

wwPDB EM Validation Summary Report (i)

Nov 19, 2022 – 11:24 pm GMT

PDB ID : 6FLQ EMDB ID : EMD-4275

Title : CryoEM structure of E.coli RNA polymerase paused elongation complex bound

to NusA

Authors : Guo, X.; Weixlbaumer, A.

Deposited on : 2018-01-26

Resolution : 4.10 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) 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)

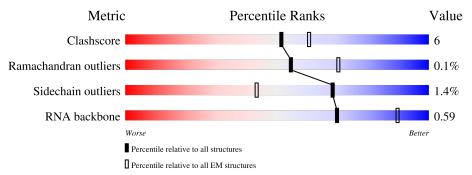
Validation Pipeline (wwPDB-VP) : 2.31.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 4.10 Å.

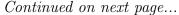
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive $(\# \mathrm{Entries})$	${ m EM\ structures} \ (\#{ m Entries})$
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 >=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	
			88%	
1	A	329	79% 14%	• 6%
			86%	
1	В	329	76% 14%	9%
			85%	
2	С	1342	82%	% •
			81%	
3	D	1407	81% 14%	5%
			92%	•
4	${ m E}$	91	93%	5% •
			100%	
5	F	495	96%	•
			94%	_
6	N	31	77% 23%	





 $Continued\ from\ previous\ page...$

Mol	Chain	Length		Quality of chain						
				95%						
7	R	21	24%	24% 43% 19% 14%						
				79%						
8	Τ	39		69% 31%						



2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 30201 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.

\mathbf{Mol}	Chain	Residues		Atoms				AltConf	Trace	
1	A	310		C 1339		O 429	S 6	0	0	
1	В	298	Total 2068	C 1280		O 408		0	0	

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

Mol	Chain	Residues		A	toms			AltConf	Trace
2	С	1341	Total 10573	C 6634	N 1841	O 2055	S 43	0	0

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

Mol	Chain	Residues		\mathbf{A}	toms			AltConf	Trace
3	D	1334	Total 10357	C 6510	N 1846	O 1952	S 49	0	0

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

Mol	Chain	Residues		At	oms			AltConf	Trace
4	Е	90	Total 709	C 430	N 136	O 142	S 1	0	0

• Molecule 5 is a protein called Transcription termination/antitermination protein NusA.

Mol	Chain	Residues		Atoms				Trace
5	E	495	Total	С	N	О	0	0
	I.	490	2447	1457	495	495	0	0

• Molecule 6 is a DNA chain called DNA (31-MER).



Mol	Chain	Residues		A	toms			AltConf	Trace
6	N	31	Total 647	C 304	N 131	O 181	P 31	0	0

• Molecule 7 is a RNA chain called RNA (5'-R(*CP*CP*UP*GP*AP*UP*CP*AP*GP*GP*CP*GP*AP*UP*GP*UP*GP*UP*GP*UP-3').

Mo	Chain	Residues		At	oms			AltConf	Trace
7	D	91	Total	С	N	О	Р	0	0
1	n	21	444	199	77	148	20	U	U

• Molecule 8 is a DNA chain called DNA (39-MER).

Mol	Chain	Residues		Atoms			AltConf	Trace	
8	Т	39	Total 785	C 375	N 135	O 236	P 39	0	0

