wwPDB EM Validation Summary Report

Nov 20, 2022 – 02:06 PM EST

PDB ID	:	3JBW
EMDB ID	:	EMD-6489
Title	:	Cryo-electron microscopy structure of RAG Paired Complex (with NBD, no symmetry)
Authors	:	Ru, H.; Chambers, M.G.; Fu, T.-M.; Tong, A.B.; Liao, M.; Wu, H.
Deposited on	:	2015-10-21
Resolution	:	4.63 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:

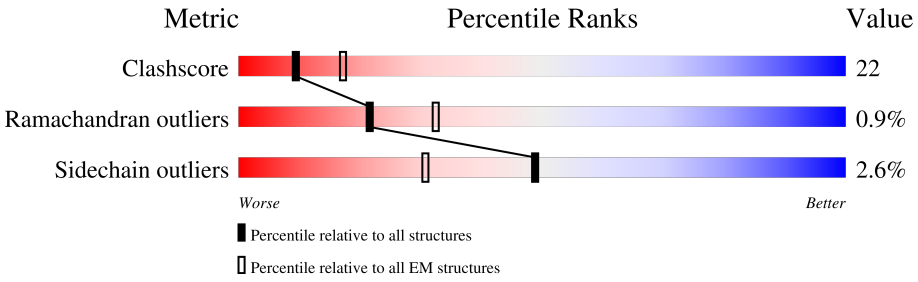
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 4.63 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.



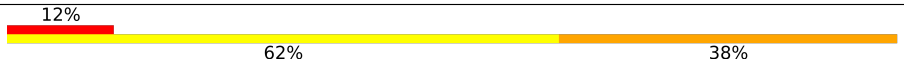
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	764	12% 51% 29% • 19%
1	C	764	12% 49% 31% •• 19%
2	B	533	14% 37% 28% • 34%
2	D	533	14% 39% 26% • 34%
3	E	34	26% 15% 41% 44%
4	F	50	22% 32% 50% 18%
5	G	61	30% 15% 66% 20%
6	H	45	42% 29% 42% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	I	16	
7	J	16	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 19997 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 V(D)J recombination-activating protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	622	Total	C	N	O	S	0	0
			5008	3134	901	936	37		
1	C	622	Total	C	N	O	S	0	0
			5008	3134	901	936	37		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	268	GLY	-	EXPRESSION TAG	UNP O13033
A	269	GLY	-	EXPRESSION TAG	UNP O13033
A	270	SER	-	EXPRESSION TAG	UNP O13033
C	268	GLY	-	EXPRESSION TAG	UNP O13033
C	269	GLY	-	EXPRESSION TAG	UNP O13033
C	270	SER	-	EXPRESSION TAG	UNP O13033

- Molecule 2 is a protein called V(D)J recombination-activating protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	351	Total	C	N	O	S	0	0
			2714	1716	470	509	19		
2	D	351	Total	C	N	O	S	0	0
			2714	1716	470	509	19		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	EXPRESSION TAG	UNP Q1RLW7
B	-1	GLY	-	EXPRESSION TAG	UNP Q1RLW7
B	0	SER	-	EXPRESSION TAG	UNP Q1RLW7
D	-2	GLY	-	EXPRESSION TAG	UNP Q1RLW7
D	-1	GLY	-	EXPRESSION TAG	UNP Q1RLW7
D	0	SER	-	EXPRESSION TAG	UNP Q1RLW7

- Molecule 3 is a DNA chain called 12-RSS signal end forward strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	34	Total	C	N	O	P	0	0
			697	330	138	195	34		

- Molecule 4 is a DNA chain called Nicked 12-RSS intermediate reverse strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	50	Total	C	N	O	P	0	0
			1027	489	183	305	50		

- Molecule 5 is a DNA chain called Nicked 23-RSS intermediate reverse strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	G	61	Total	C	N	O	P	0	0
			1256	596	235	364	61		

- Molecule 6 is a DNA chain called Nicked 23-RSS intermediate forward strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	45	Total	C	N	O	P	0	0
			919	437	169	268	45		

- Molecule 7 is a DNA chain called 5'-D(P*GP*AP*TP*CP*TP*GP*GP*CP*CP*TP*GP*TP*CP*TP*TP*A)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	16	Total	C	N	O	P	0	0
			326	156	54	100	16		
7	J	16	Total	C	N	O	P	0	0
			326	156	54	100	16		

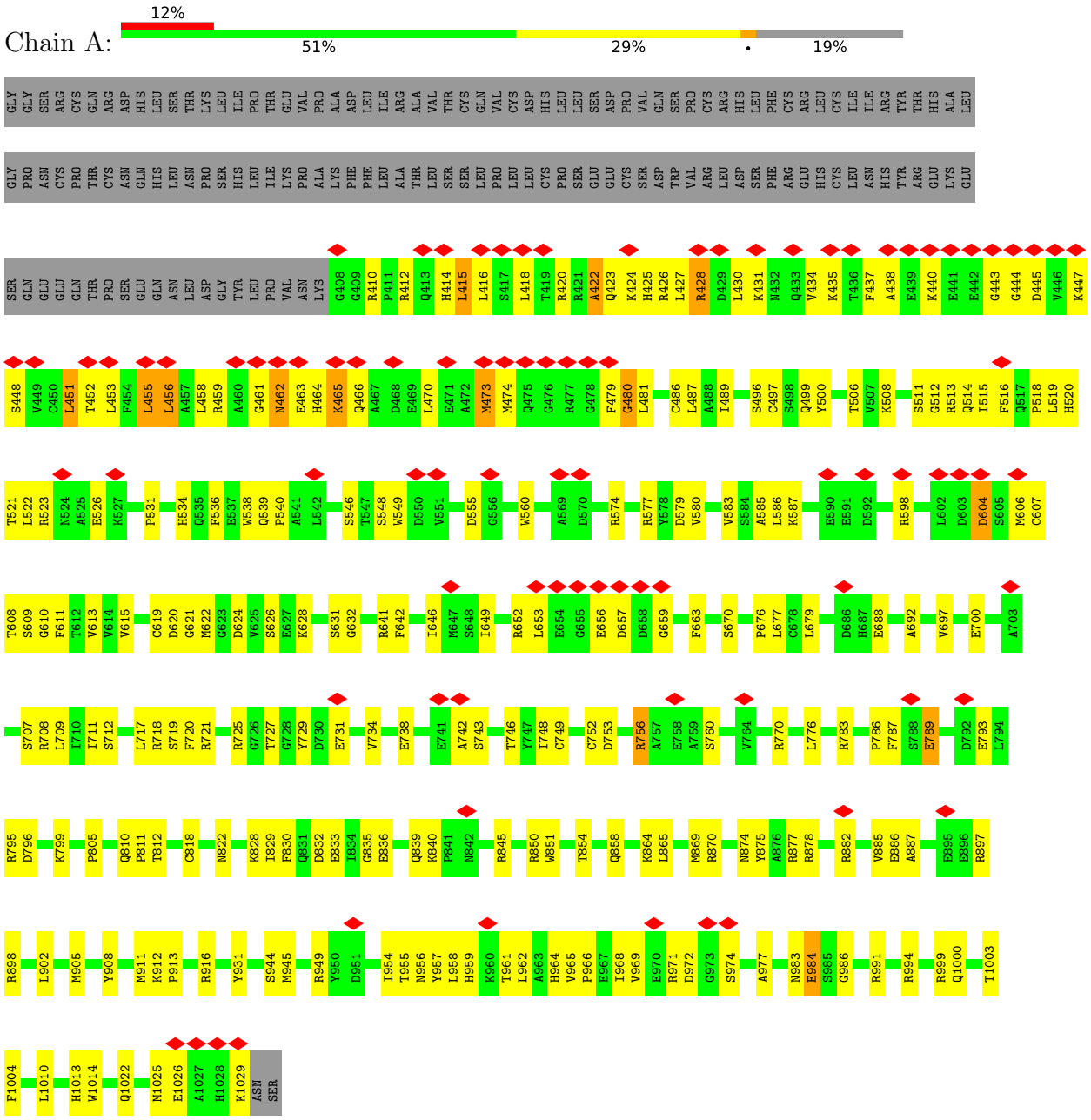
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
8	A	1	Total	Zn	0
			1	1	
8	C	1	Total	Zn	0
			1	1	

