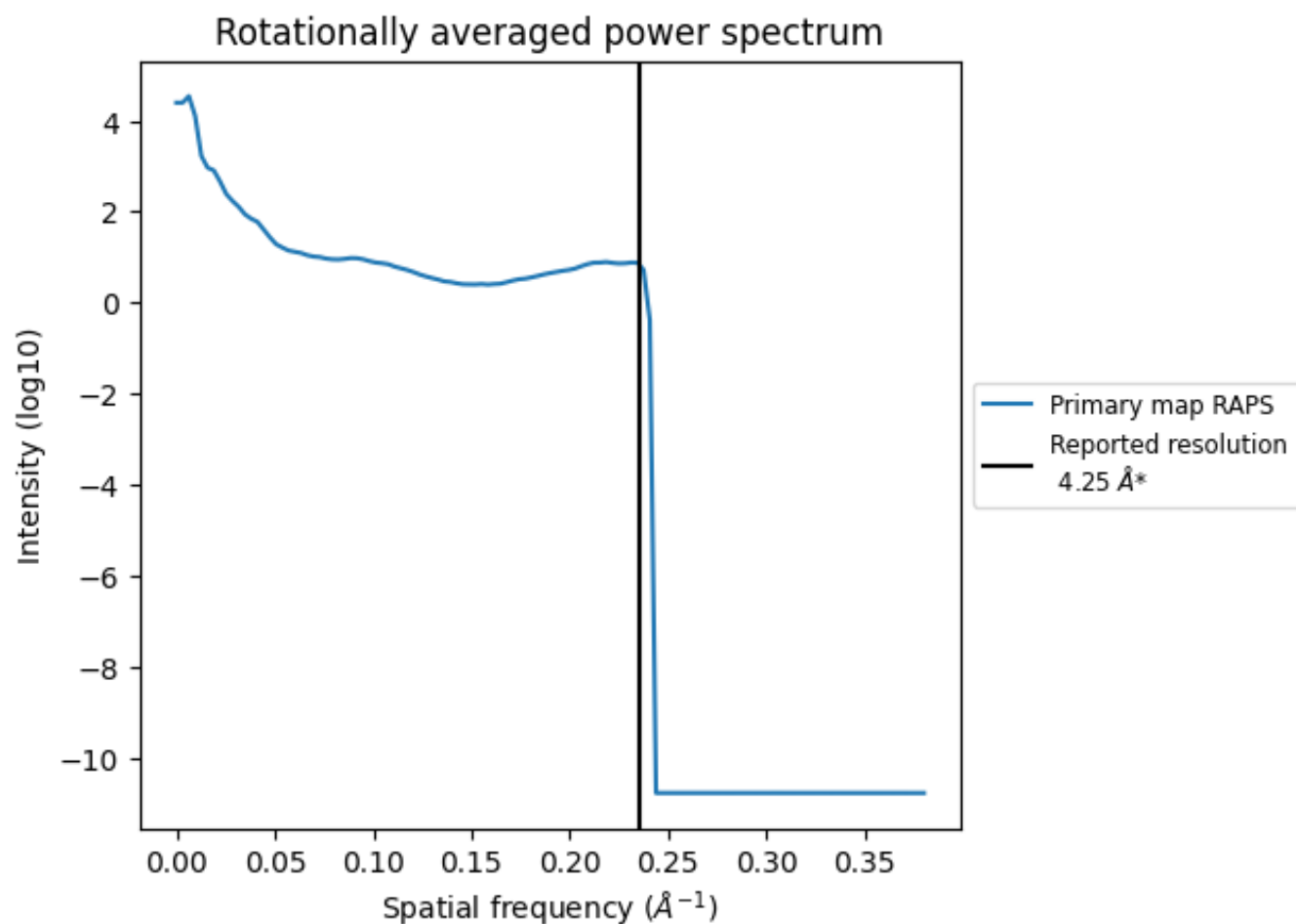
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 206 nm<sup>3</sup>; this corresponds to an approximate mass of 186 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 0.235  $\text{\AA}^{-1}$

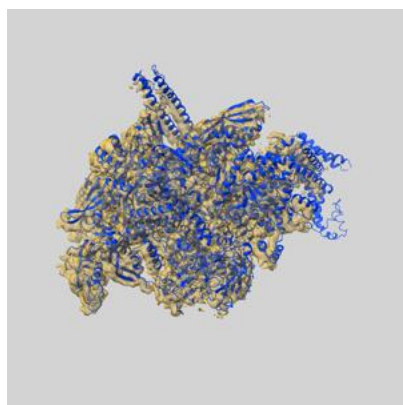
## 8 Fourier-Shell correlation

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

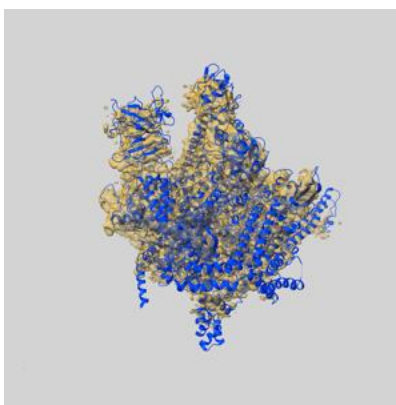
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-7438 and PDB model 6C9Y. Per-residue inclusion information can be found in [section 3](#) on [page 5](#).

