wwPDB EM Validation Summary Report

Nov 27, 2022 – 06:33 AM EST

PDB ID	:	6OMF
EMDB ID	:	EMD-20090
Title	:	CryoEM structure of SigmaS-transcription initiation complex with activator Crl
Authors	:	Jaramillo Cartagena, A.; Darst, S.A.; Campbell, E.A.
Deposited on	:	2019-04-18
Resolution	:	3.26 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:

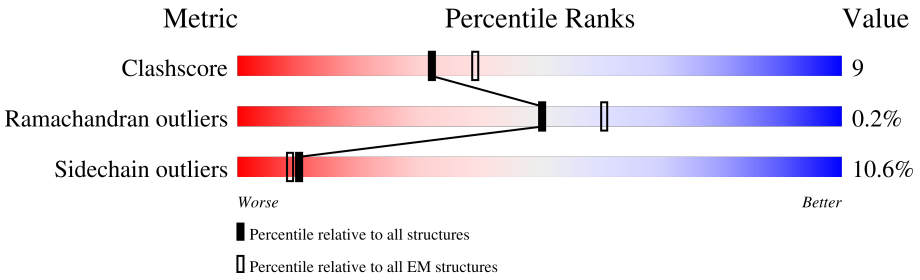
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance

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

The reported resolution of this entry is 3.26 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.



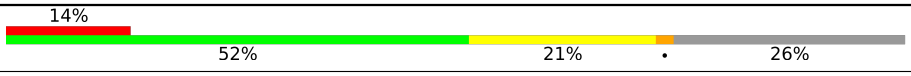
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 <=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	239	7% 65% 27% 5%
1	B	239	64% 27% 7%
2	C	1342	6% 71% 25%
3	D	1407	10% 68% 24% 5%
4	E	91	73% 10% 13%
5	F	331	5% 65% 16% 16%
6	J	136	35% 71% 22% 7%
7	T	66	11% 53% 11% 35%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	N	66	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 30251 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	233	Total	C	N	O	S	0	0
			1806	1127	317	356	6		
1	B	223	Total	C	N	O	S	0	0
			1714	1070	302	336	6		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	GLU	-	expression tag	UNP P0A7Z6
A	236	VAL	-	expression tag	UNP P0A7Z6
A	237	LEU	-	expression tag	UNP P0A7Z6
A	238	PHE	-	expression tag	UNP P0A7Z6
A	239	GLN	-	expression tag	UNP P0A7Z6
B	235	GLU	-	expression tag	UNP P0A7Z6
B	236	VAL	-	expression tag	UNP P0A7Z6
B	237	LEU	-	expression tag	UNP P0A7Z6
B	238	PHE	-	expression tag	UNP P0A7Z6
B	239	GLN	-	expression tag	UNP P0A7Z6

- 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	1336	Total	C	N	O	S	0	0
			10382	6522	1851	1960	49		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	79	Total	C	N	O	S	0	0
			627	382	118	126	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	278	Total	C	N	O	S	0	0
			2244	1406	415	419	4		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	SER	-	expression tag	UNP S5ZIY8

- Molecule 6 is a protein called Sigma factor-binding protein Crl.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	127	Total	C	N	O	S	0	0
			1024	667	166	187	4		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	-2	GLY	-	expression tag	UNP A0A0F7J6B7
J	-1	PRO	-	expression tag	UNP A0A0F7J6B7
J	0	HIS	-	expression tag	UNP A0A0F7J6B7

- Molecule 7 is a DNA chain called Template DNA strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	T	43	Total	C	N	O	P	0	0
			876	418	155	260	43		

- Molecule 8 is a DNA chain called Non-template DNA strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	49	Total	C	N	O	P	0	0
			1008	481	191	288	48		

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

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

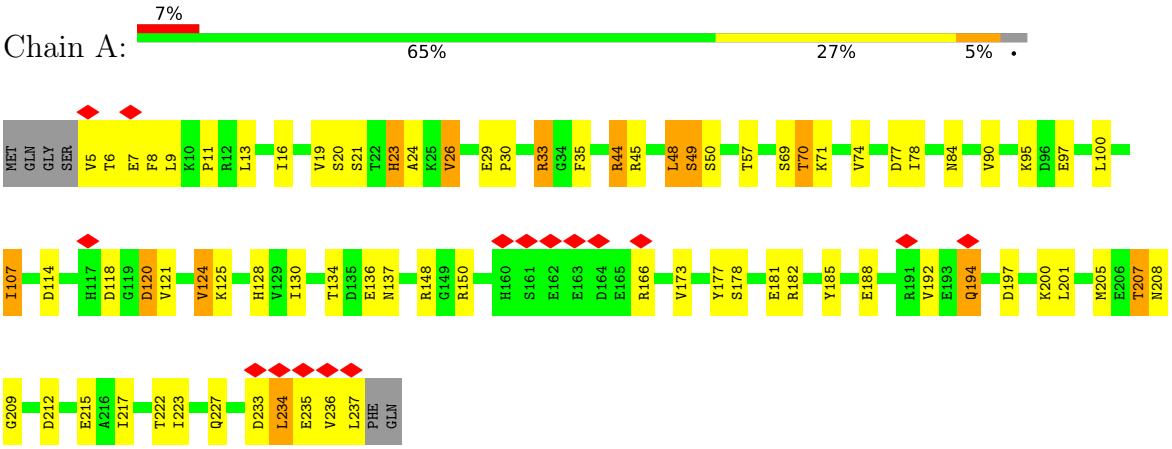
- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

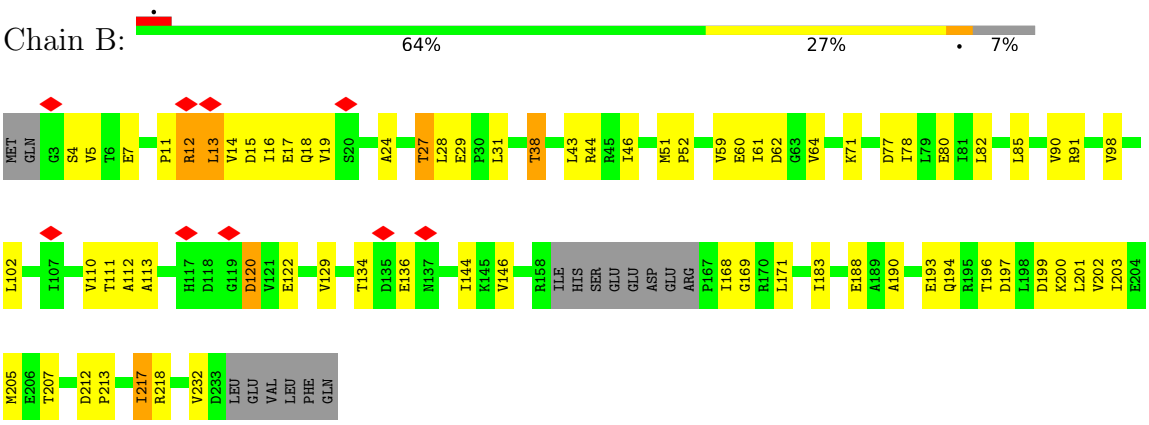
3 Residue-property plots (i)

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.