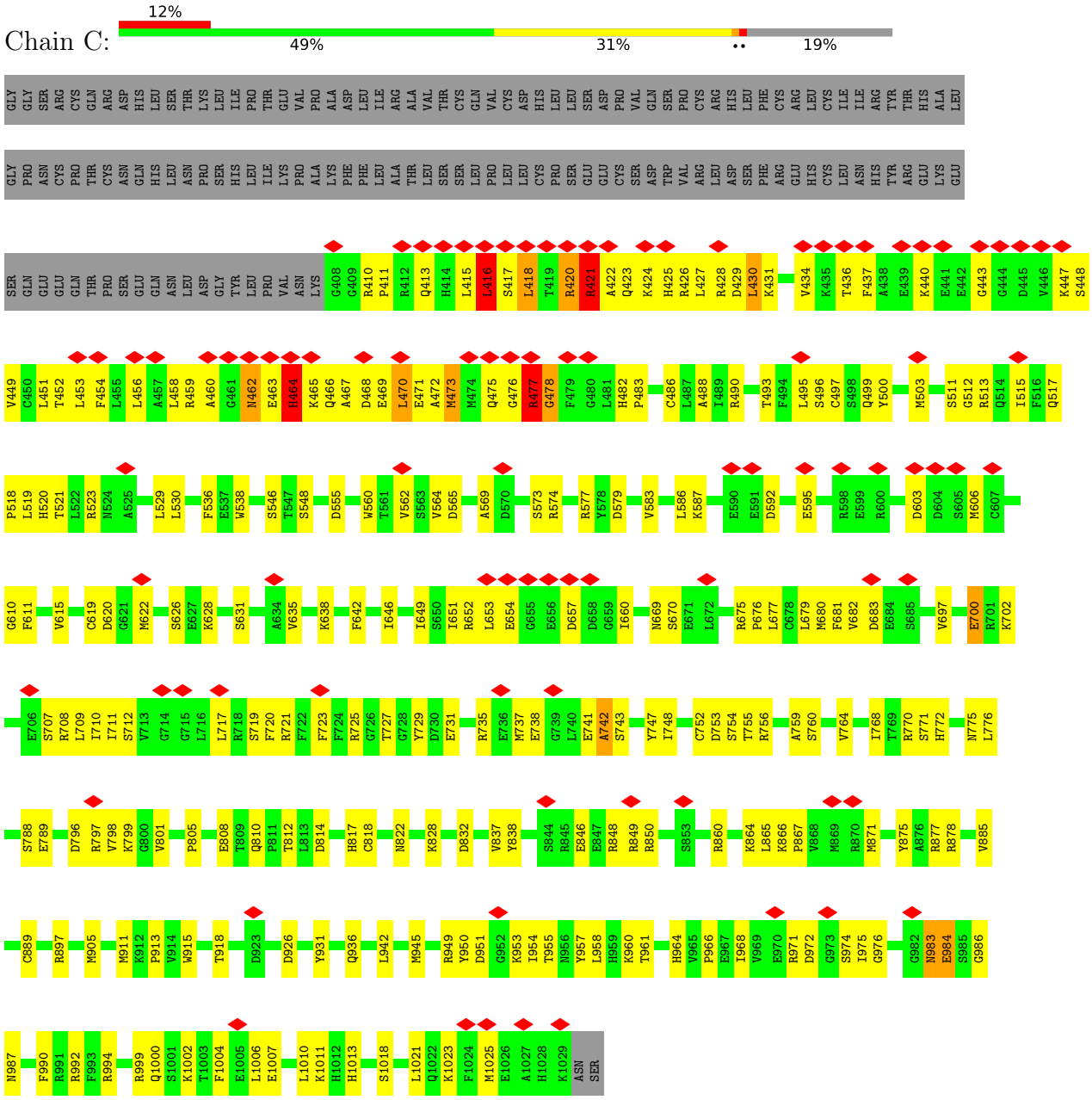
3 Residue-property plots (i)

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.

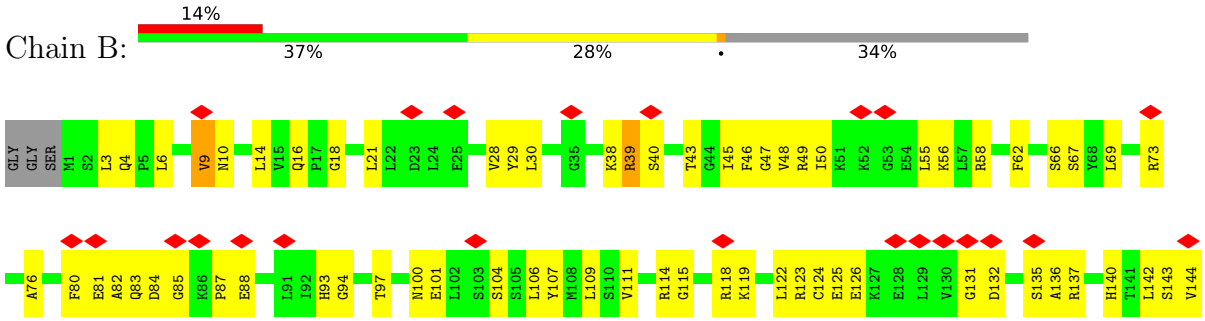
- Molecule 1: V(D)J recombination-activating protein 1