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

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

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

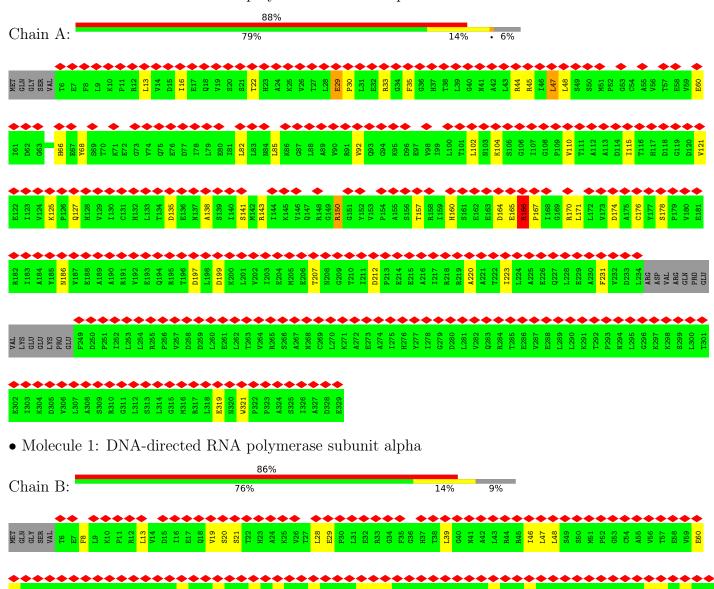
Mol	Chain	Residues	Atoms	AltConf
10	D	2	Total Zn 2 2	0



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: DNA-directed RNA polymerase subunit alpha