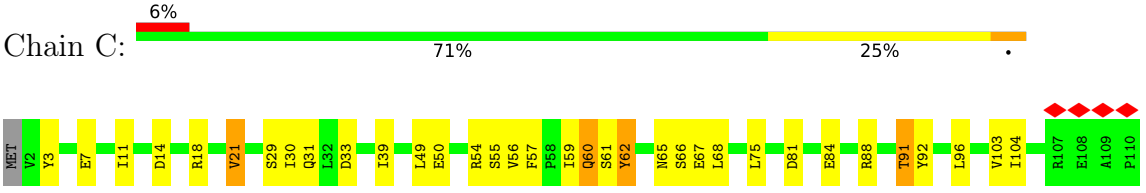
• Molecule 1: DNA-directed RNA polymerase subunit alpha

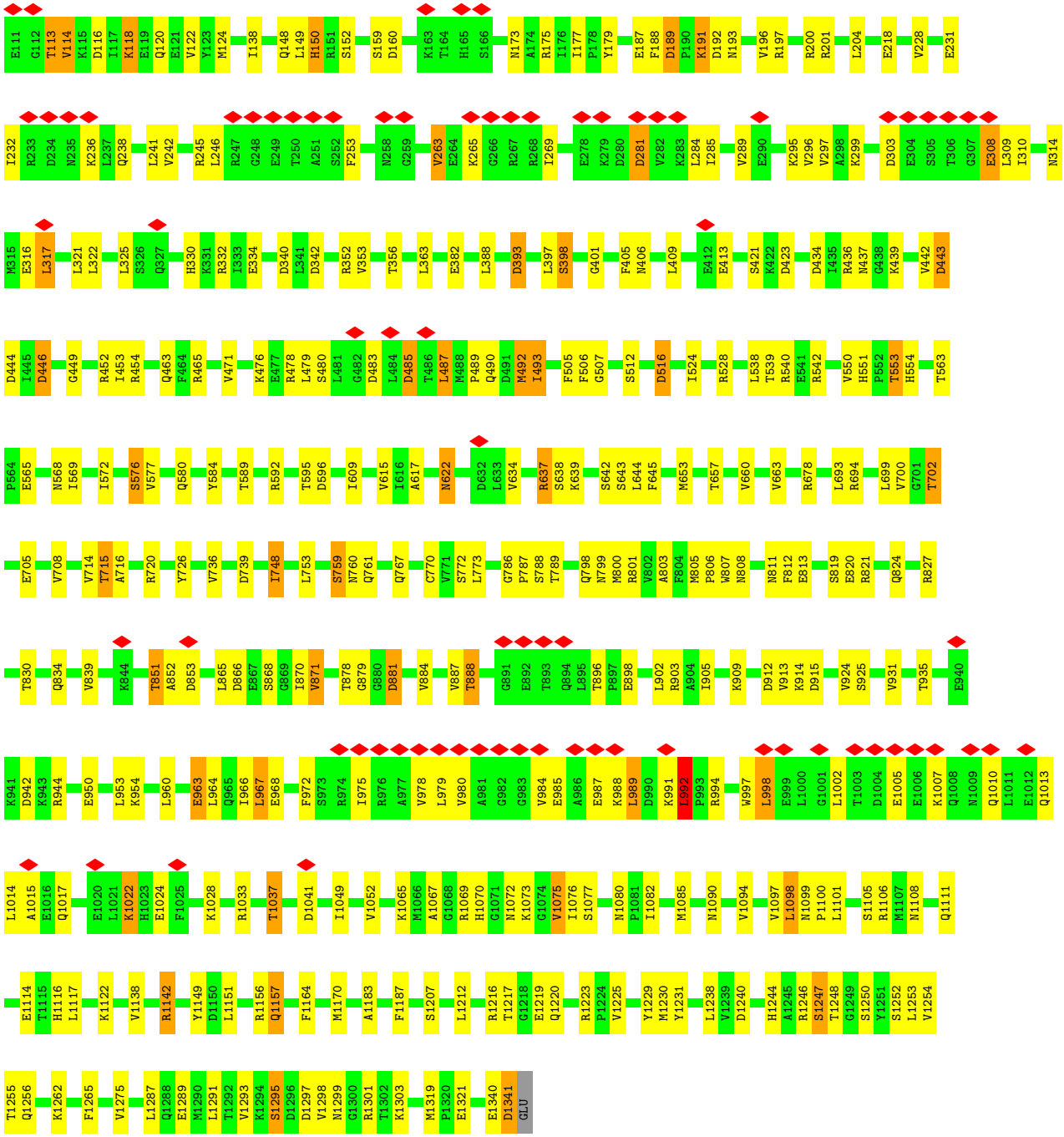


• Molecule 1: DNA-directed RNA polymerase subunit alpha



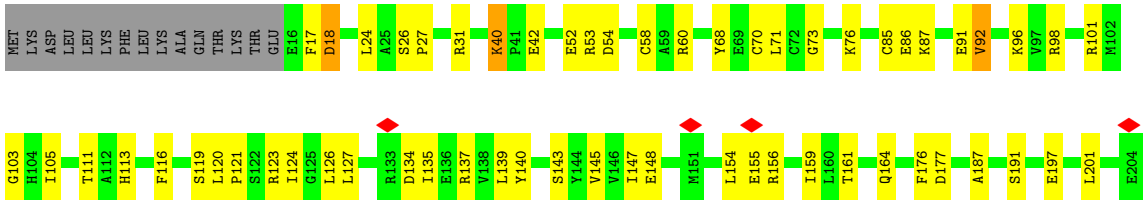
• Molecule 2: DNA-directed RNA polymerase subunit beta