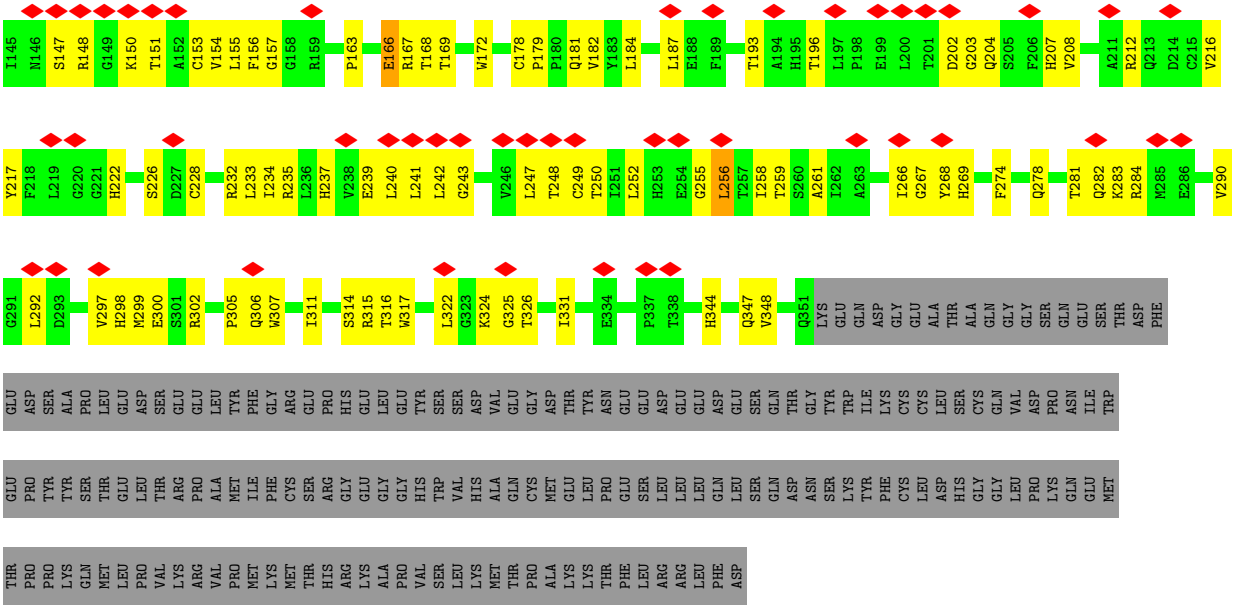


• Molecule 1: V(D)J recombination-activating protein 1

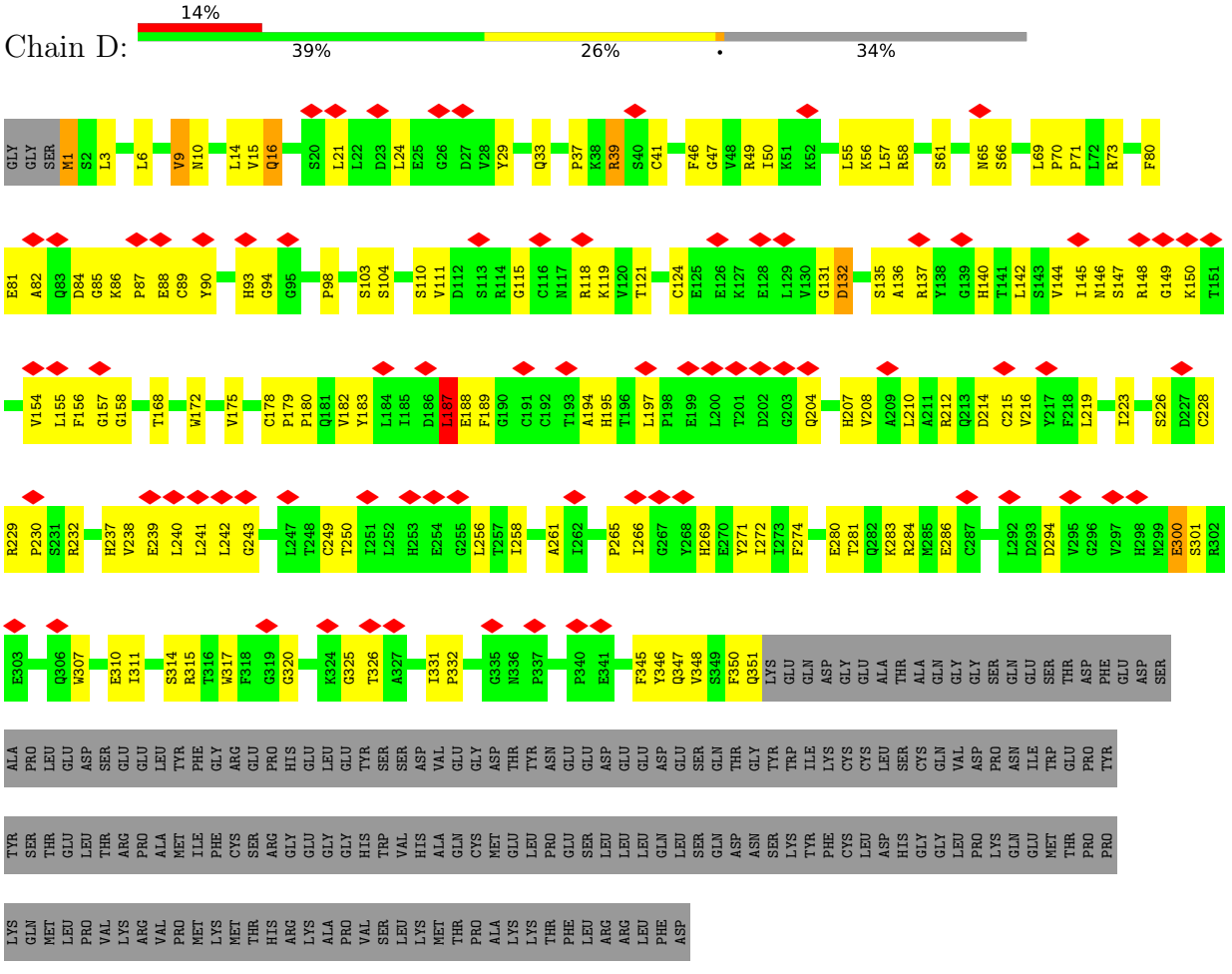


• Molecule 2: V(D)J recombination-activating protein 2



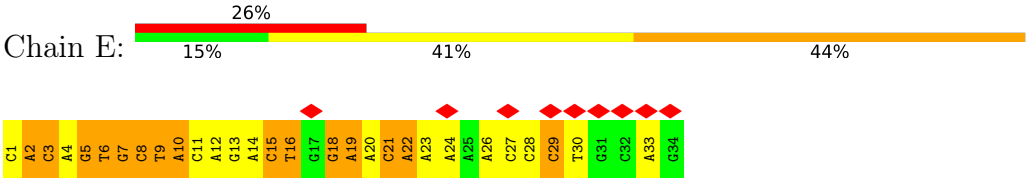


• Molecule 2: V(D)J recombination-activating protein 2



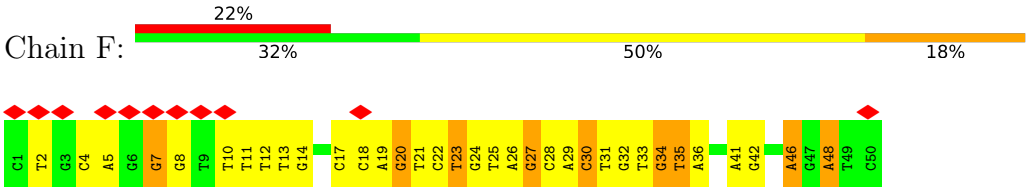
• Molecule 3: 12-RSS signal end forward strand

Chain E:



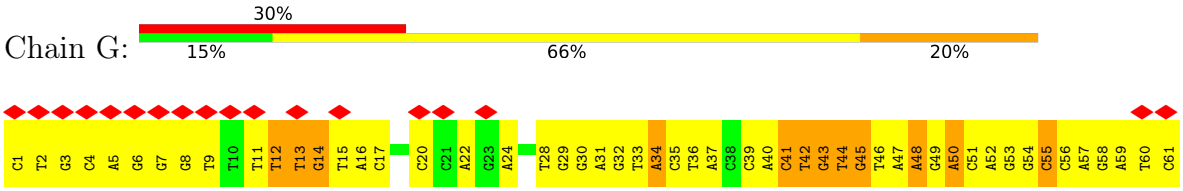
• Molecule 4: Nicked 12-RSS intermediate reverse strand

Chain F:



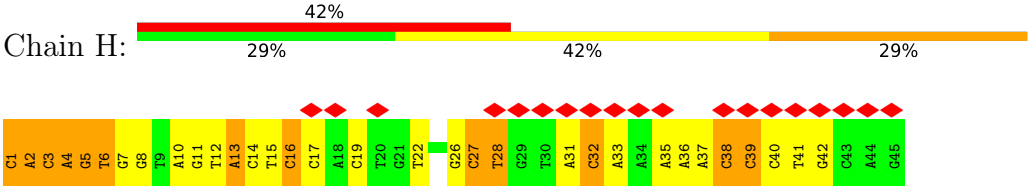
• Molecule 5: Nicked 23-RSS intermediate reverse strand