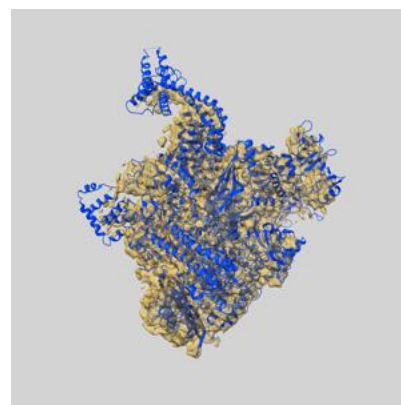
### 9.1 Map-model overlay [i](#)



X



Y

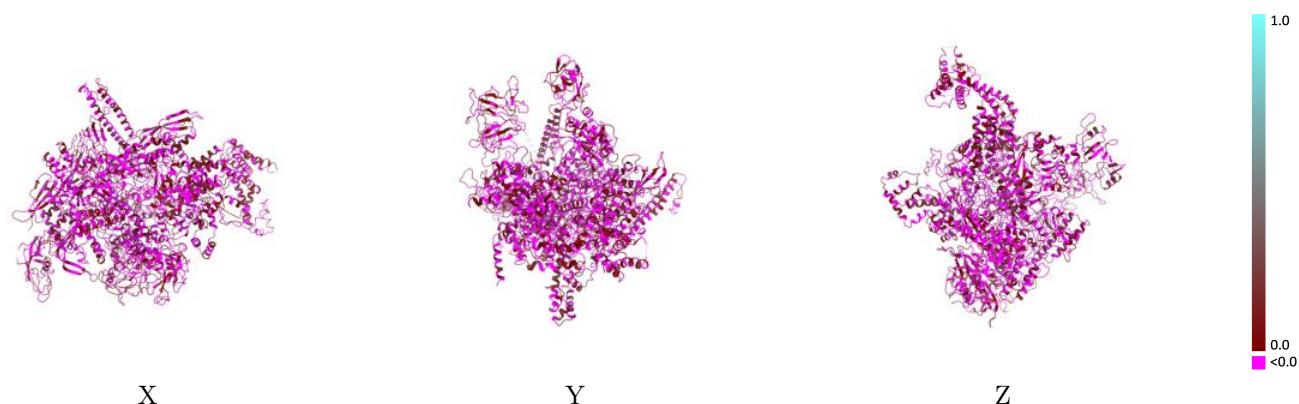


Z

The images above show the 3D surface view of the map at the recommended contour level 0.032 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.