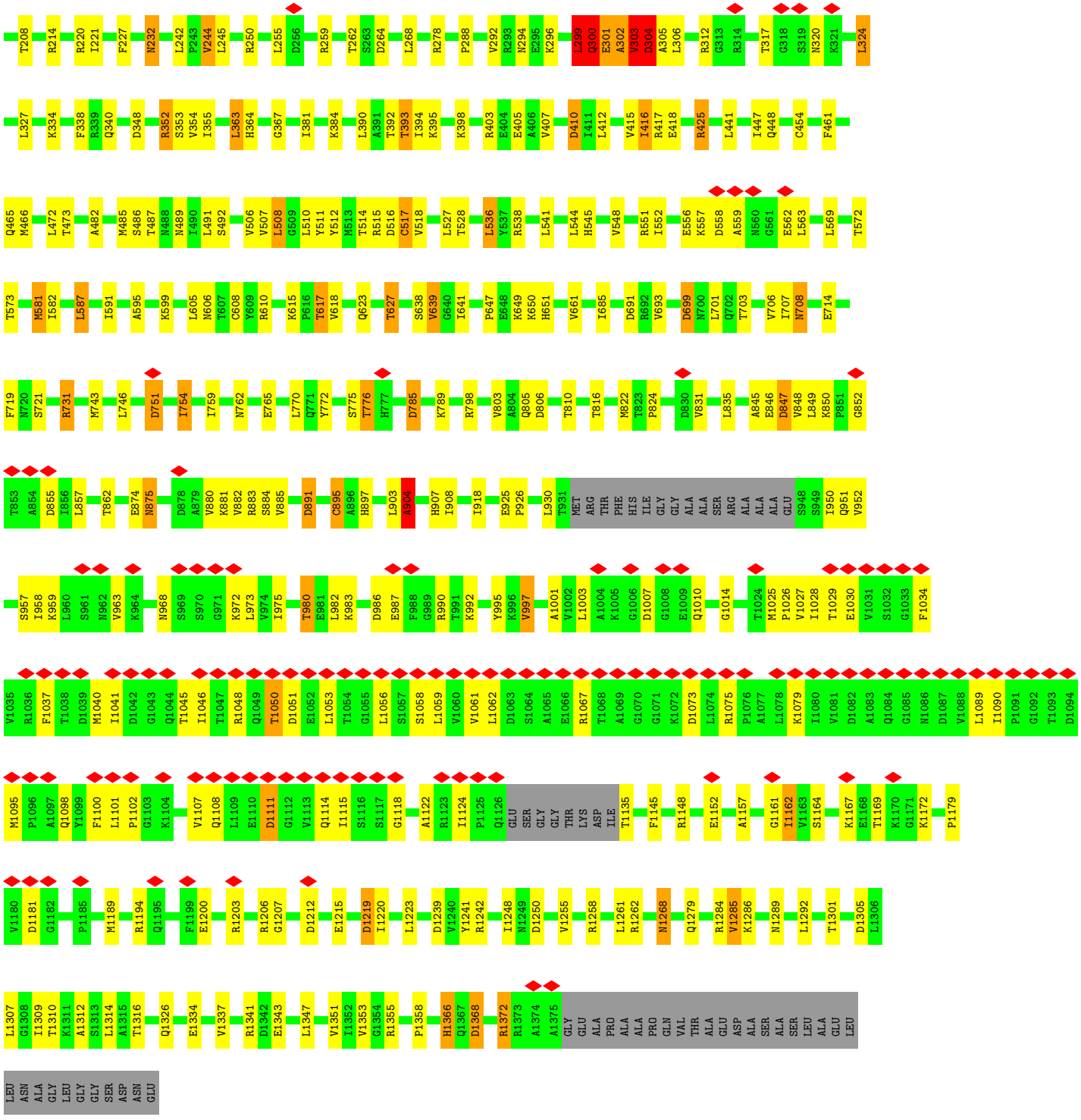




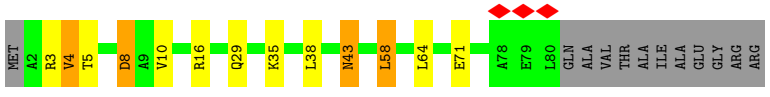
• Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain D:



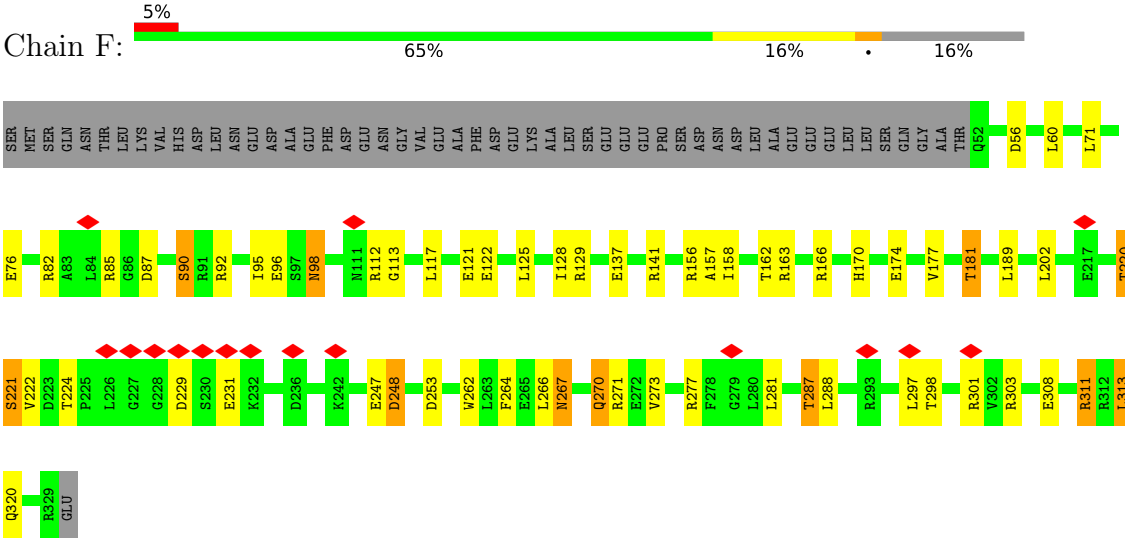


• Molecule 4: DNA-directed RNA polymerase subunit omega



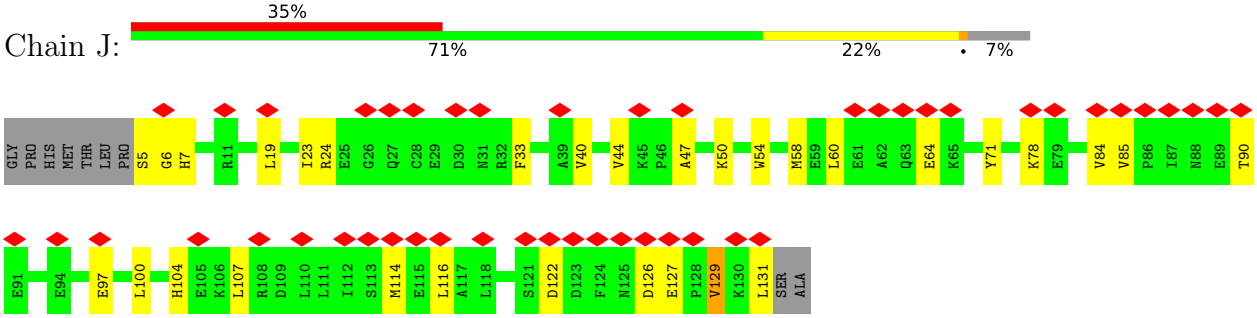
• Molecule 5: RNA polymerase sigma factor RpoS