Chain G:



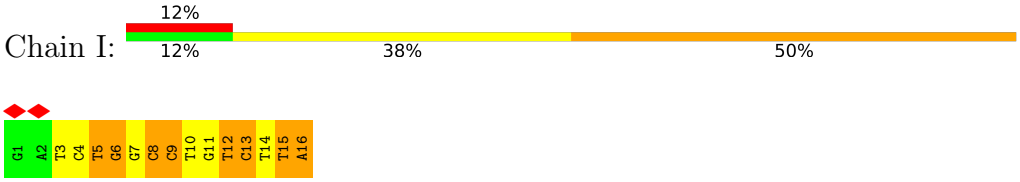
• Molecule 6: Nicked 23-RSS intermediate forward strand

Chain H:



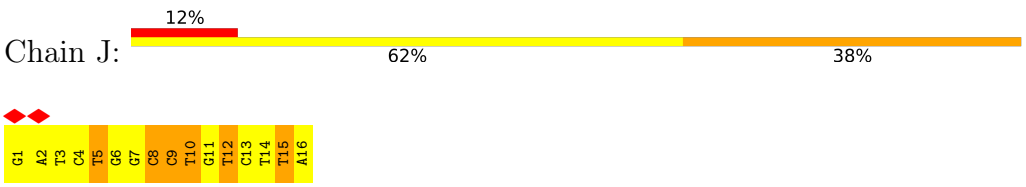
• Molecule 7: 5'-D(P*GP*AP*TP*CP*TP*GP*GP*CP*CP*TP*GP*TP*CP*TP*TP*A)-3
'

Chain I:



• Molecule 7: 5'-D(P*GP*AP*TP*CP*TP*GP*GP*CP*CP*TP*GP*TP*CP*TP*TP*A)-3
'

Chain J:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	14129	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Each particle	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	41	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	40607	Depositor
Image detector	GATAN K2 (4k x 4k)	Depositor
Maximum map value	0.096	Depositor
Minimum map value	-0.049	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.025	Depositor
Map size (Å)	236.16, 236.16, 236.16	wwPDB
Map dimensions	192, 192, 192	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.23, 1.23, 1.23	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

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	0.47	1/5106 (0.0%)	0.77	5/6867 (0.1%)
1	C	0.48	1/5106 (0.0%)	0.82	11/6867 (0.2%)
2	B	0.40	0/2784	0.66	1/3784 (0.0%)
2	D	0.41	0/2784	0.65	1/3784 (0.0%)
3	E	1.65	15/784 (1.9%)	1.33	11/1206 (0.9%)
4	F	1.42	8/1150 (0.7%)	1.32	7/1774 (0.4%)
5	G	1.74	24/1410 (1.7%)	1.41	23/2175 (1.1%)
6	H	2.29	44/1030 (4.3%)	1.65	26/1586 (1.6%)
7	I	2.15	13/363 (3.6%)	1.54	6/558 (1.1%)
7	J	1.57	6/363 (1.7%)	1.31	4/558 (0.7%)
All	All	0.98	112/20880 (0.5%)	0.98	95/29159 (0.3%)

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	C	0	2
All	All	0	3

The worst 5 of 112 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	42	DT	P-O5'	-19.75	1.40	1.59
6	H	2	DA	C3'-O3'	19.18	1.68	1.44
5	G	45	DG	C3'-O3'	17.23	1.66	1.44
6	H	5	DG	C3'-O3'	17.20	1.66	1.44
6	H	3	DC	N3-C4	15.62	1.44	1.33

The worst 5 of 95 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	H	6	DT	O5'-P-OP2	22.18	137.32	110.70
6	H	3	DC	OP1-P-OP2	-14.91	97.24	119.60
6	H	3	DC	O4'-C1'-N1	12.66	116.86	108.00
5	G	41	DC	OP2-P-O3'	11.06	129.54	105.20
5	G	42	DT	O5'-P-OP1	-10.51	96.25	105.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	462	ASN	Peptide
1	C	477	ARG	Peptide
1	C	742	ALA	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	5008	0	4968	224	0
1	C	5008	0	4968	233	0
2	B	2714	0	2665	124	0
2	D	2714	0	2665	115	0
3	E	697	0	379	36	0
4	F	1027	0	566	57	0
5	G	1256	0	686	82	0
6	H	919	0	506	60	0
7	I	326	0	183	32	0
7	J	326	0	183	36	0
8	A	1	0	0	0	0
8	C	1	0	0	0	0
All	All	19997	0	17769	828	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 22.

The worst 5 of 828 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)
6:H:5:DG:C3'	6:H:5:DG:O3'	1.66	1.44
5:G:45:DG:O3'	5:G:45:DG:C3'	1.66	1.44
6:H:2:DA:C3'	6:H:2:DA:O3'	1.68	1.38
2:D:58:ARG:HH12	7:I:6:DG:H5"	1.20	1.02
5:G:12:DT:H2"	5:G:13:DT:H5'	1.41	1.01

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	620/764 (81%)	574 (93%)	38 (6%)	8 (1%)	12	48
1	C	620/764 (81%)	572 (92%)	42 (7%)	6 (1%)	15	54
2	B	349/533 (66%)	332 (95%)	15 (4%)	2 (1%)	25	65
2	D	349/533 (66%)	335 (96%)	13 (4%)	1 (0%)	41	76
All	All	1938/2594 (75%)	1813 (94%)	108 (6%)	17 (1%)	21	56

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	424	LYS
1	A	465	LYS
1	A	473	MET
1	C	416	LEU
1	C	464	HIS

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	548/681 (80%)	538 (98%)	10 (2%)	59	77
1	C	548/681 (80%)	528 (96%)	20 (4%)	35	59
2	B	303/465 (65%)	296 (98%)	7 (2%)	50	70
2	D	303/465 (65%)	295 (97%)	8 (3%)	46	67
All	All	1702/2292 (74%)	1657 (97%)	45 (3%)	49	67

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	649	ILE
1	C	889	CYS
1	C	677	LEU
1	C	707	SER
2	D	1	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	466	GLN
1	C	669	ASN
1	C	1000	GLN
1	A	983	ASN
2	B	100	ASN

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

Of 2 ligands modelled in this entry, 2 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	H	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	2:DA	O3'	3:DC	P	1.77

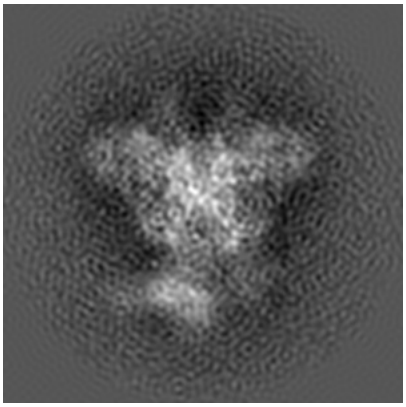
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-6489. These allow visual inspection of the internal detail of the map and identification of artifacts.