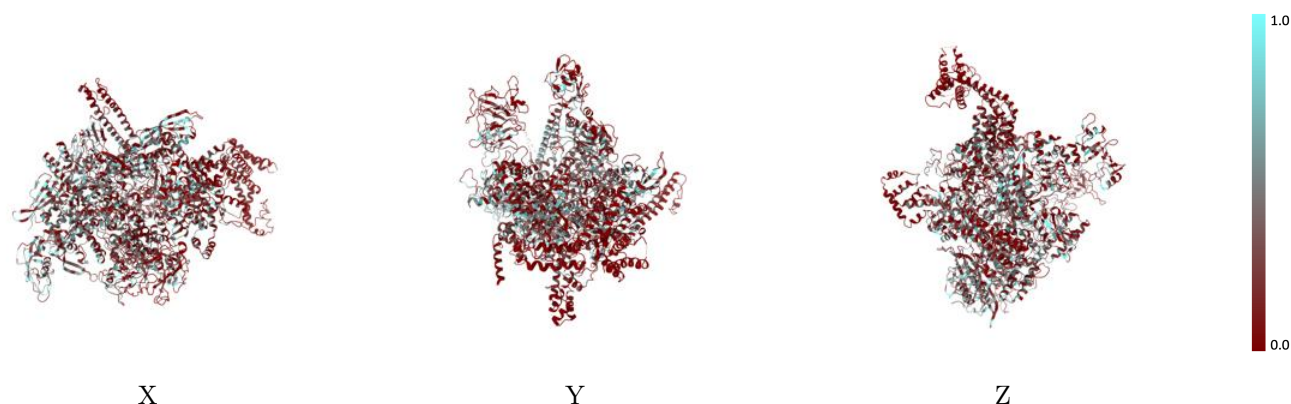


## 9.2 Q-score mapped to coordinate model [i](#)



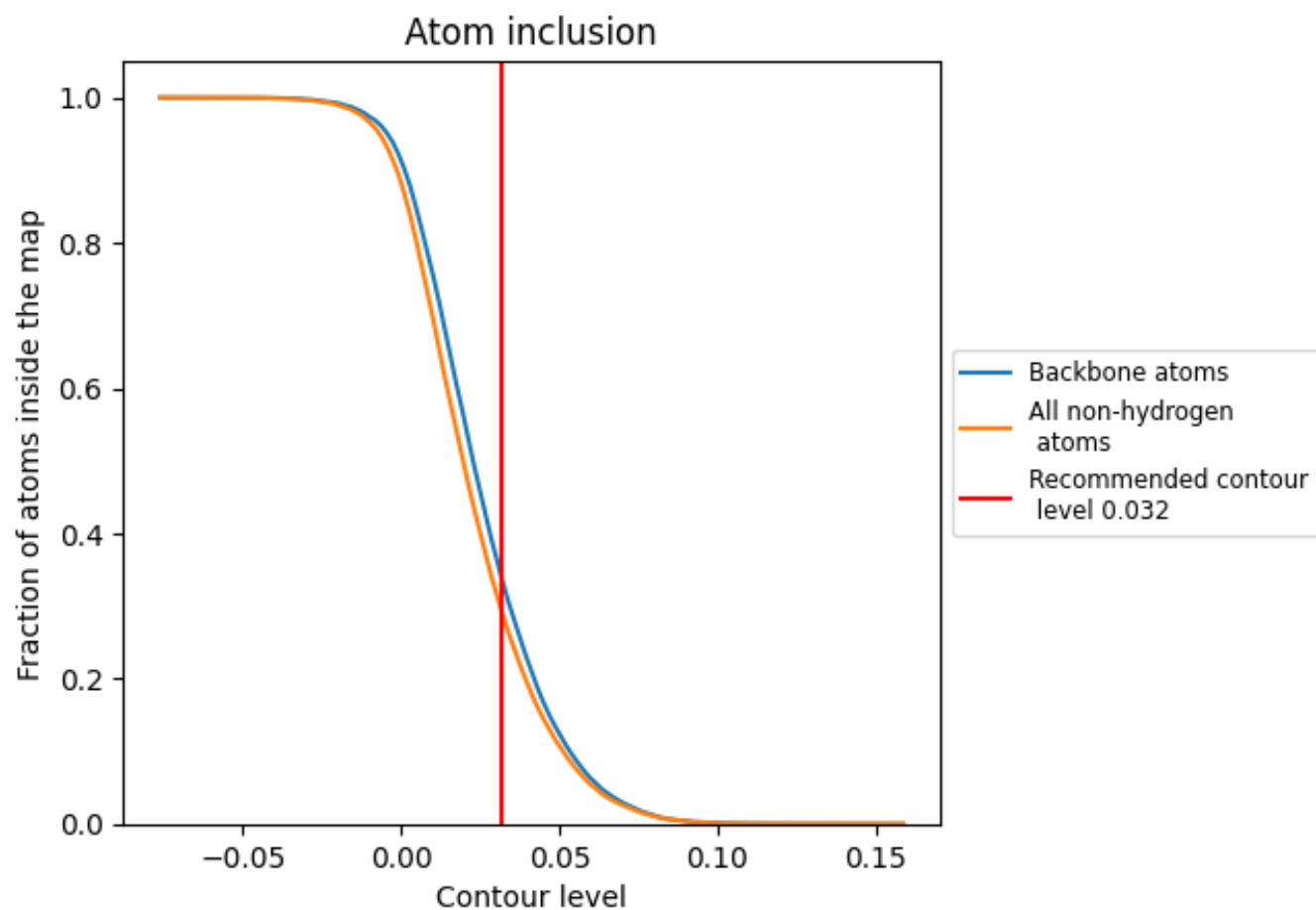
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.032).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 34% of all backbone atoms, 29% 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 (0.032) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.2898	<div></div> -0.0160
A	<div></div> 0.4002	<div></div> -0.0020
B	<div></div> 0.4188	<div></div> -0.0130
C	<div></div> 0.3114	<div></div> -0.0300
D	<div></div> 0.3098	<div></div> -0.0120
E	<div></div> 0.1138	<div></div> -0.0490
F	<div></div> 0.0924	<div></div> 0.0030

1.0  
0.0  
-0.0