Chain F:



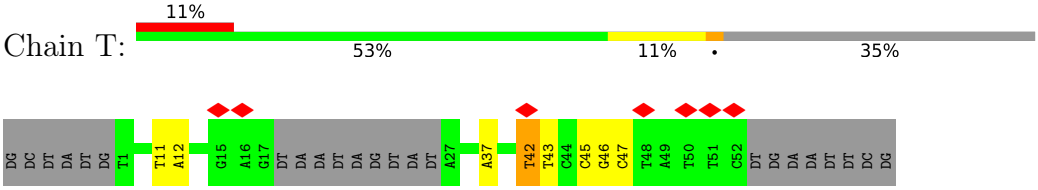
• Molecule 6: Sigma factor-binding protein Crl

Chain J:



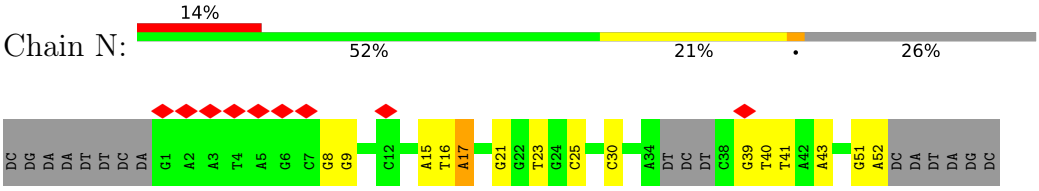
• Molecule 7: Template DNA strand

Chain T:



• Molecule 8: Non-template DNA strand

Chain N:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	292000	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 ($e^-/Å^2$)	1.42	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	3.935	Depositor
Minimum map value	-2.803	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.097	Depositor
Recommended contour level	0.45	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: 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	0.55	0/1828	0.62	0/2479
1	B	0.48	0/1734	0.60	0/2349
2	C	0.58	0/10736	0.59	2/14487 (0.0%)
3	D	0.60	6/10539 (0.1%)	0.62	7/14232 (0.0%)
4	E	0.48	0/629	0.58	0/847
5	F	0.43	0/2273	0.53	0/3065
6	J	0.34	0/1054	0.50	0/1433
7	T	0.93	0/979	1.06	2/1505 (0.1%)
8	N	0.96	0/1132	1.08	2/1744 (0.1%)
All	All	0.59	6/30904 (0.0%)	0.65	13/42141 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
2	C	0	2
3	D	0	3
All	All	0	7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	300	GLN	CA-C	-5.92	1.37	1.52
3	D	302	ALA	CA-CB	-5.85	1.40	1.52
3	D	300	GLN	CG-CD	-5.19	1.39	1.51
3	D	303	VAL	CB-CG2	-5.13	1.42	1.52
3	D	304	ASP	C-O	-5.08	1.13	1.23

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	299	LEU	CB-CG-CD1	-9.40	95.02	111.00
3	D	299	LEU	CB-CG-CD2	-7.89	97.58	111.00
3	D	304	ASP	CB-CG-OD2	7.42	124.98	118.30
3	D	299	LEU	CA-CB-CG	6.97	131.34	115.30
3	D	300	GLN	CB-CA-C	-6.35	97.69	110.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	7	GLU	Peptide
1	B	19	VAL	Peptide
2	C	398	SER	Peptide
2	C	485	ASP	Peptide
3	D	119	SER	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	1806	0	1835	46	0
1	B	1714	0	1748	43	0
2	C	10567	0	10585	202	0
3	D	10382	0	10594	229	0
4	E	627	0	634	10	0
5	F	2244	0	2279	45	0
6	J	1024	0	949	14	0
7	T	876	0	487	7	0
8	N	1008	0	554	11	0
9	D	1	0	0	0	0
10	D	2	0	0	0	0
All	All	30251	0	29665	561	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 561 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)
1:A:236:VAL:HG13	1:B:16:ILE:CG2	1.79	1.13
3:D:300:GLN:NE2	3:D:304:ASP:OD2	1.85	1.08
3:D:299:LEU:O	3:D:302:ALA:N	1.85	1.06
3:D:300:GLN:NE2	3:D:300:GLN:O	1.90	1.05
1:A:233:ASP:O	1:A:235:GLU:OE1	1.80	1.00

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	231/239 (97%)	208 (90%)	22 (10%)	1 (0%)	34 67
1	B	219/239 (92%)	203 (93%)	16 (7%)	0	100 100
2	C	1338/1342 (100%)	1234 (92%)	104 (8%)	0	100 100
3	D	1330/1407 (94%)	1233 (93%)	92 (7%)	5 (0%)	34 67
4	E	77/91 (85%)	69 (90%)	8 (10%)	0	100 100
5	F	276/331 (83%)	266 (96%)	10 (4%)	0	100 100
6	J	125/136 (92%)	118 (94%)	7 (6%)	0	100 100
All	All	3596/3785 (95%)	3331 (93%)	259 (7%)	6 (0%)	50 77

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	300	GLN
3	D	304	ASP
3	D	904	ALA
1	A	8	PHE
3	D	299	LEU

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	201/206 (98%)	168 (84%)	33 (16%)	2 10
1	B	190/206 (92%)	171 (90%)	19 (10%)	7 27
2	C	1155/1157 (100%)	1030 (89%)	125 (11%)	6 24
3	D	1118/1168 (96%)	1009 (90%)	109 (10%)	8 29
4	E	67/75 (89%)	63 (94%)	4 (6%)	19 49
5	F	236/287 (82%)	216 (92%)	20 (8%)	10 34
6	J	104/123 (85%)	89 (86%)	15 (14%)	3 14
All	All	3071/3222 (95%)	2746 (89%)	325 (11%)	10 25

5 of 325 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	691	ASP
4	E	4	VAL
3	D	751	ASP
3	D	987	GLU
5	F	224	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1289	ASN
4	E	31	GLN
6	J	63	GLN
2	C	684	ASN
2	C	649	GLN

5.3.3 RNA (i)

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.

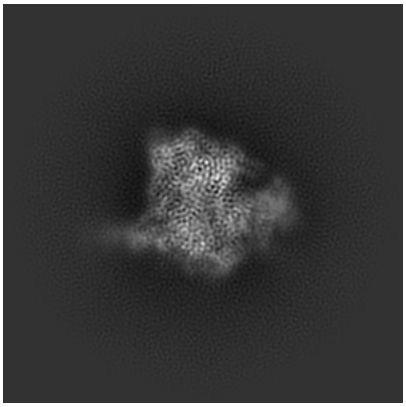
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-20090. These allow visual inspection of the internal detail of the map and identification of artifacts.