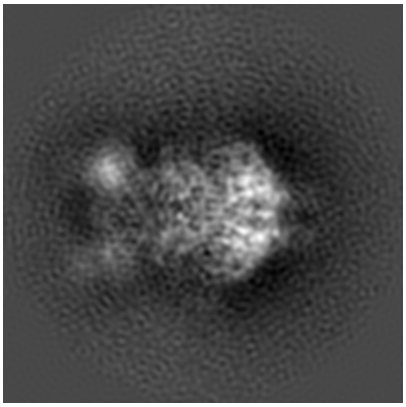
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

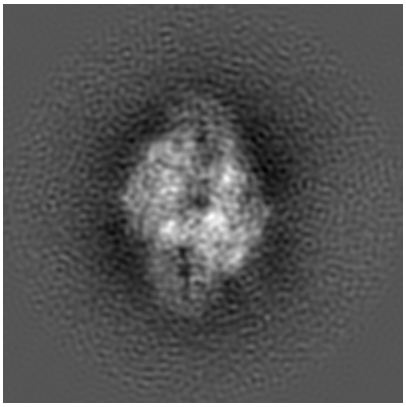
6.1.1 Primary map



X



Y

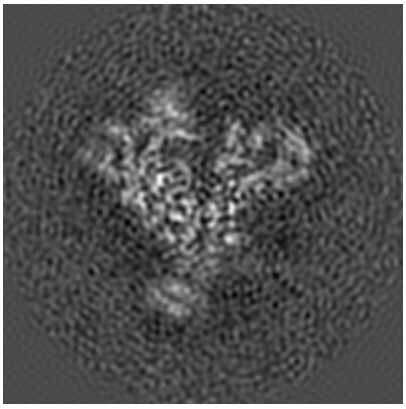


Z

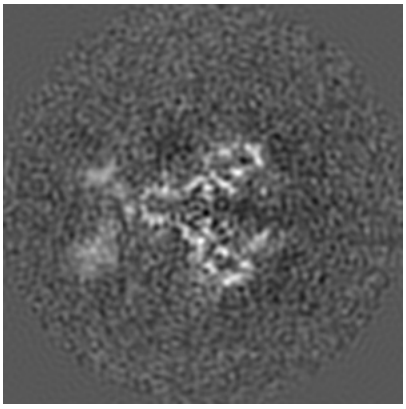
The images above show the map projected in three orthogonal directions.

6.2 Central slices

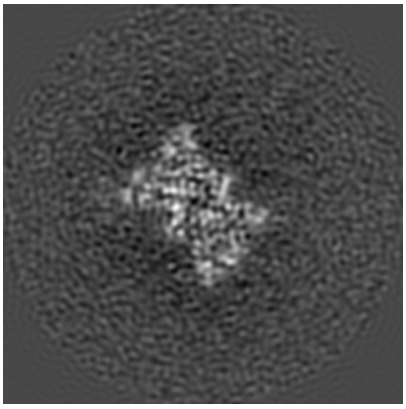
6.2.1 Primary map



X Index: 96



Y Index: 96

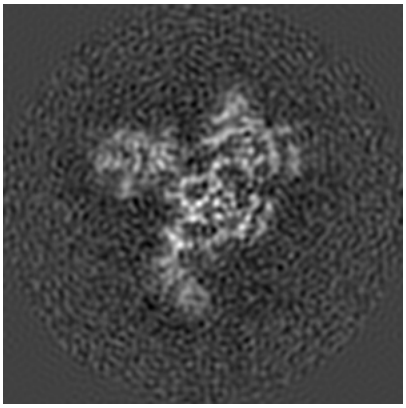


Z Index: 96

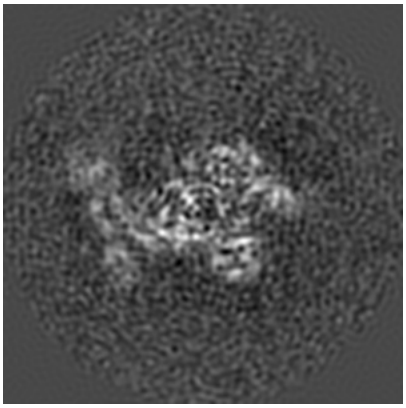
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

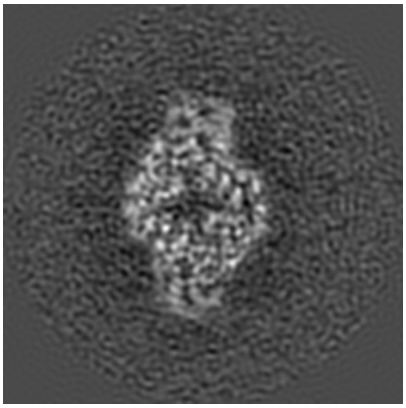
6.3.1 Primary map



X Index: 82



Y Index: 84



Z Index: 115

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 0.025. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

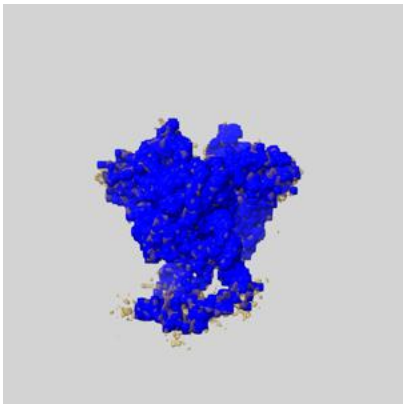
6.5 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

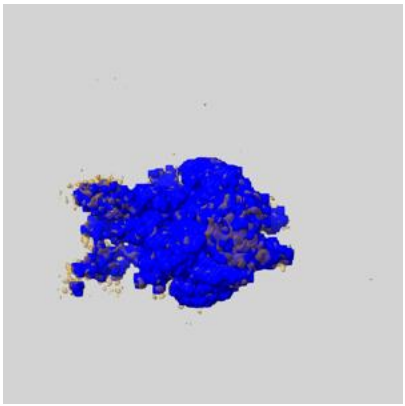
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

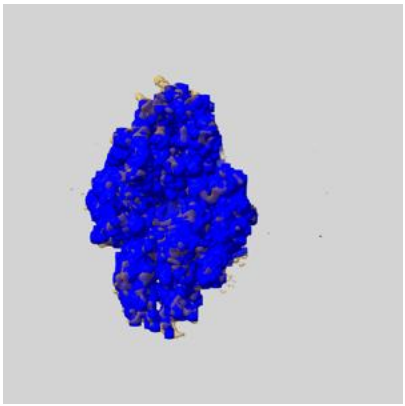
6.5.1 emd_6489_msk.map (i)