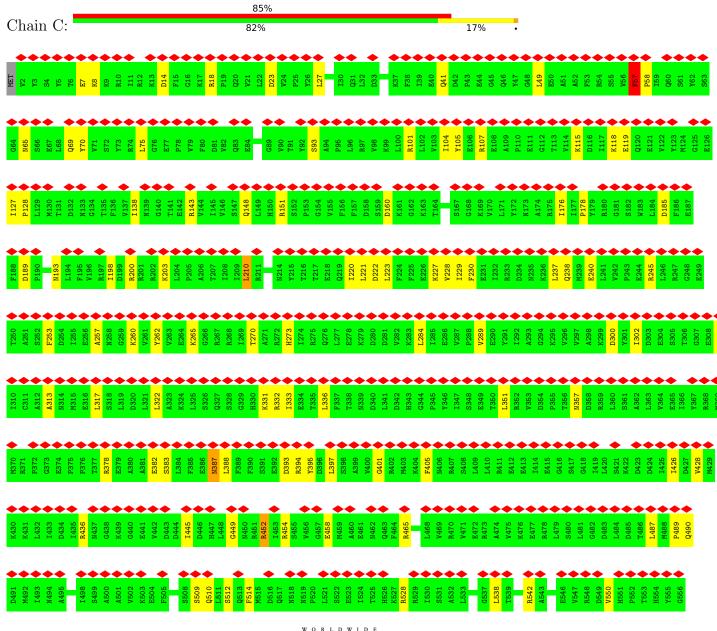




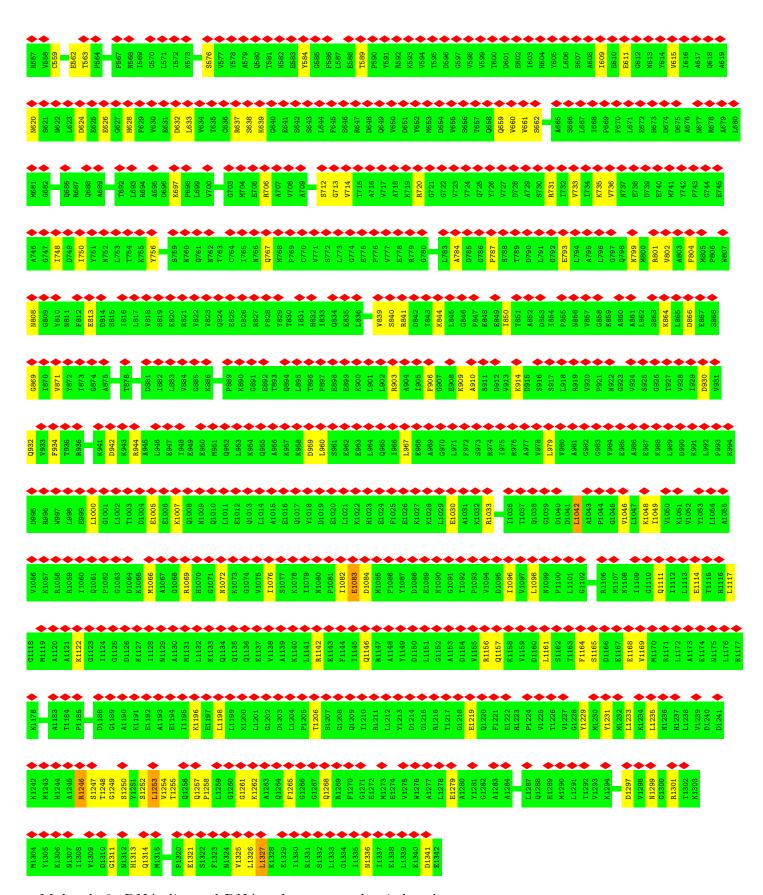
EMD-4275, 6FLQ



• Molecule 2: DNA-directed RNA polymerase subunit beta

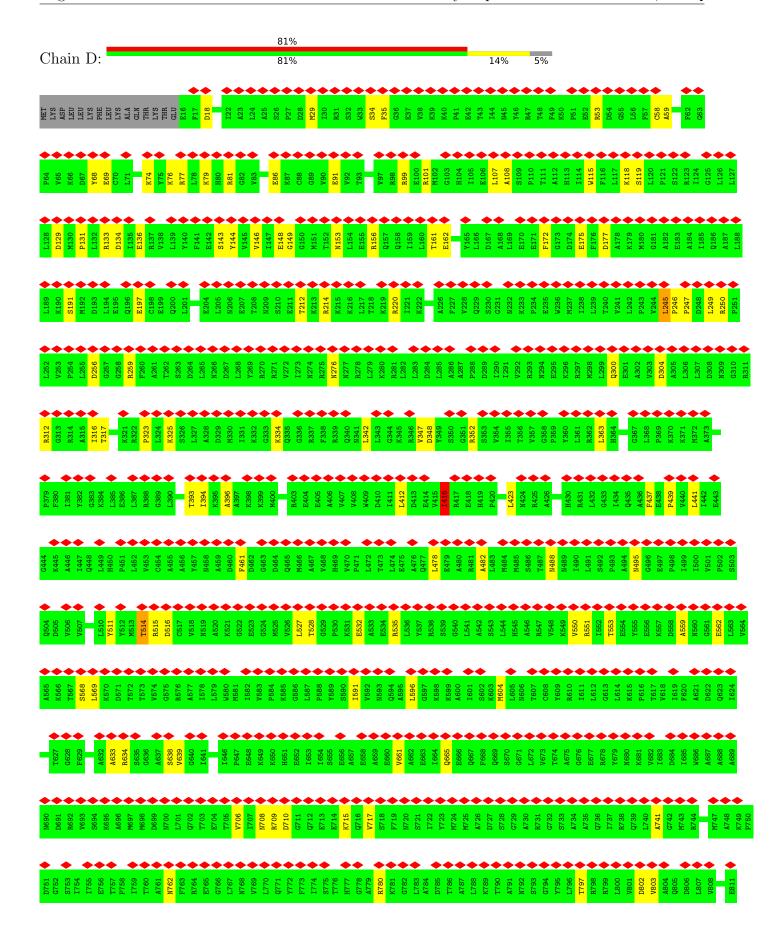




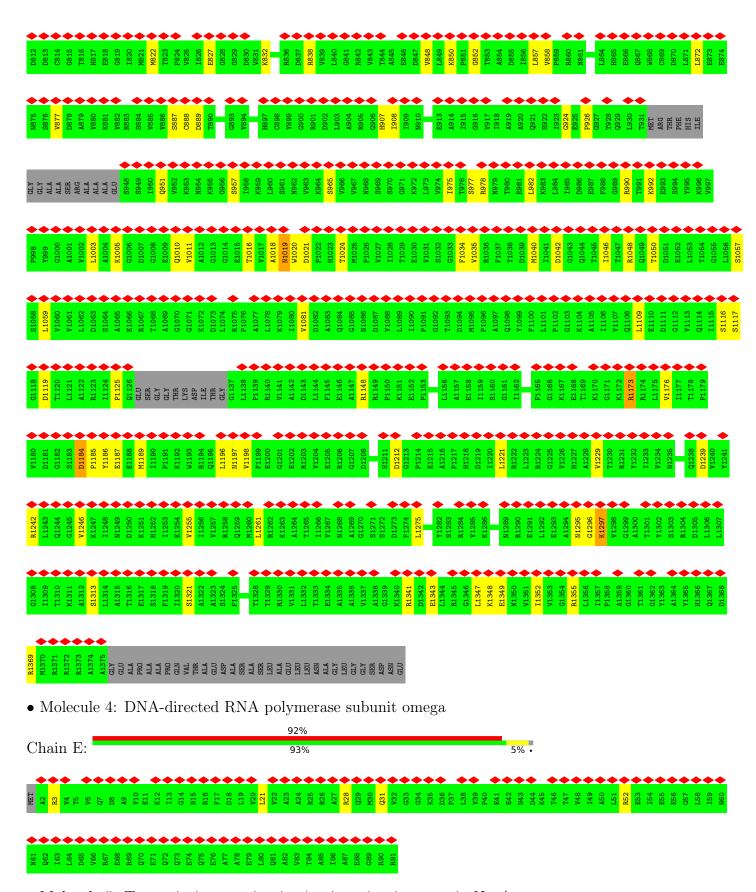