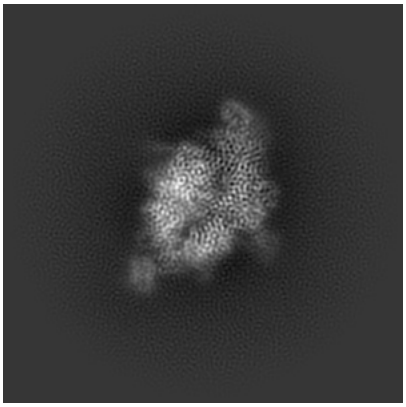
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

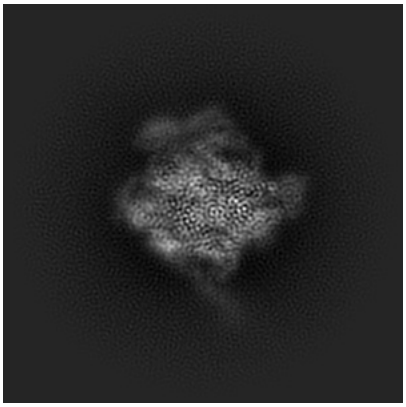
6.1.1 Primary map



X



Y

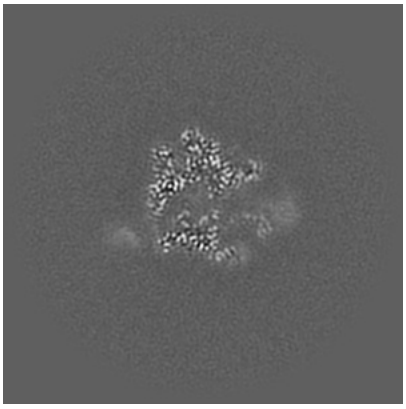


Z

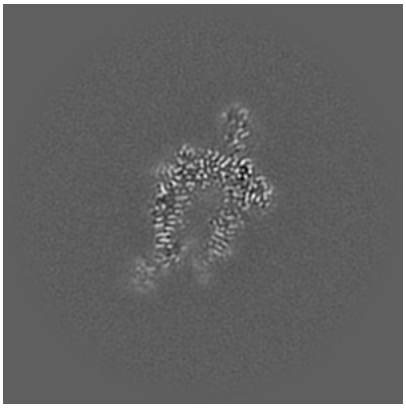
The images above show the map projected in three orthogonal directions.

6.2 Central slices

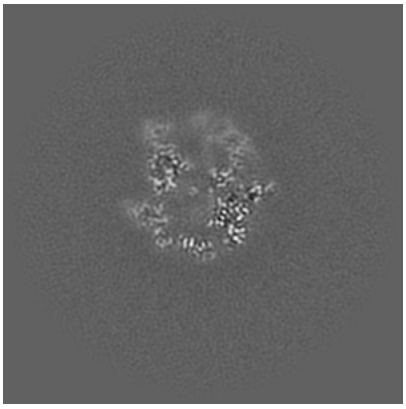
6.2.1 Primary map



X Index: 128



Y Index: 128

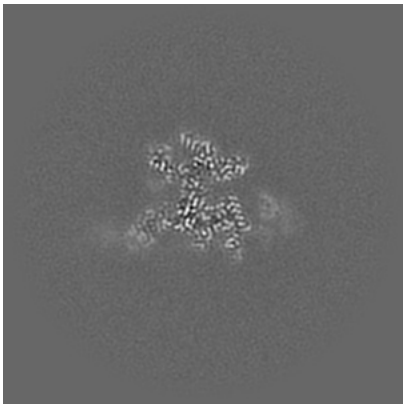


Z Index: 128

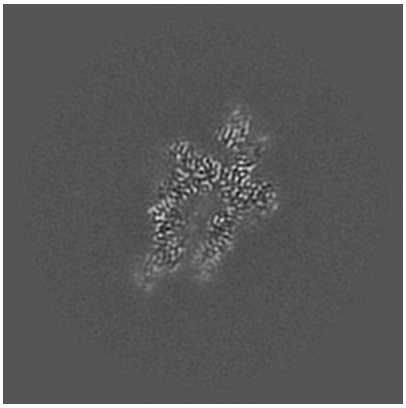
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

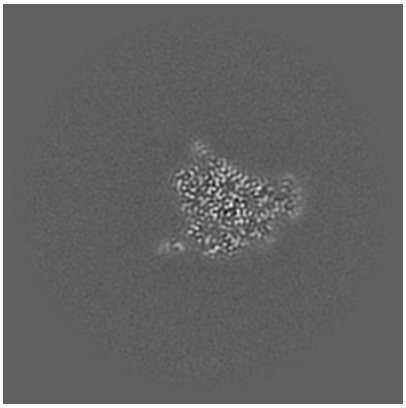
6.3.1 Primary map



X Index: 141



Y Index: 123



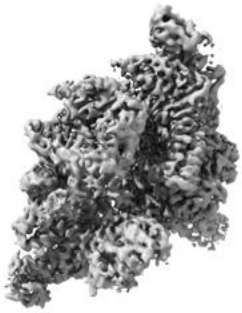
Z Index: 151

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

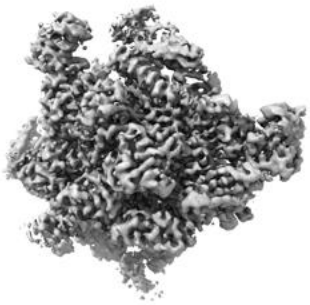
6.4 Orthogonal surface views (i)