X



Y

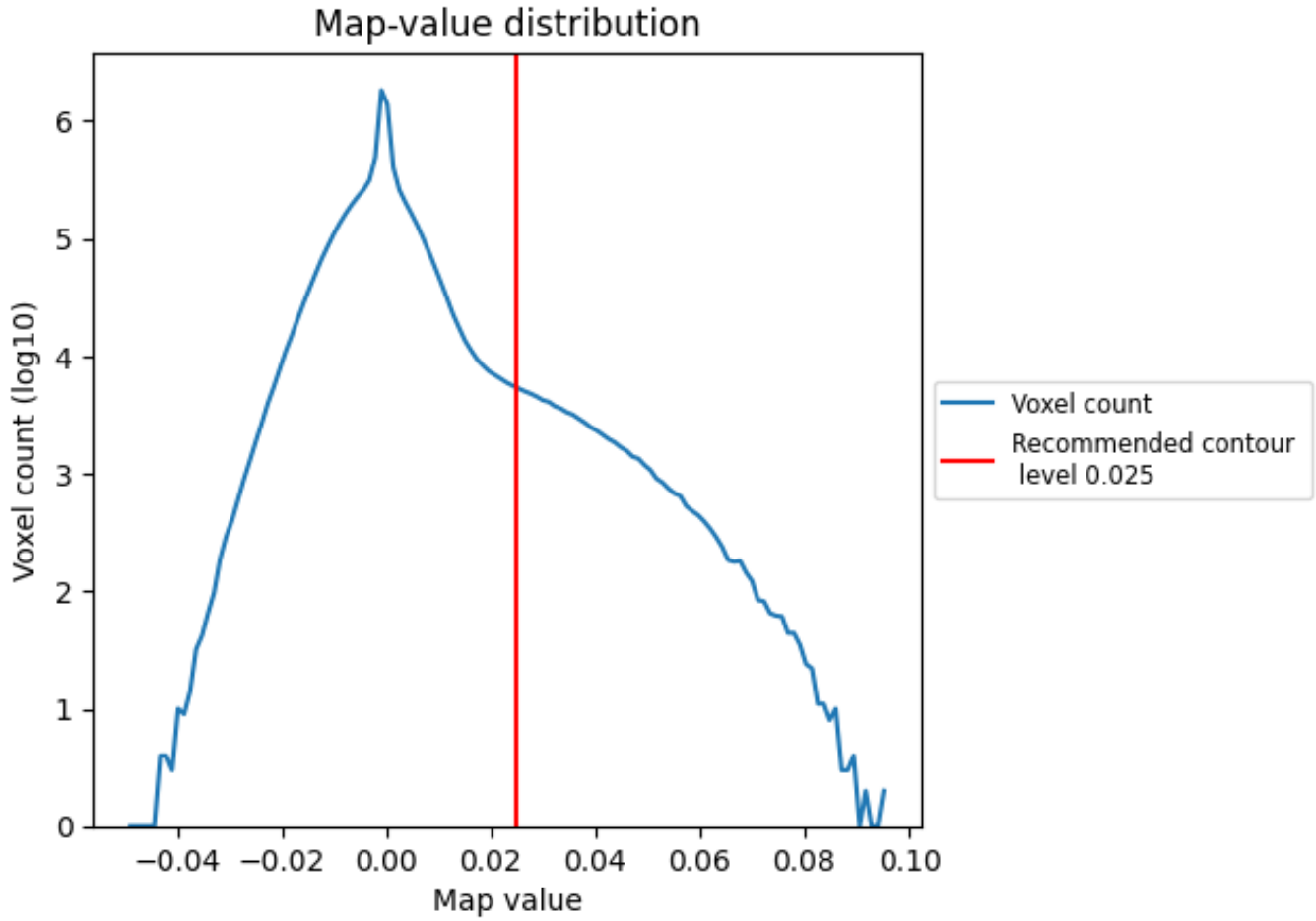


Z

7 Map analysis

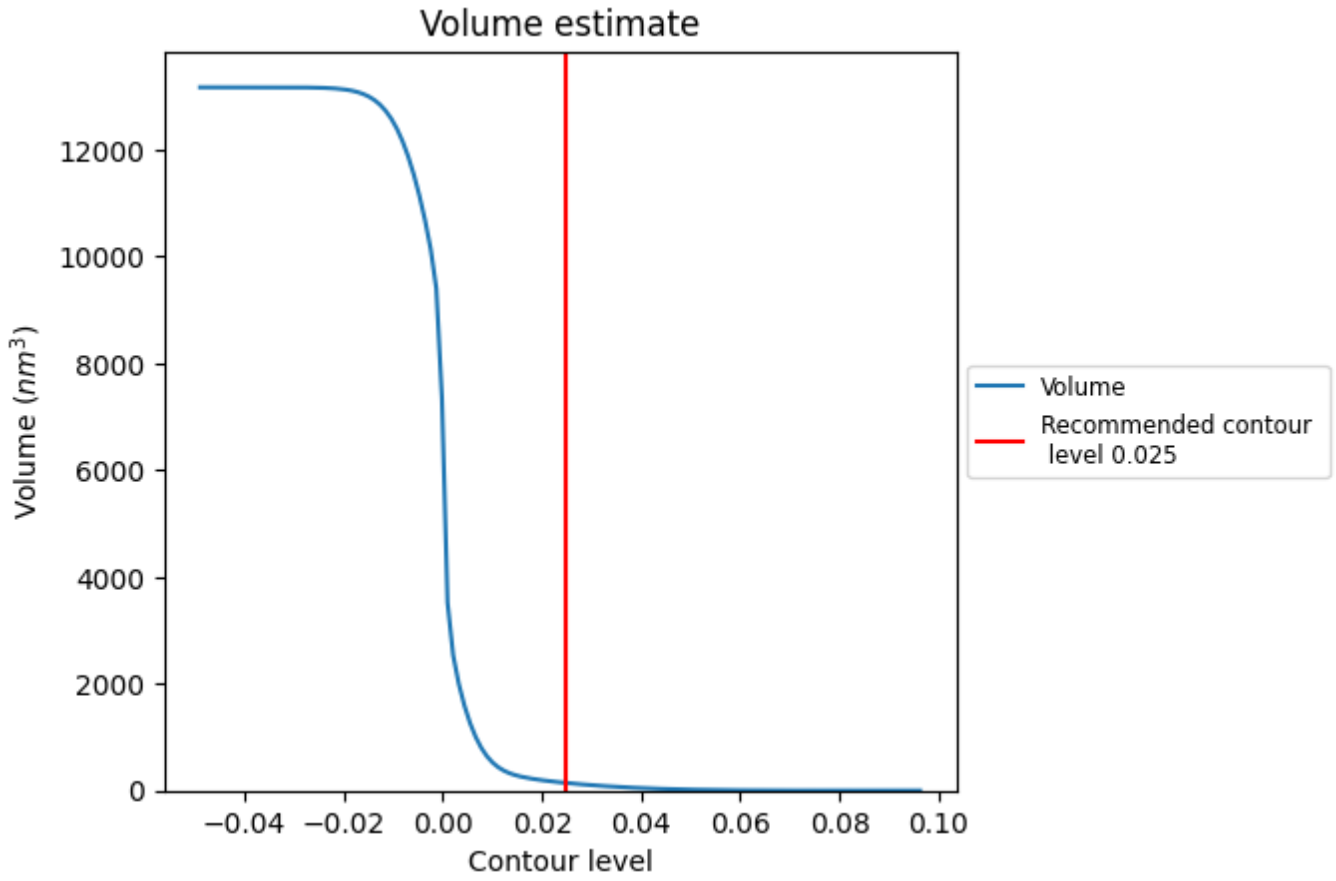
This section contains the results of statistical analysis of the map.

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.