• Molecule 3: DNA-directed RNA polymerase subunit beta'



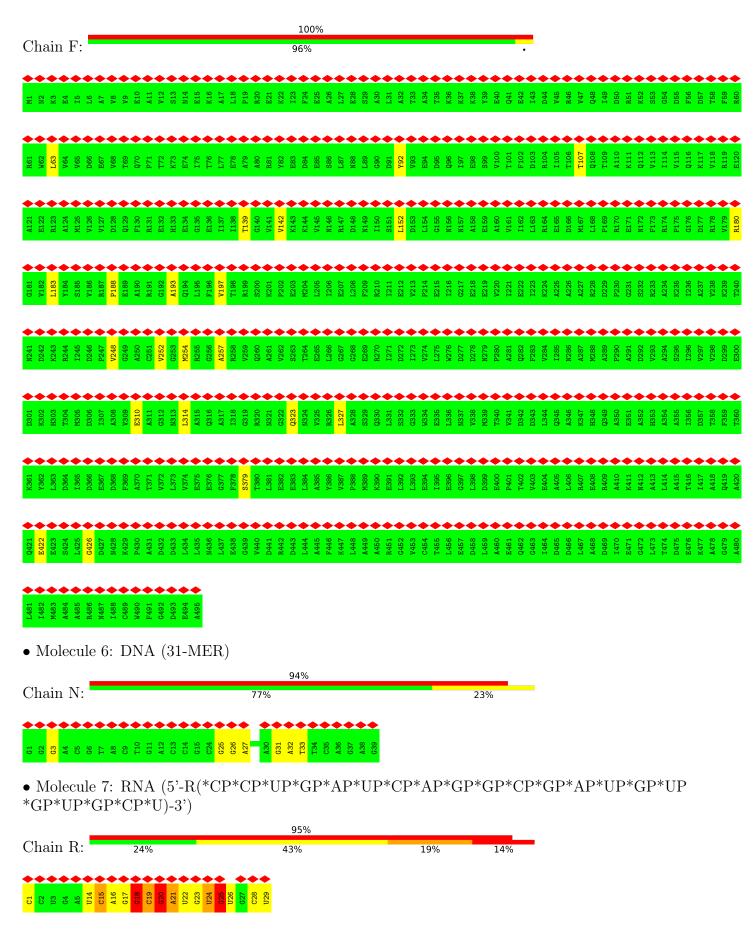






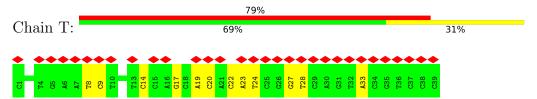
• Molecule 5: Transcription termination/antitermination protein NusA