6.4.1 Primary map

X



Y



Z

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

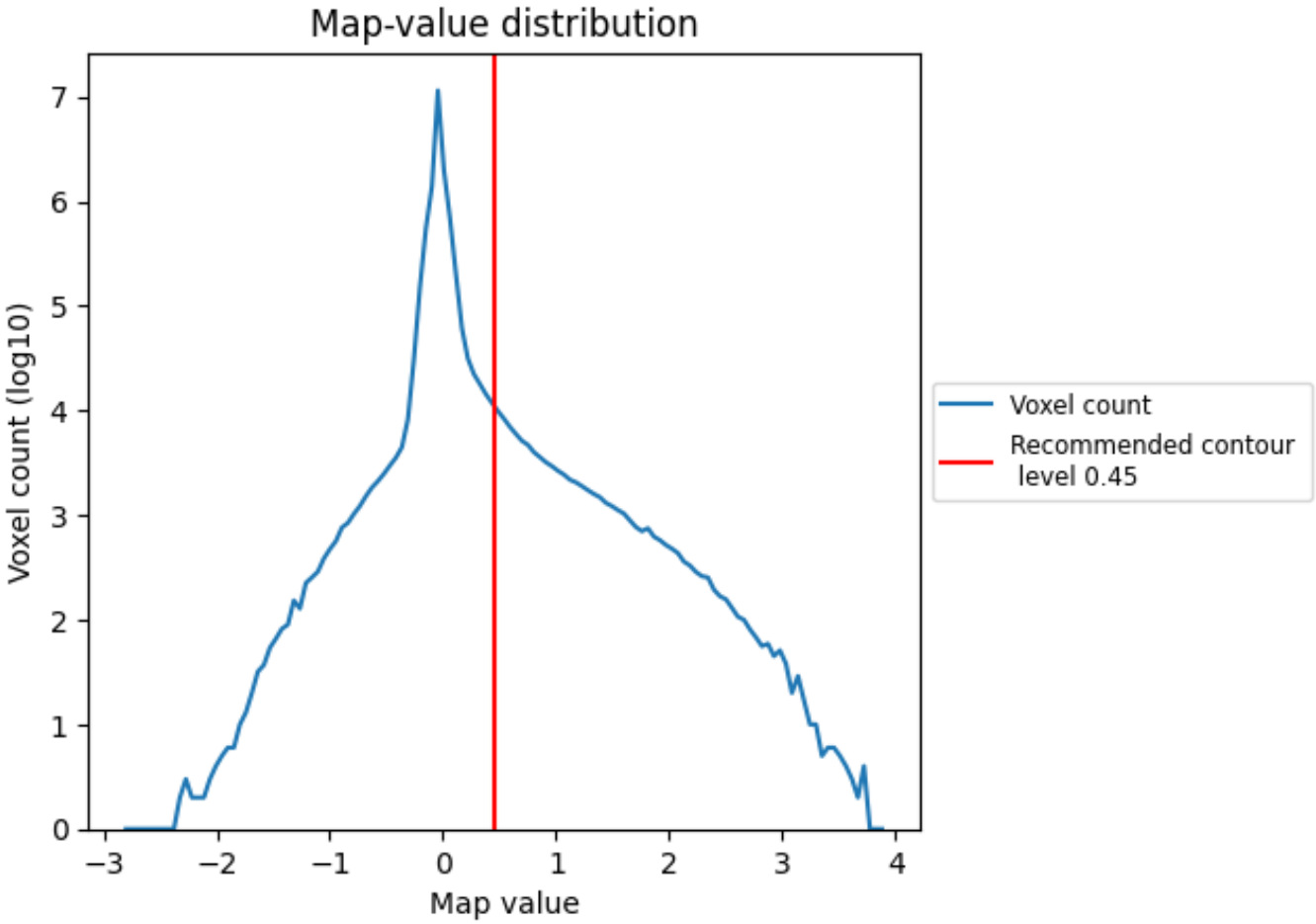
6.5 Mask visualisation (i)

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

7 Map analysis

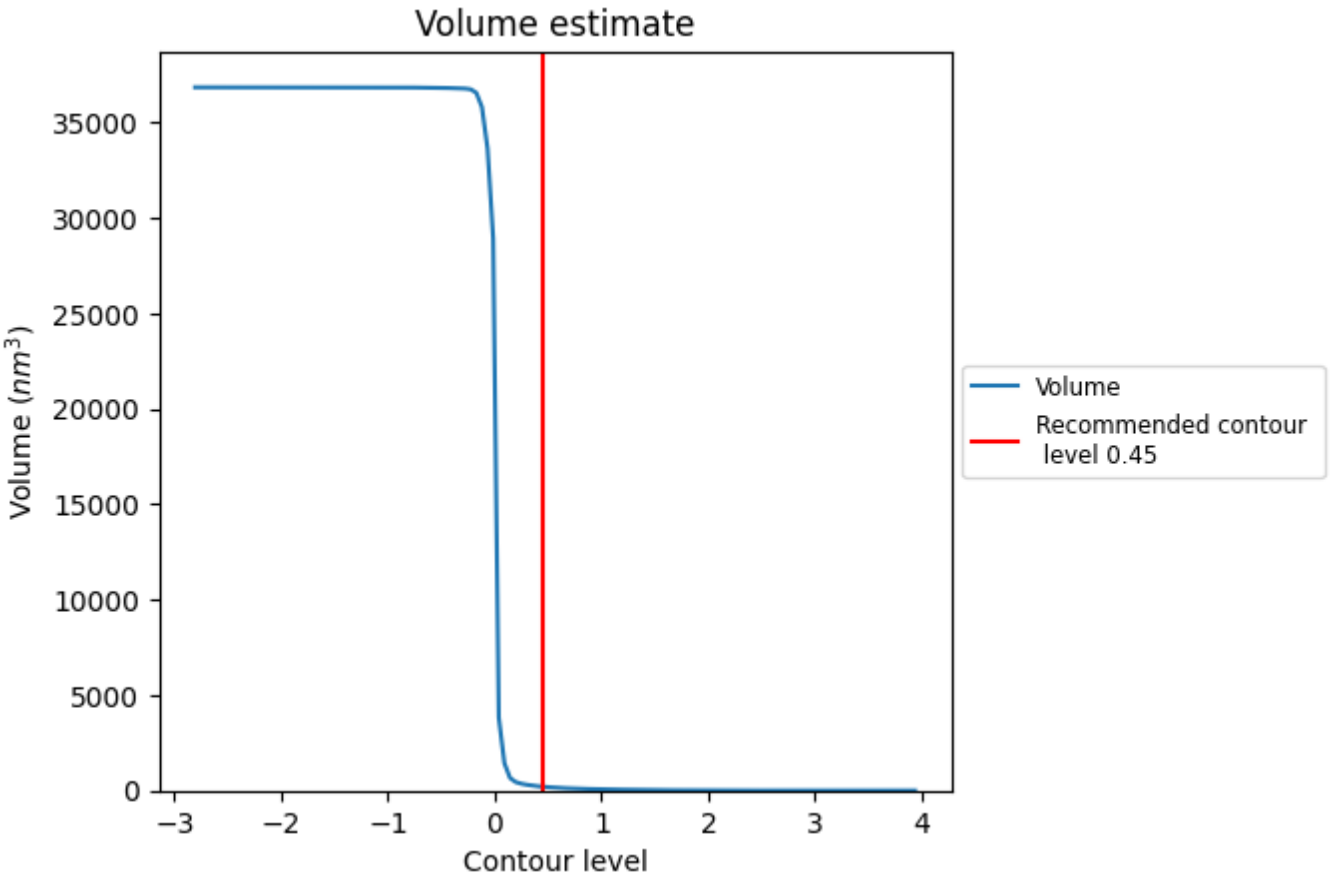
This section contains the results of statistical analysis of the map.