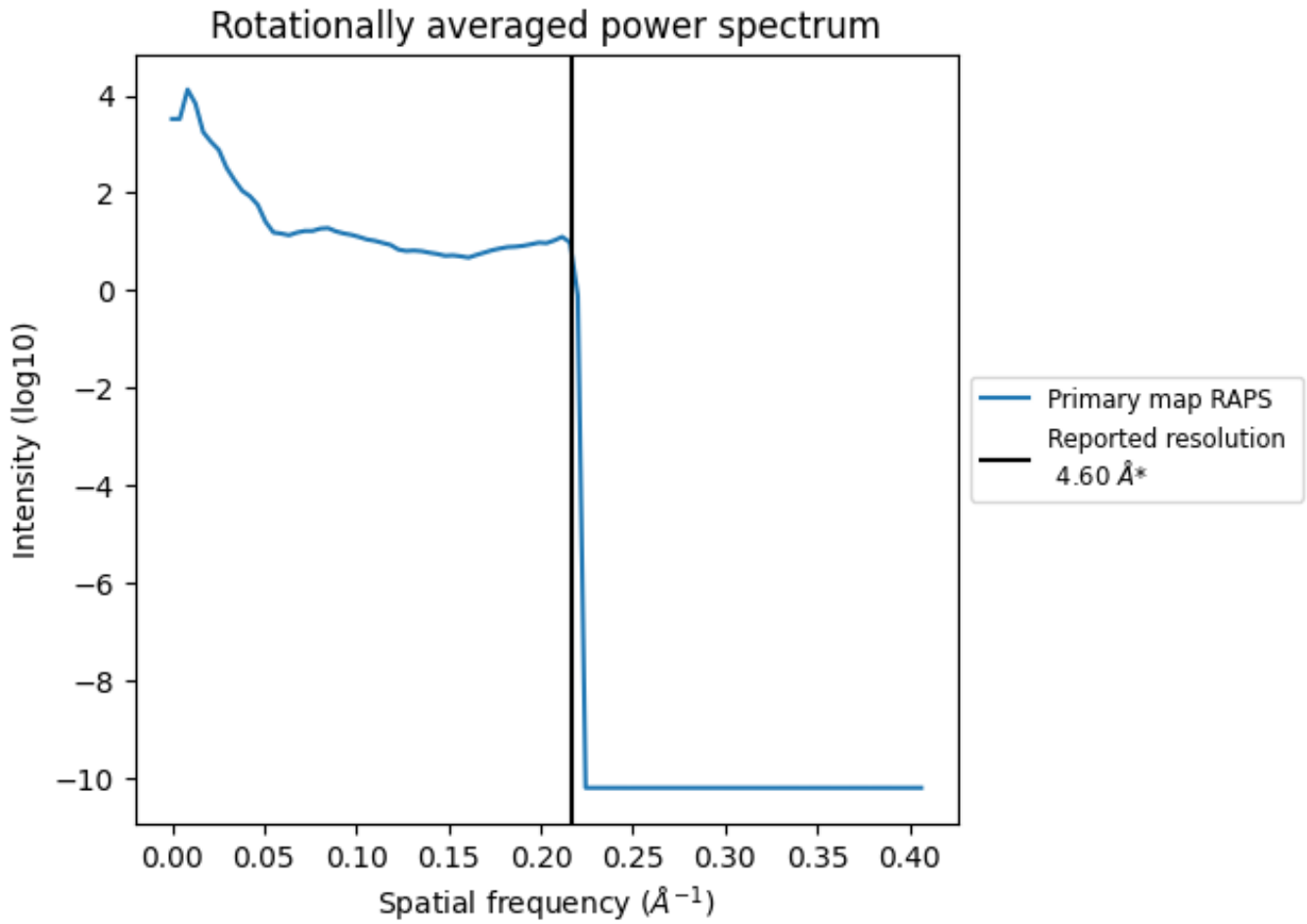
7.2 Volume estimate (i)



The volume at the recommended contour level is 142 nm^3; this corresponds to an approximate mass of 128 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)

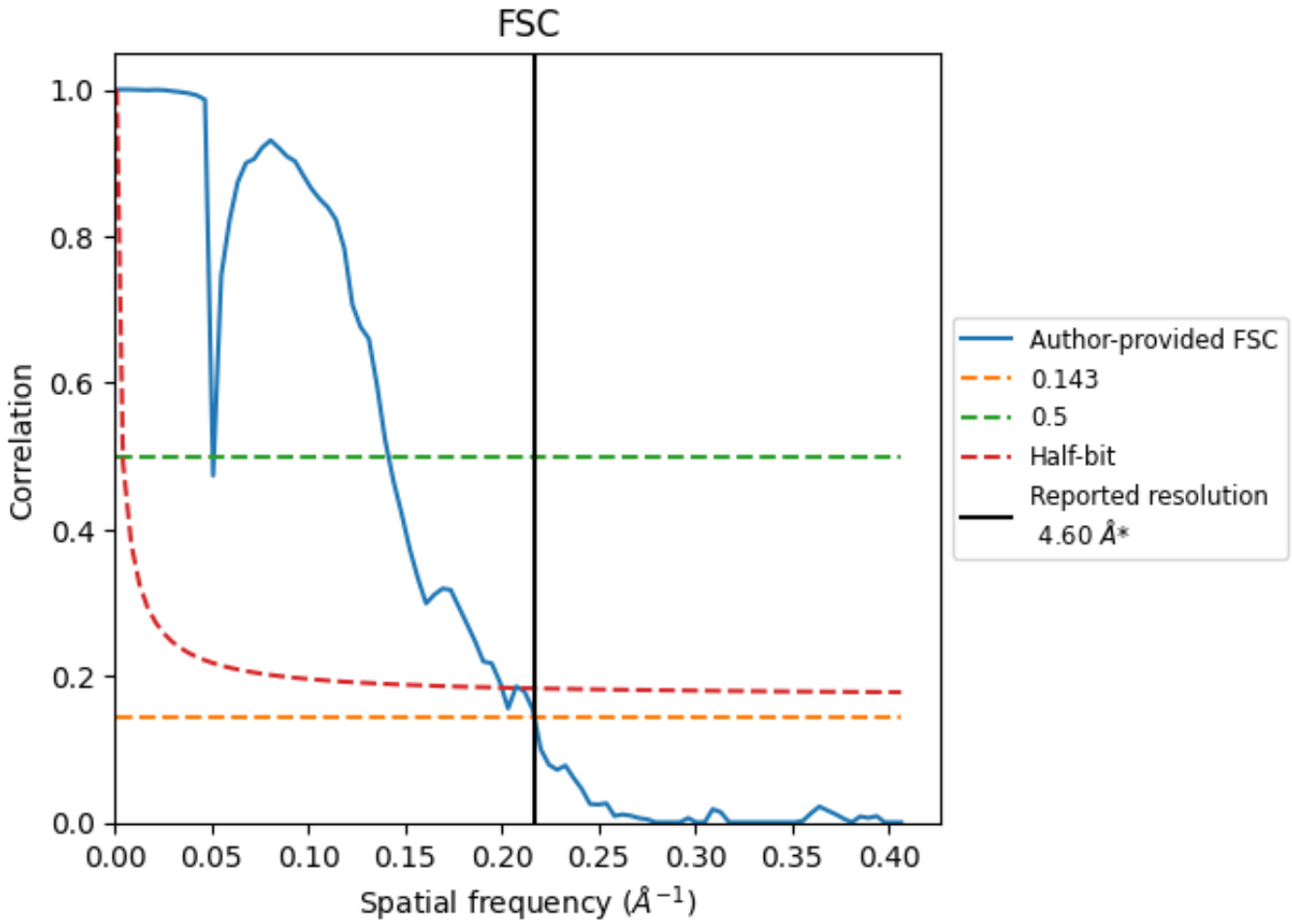


*Reported resolution corresponds to spatial frequency of 0.217 Å^{-1}

8 Fourier-Shell correlation

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

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.217 Å^{-1}

8.2 Resolution estimates (i)

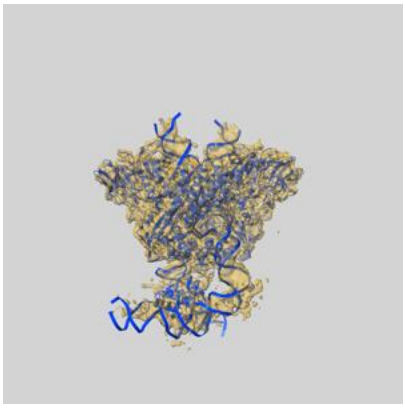
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.60	-	-
Author-provided FSC curve	4.61	19.76	5.00
Unmasked-calculated*	-	-	-

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

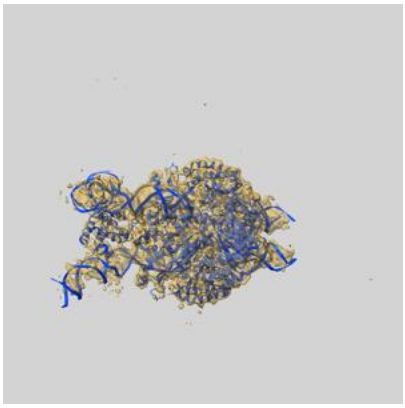
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-6489 and PDB model 3JBW. Per-residue inclusion information can be found in section 3 on page 6.