• Molecule 8: DNA (39-MER)





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	157100	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE	Depositor
	CORRECTION	
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	53	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	105000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	2.869	Depositor
Minimum map value	-1.326	Depositor
Average map value	0.006	Depositor
Map value standard deviation	0.086	Depositor
Recommended contour level	0.45	Depositor
Map size (Å)	308.0, 308.0, 308.0	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.1, 1.1, 1.1	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: 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	Bo	nd lengths	В	ond angles
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z >5
1	A	0.65	0/2189	0.73	3/2981 (0.1%)
1	В	0.58	0/2086	0.77	$3/2841 \; (0.1\%)$
2	С	0.78	1/10742~(0.0%)	0.75	8/14494 (0.1%)
3	D	0.73	3/10514 (0.0%)	0.75	4/14199~(0.0%)
4	Е	0.57	0/711	0.66	0/956
5	F	0.26	0/2446	0.46	0/3406
6	N	1.05	0/728	1.06	1/1121 (0.1%)
7	R	0.98	0/494	1.75	18/766 (2.3%)
8	Τ	1.26	$2/876 \ (0.2\%)$	1.09	1/1346~(0.1%)
All	All	0.74	$6/30786 \ (0.0\%)$	0.78	38/42110 (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 maintain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	В	0	4
2	С	0	4
3	D	0	5
5	F	0	1
All	All	0	18

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\mathring{\mathrm{A}})$	$Ideal(\AA)$
2	С	1084	ASP	C-N	-11.21	1.08	1.34
3	D	115	TRP	CB-CG	-5.71	1.40	1.50
3	D	108	ALA	C-N	-5.60	1.21	1.34
3	D	1255	VAL	CB-CG1	-5.28	1.41	1.52

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	Observed(A)	Ideal(A)	
8	Т	14	DC	C3'-O3'	-5.27	1.37	1.44	

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
7	R	25	G	C4-N9-C1'	9.15	138.39	126.50
7	R	25	G	C8-N9-C1'	-8.60	115.82	127.00
1	В	13	LEU	CA-CB-CG	8.46	134.76	115.30
3	D	982	LEU	CA-CB-CG	8.29	134.37	115.30
7	R	28	С	C5-C6-N1	8.11	125.06	121.00

There are no chirality outliers.

5 of 18 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	166	ARG	Peptide
1	A	29	GLU	Peptide
1	A	319	GLU	Peptide
1	A	321	TRP	Peptide
1	В	19	VAL	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	2168	0	1960	31	0
1	В	2068	0	1867	24	0
2	С	10573	0	10584	145	0
3	D	10357	0	10571	133	0
4	Е	709	0	719	7	0
5	F	2447	0	1180	11	0
6	N	647	0	347	6	0
7	R	444	0	228	9	0
8	Τ	785	0	440	8	0
9	D	1	0	0	0	0
10	D	2	0	0	0	0
All	All	30201	0	27896	334	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 334 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{array}{ll} ext{Interatomic} \ ext{distance} \ (ext{Å}) \end{array}$	$egin{aligned} ext{Clash} \ ext{overlap } (ext{Å}) \end{aligned}$
2:C:393:ASP:OD2	2:C:394:ARG:NH1	1.90	1.04
3:D:1341:ARG:NH1	3:D:1343:GLU:OE2	1.98	0.97
3:D:129:ASP:OD2	3:D:220:ARG:NH1	1.98	0.96
1:A:33:ARG:NH1	1:A:199:ASP:OD2	2.02	0.91
3:D:832:LYS:HZ2	3:D:1242:ARG:HH12	1.12	0.90