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.

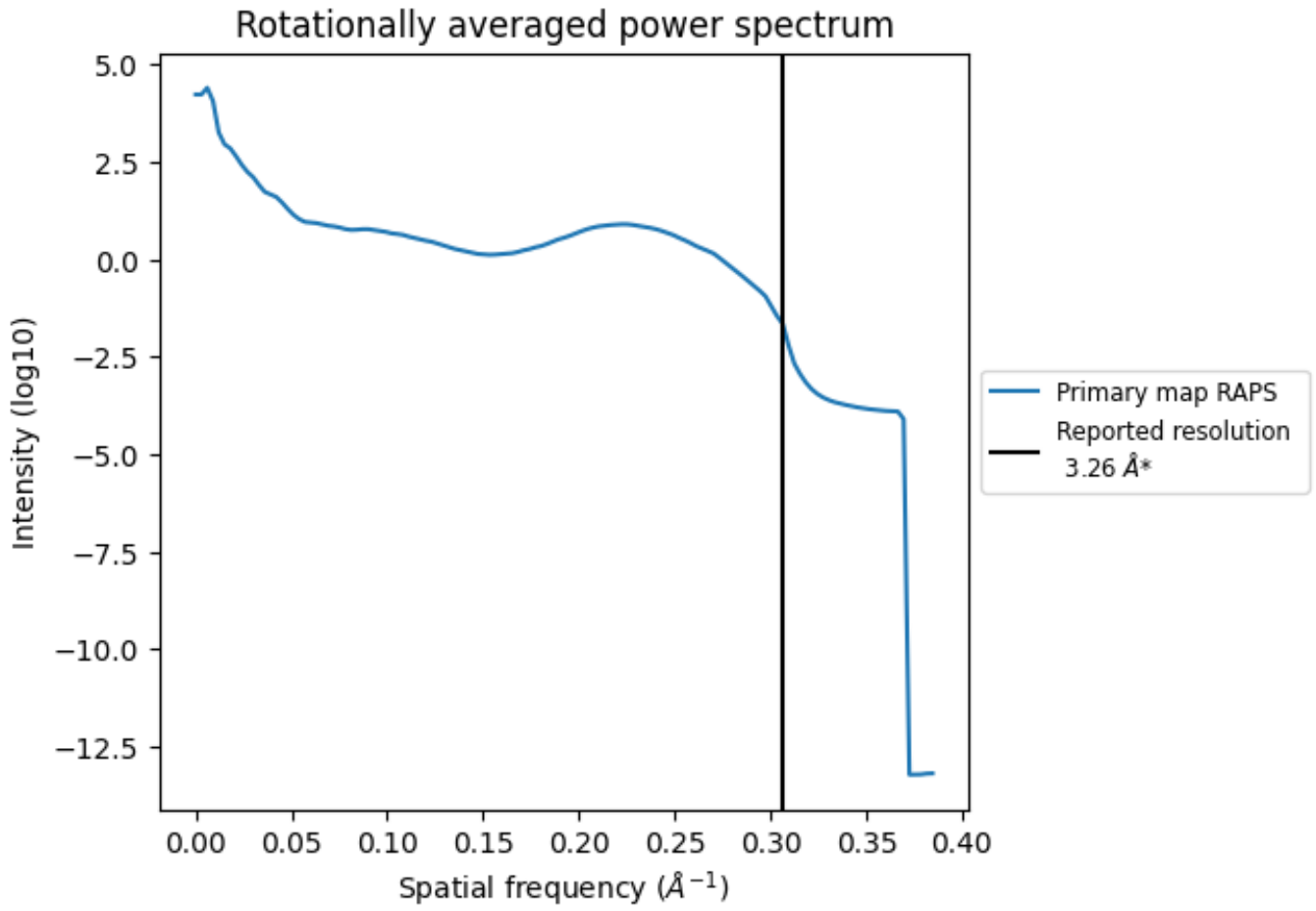
7.2 Volume estimate (i)



The volume at the recommended contour level is 203 nm^3; this corresponds to an approximate mass of 183 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 0.307 $Å^{-1}$

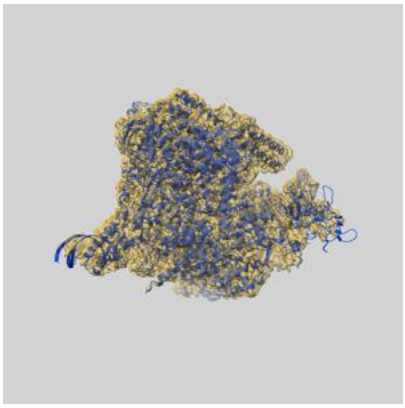
8 Fourier-Shell correlation (i)

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

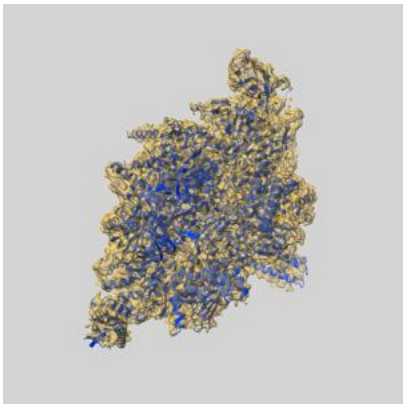
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-20090 and PDB model 6OMF. Per-residue inclusion information can be found in section 3 on page 7.

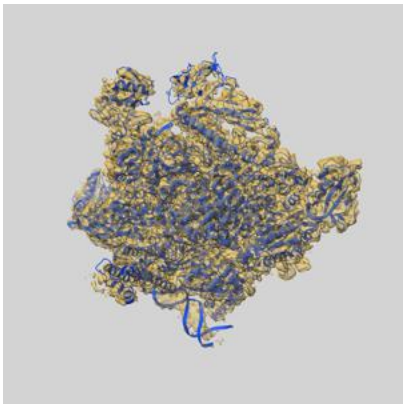
9.1 Map-model overlay



X



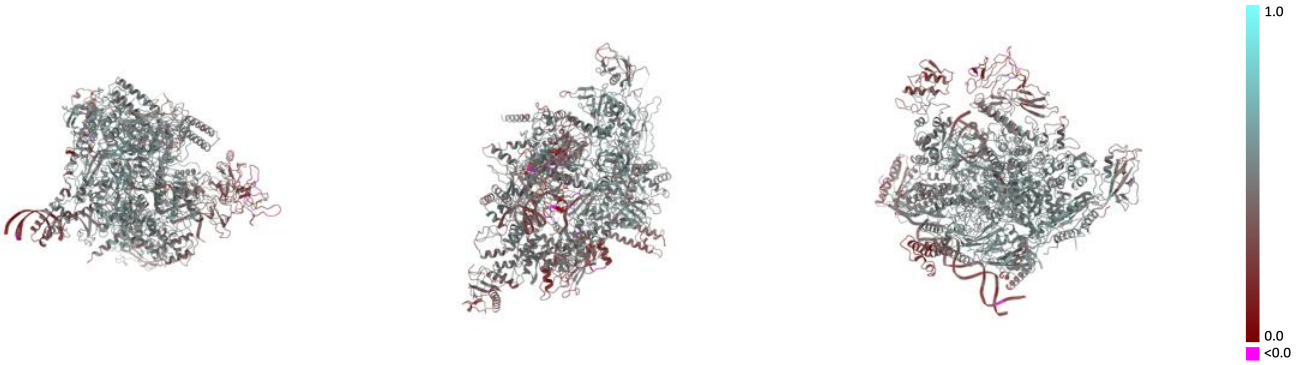
Y



Z

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.