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	Perce	ntiles
1	A	306/329~(93%)	262 (86%)	44 (14%)	0	100	100
1	В	292/329 (89%)	243 (83%)	47 (16%)	2 (1%)	22	60
2	С	1339/1342 (100%)	1190 (89%)	147 (11%)	2 (0%)	51	84
3	D	1328/1407 (94%)	1191 (90%)	136 (10%)	1 (0%)	51	84
4	E	88/91 (97%)	82 (93%)	6 (7%)	0	100	100
5	F	493/495 (100%)	458 (93%)	35 (7%)	0	100	100
All	All	3846/3993 (96%)	3426 (89%)	415 (11%)	5 (0%)	54	84

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	С	57	PHE
2	С	58	PRO
1	В	61	ILE

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
1	В	117	HIS
3	D	1185	PRO

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	Perce	ntiles
1	A	$194/286\ (68\%)$	192 (99%)	2 (1%)	76	85
1	В	183/286~(64%)	182 (100%)	1 (0%)	88	93
2	С	$1155/1157\ (100\%)$	1128 (98%)	27 (2%)	50	70
3	D	1114/1168~(95%)	1107 (99%)	7 (1%)	86	92
4	E	74/75~(99%)	74 (100%)	0	100	100
All	All	$2720/2972\ (92\%)$	2683 (99%)	37 (1%)	68	80

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	С	1301	ARG
3	D	1173	ARG
2	С	1327	LEU
3	D	416	ILE
2	С	550	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1019	ASN
3	D	1326	GLN
2	С	1146	GLN
2	С	1237	HIS
2	С	1268	GLN



5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
7	R	19/21 (90%)	6 (31%)	1 (5%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
7	R	18	G
7	R	19	С
7	R	20	G
7	R	21	A
7	R	25	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
7	R	20	G

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	N	1
7	R	1
2	С	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	N	15:DG	O3'	24:DC	Р	28.78
1	R	5:A	O3'	14:U	P	17.75
1	С	1084:ASP	С	1085:MET	N	1.08



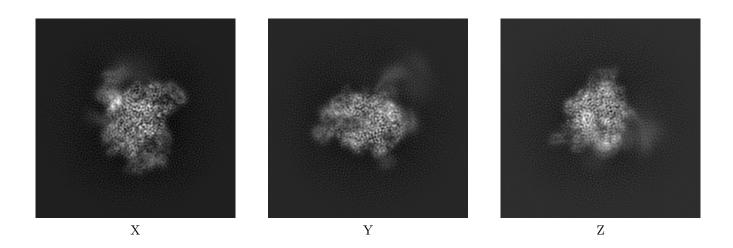
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-4275. 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)

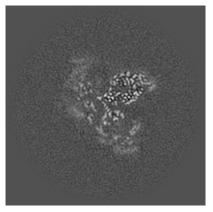
6.1.1 Primary map



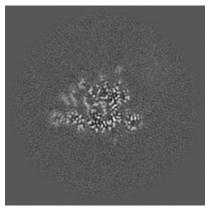
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

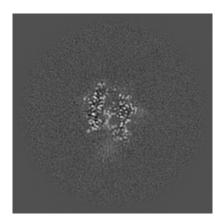
6.2.1 Primary map



X Index: 140



Y Index: 140



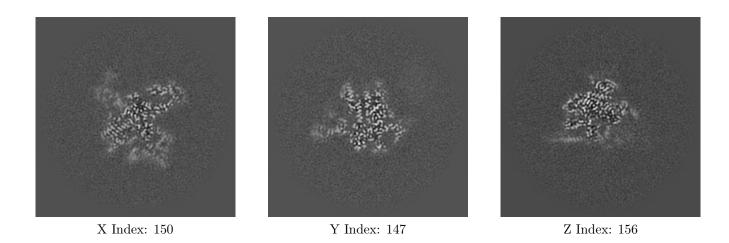
Z Index: 140



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

6.3 Largest variance slices (i)

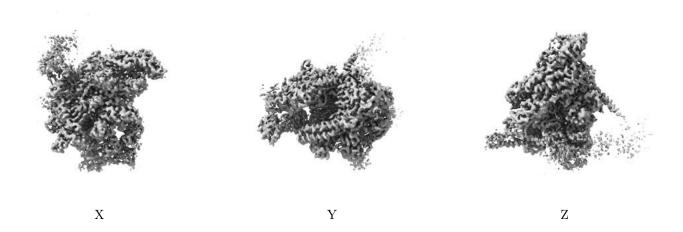
6.3.1 Primary map



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



The images above show the 3D surface view of the map at the recommended contour level 0.45. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.



6.5 Mask visualisation (i)

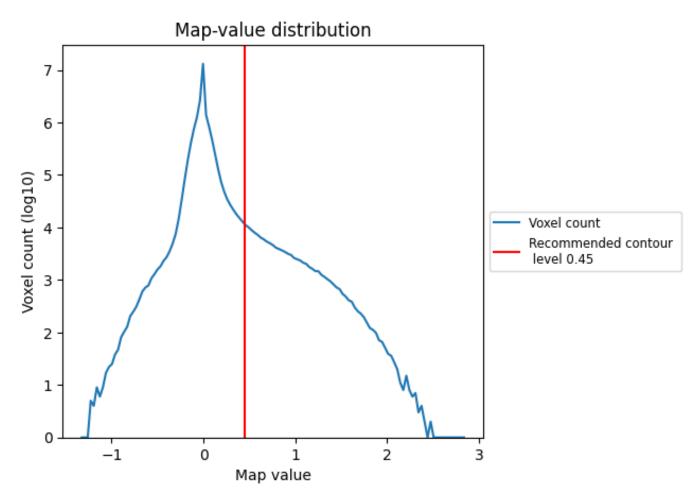
This section was not generated. No masks/segmentation were deposited.



7 Map analysis (i)

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

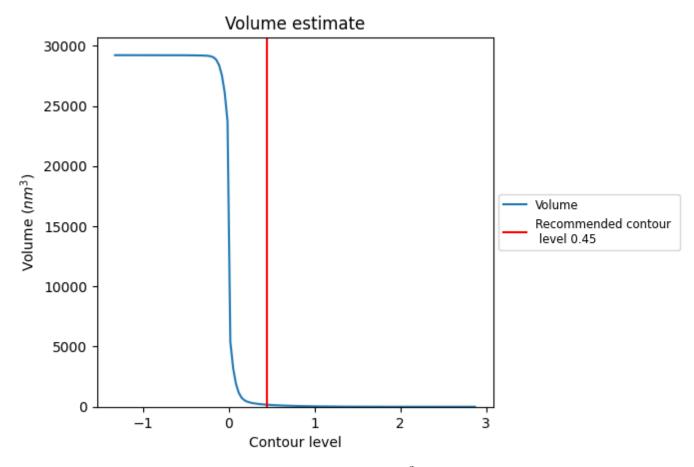
7.1 Map-value distribution (i)



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)

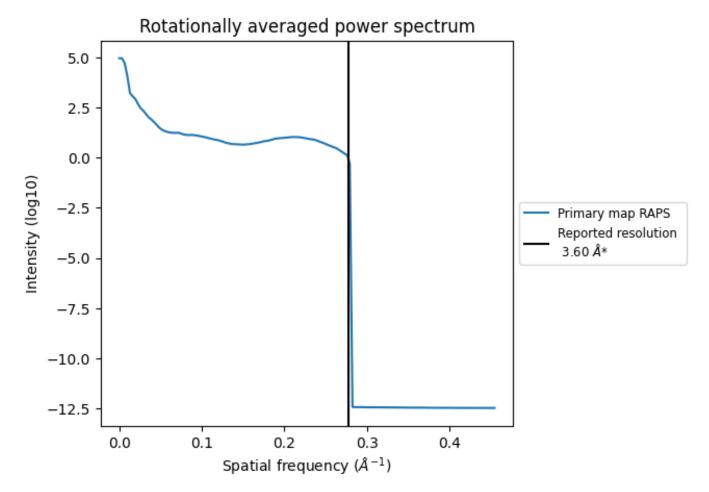


The volume at the recommended contour level is $166~\mathrm{nm^3}$; this corresponds to an approximate mass of $150~\mathrm{kDa}$.