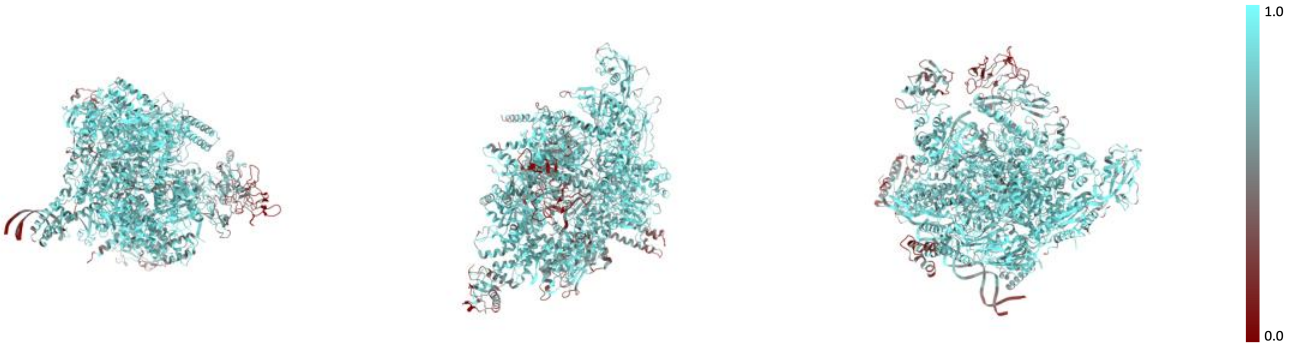
9.2 Q-score mapped to coordinate model (i)



X Y Z

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.

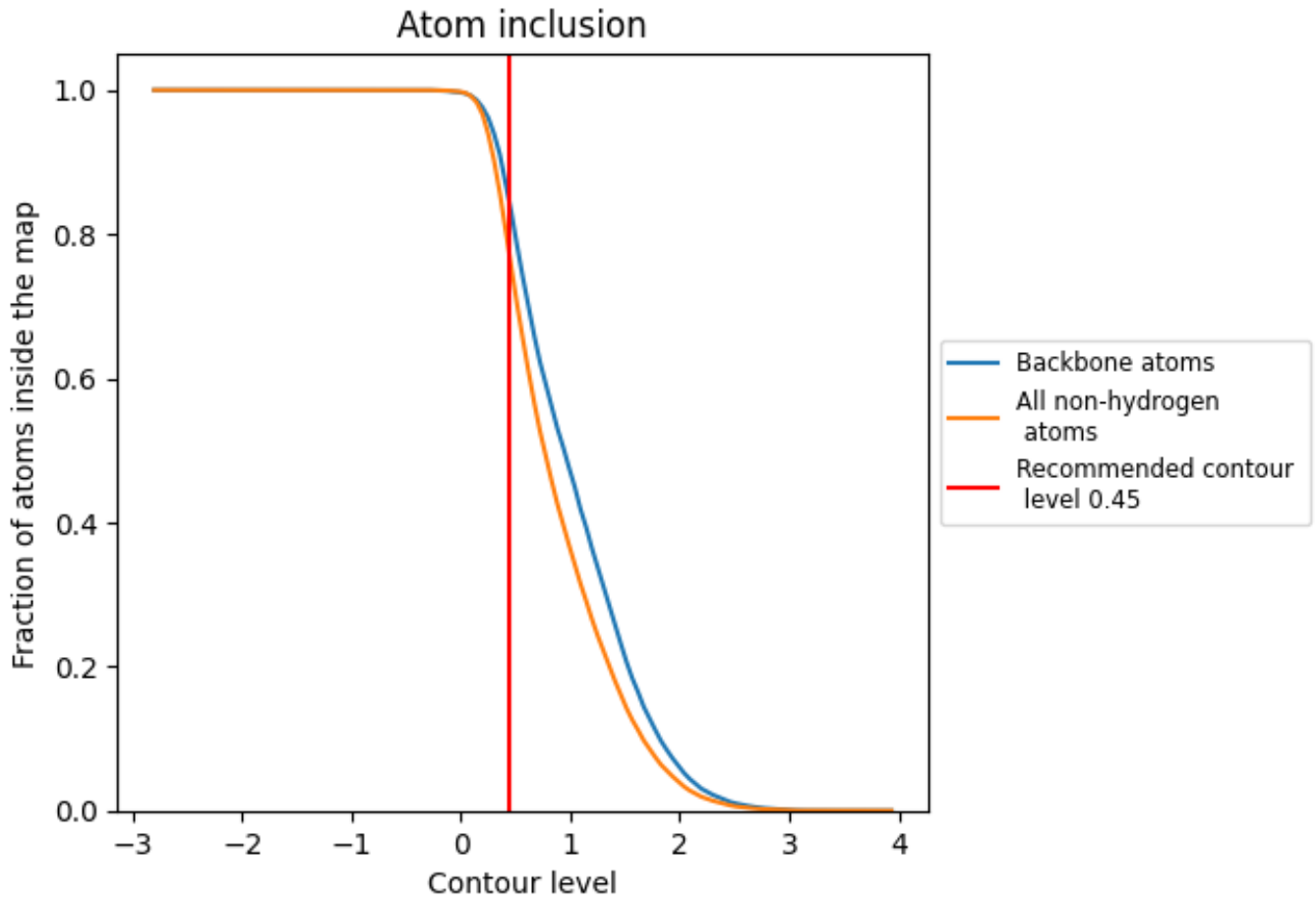
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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, 84% of all backbone atoms, 77% 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 (0.45) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7651	0.4630
A	0.8209	0.5020
B	0.7810	0.4680
C	0.7877	0.4830
D	0.7680	0.4740
E	0.7643	0.4840
F	0.7685	0.4560
J	0.5189	0.3690
N	0.6736	0.3060
T	0.7021	0.3110