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 $\rm \mathring{A}^{-1}$



8 Fourier-Shell correlation (i)

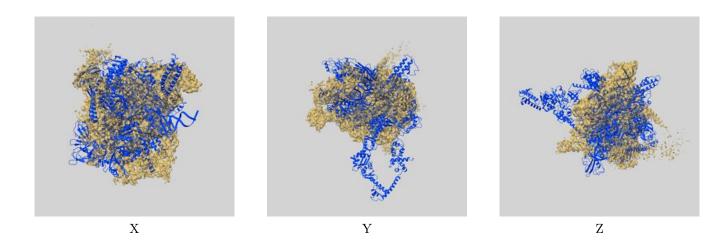
This section was not generated. No FSC curve or half-maps provided.



9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-4275 and PDB model 6FLQ. Per-residue inclusion information can be found in section 3 on page 6.

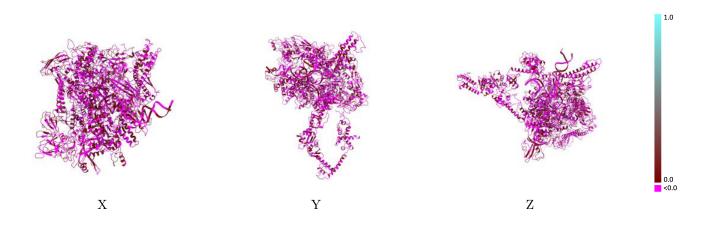
9.1 Map-model overlay (i)



The images above show the 3D surface view of the map at the recommended contour level 0.45 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.

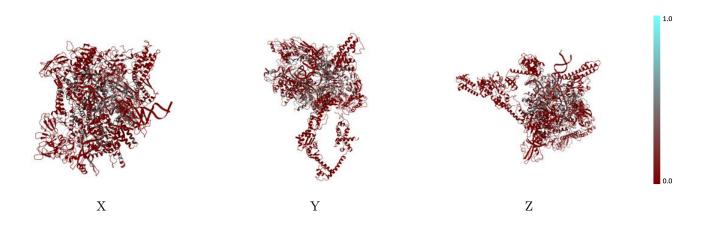


9.2 Q-score mapped to coordinate model (i)



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.

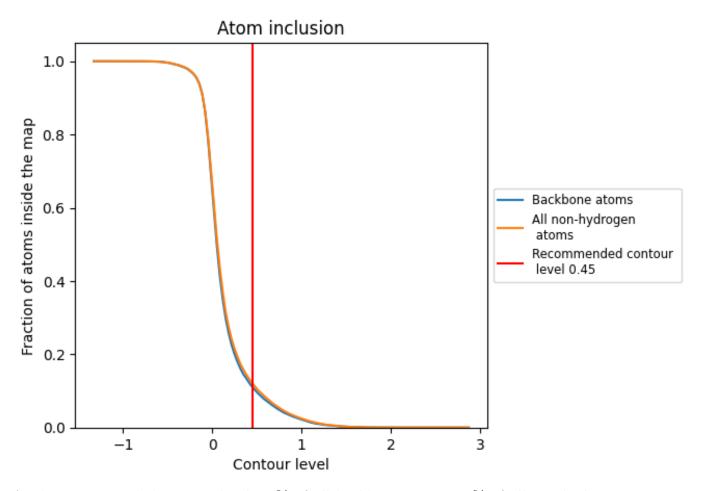
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.45).



9.4 Atom inclusion (i)



At the recommended contour level, 11% of all backbone atoms, 12% of all non-hydrogen atoms, are inside the map.



9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.45) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.1200	-0.0110
A	0.0840	-0.0070
В	0.0486	-0.0030
С	0.1385	-0.0150
D	0.1486	-0.0060
E	0.0682	-0.0180
F	0.0008	-0.0220
N	0.1128	-0.0070
R	0.2275	-0.0290
Т	0.1529	-0.0050