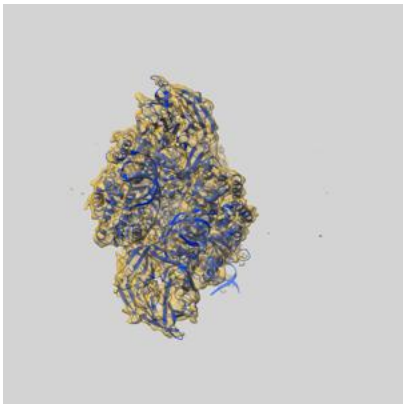
9.1 Map-model overlay



X



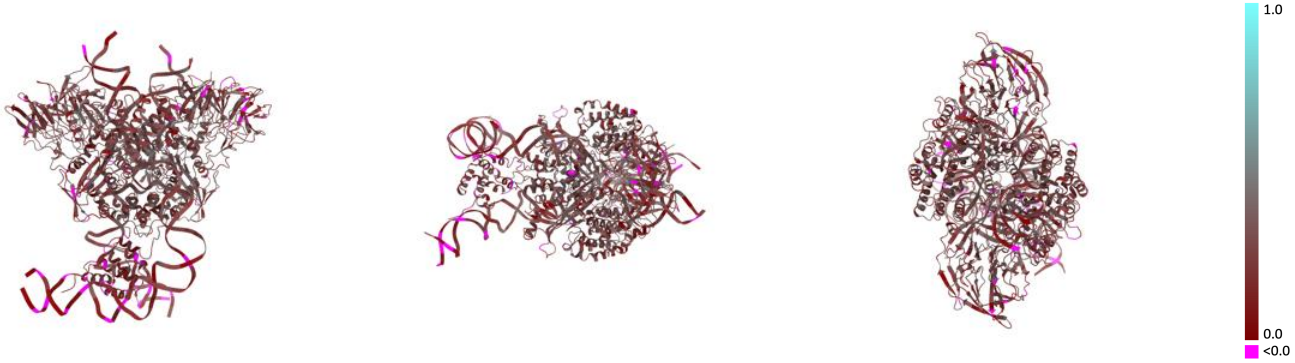
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.025 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.

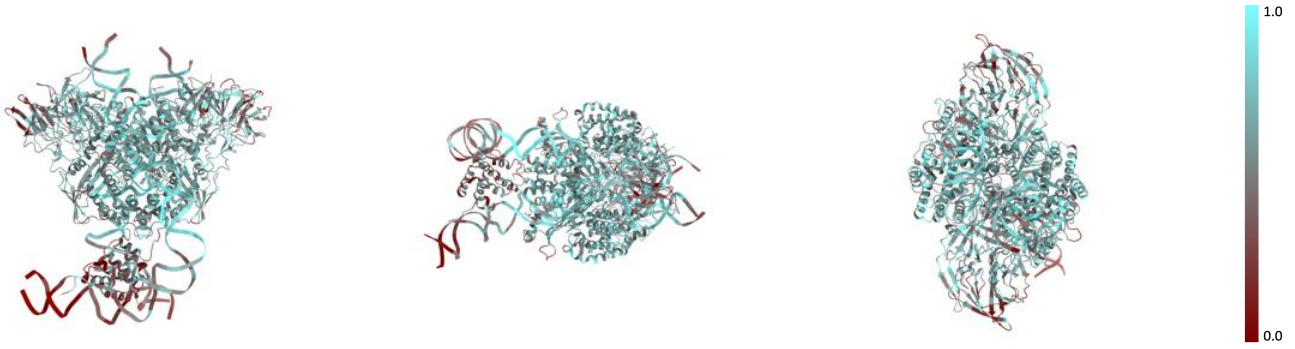
9.2 Q-score mapped to coordinate model



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.

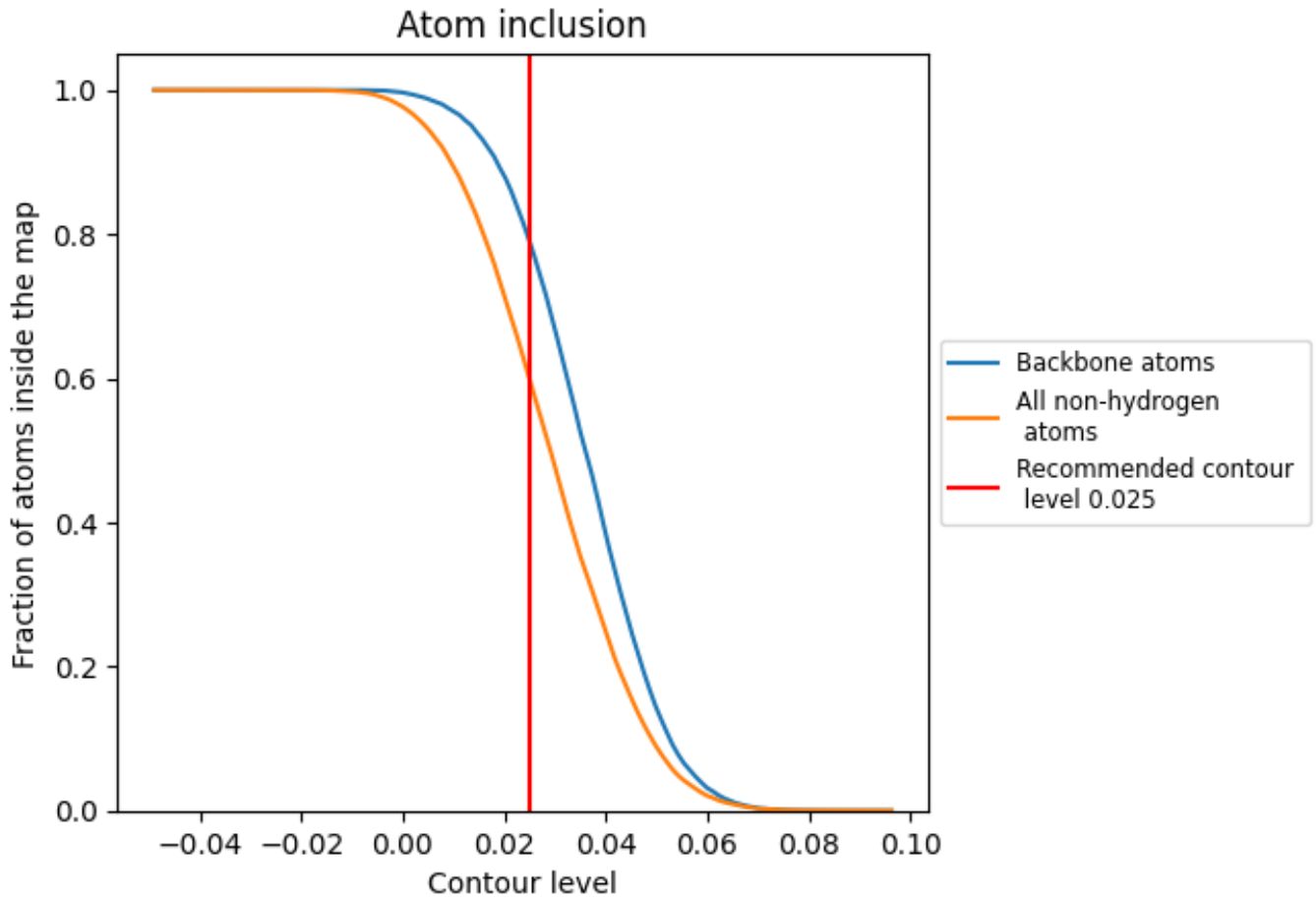
9.3 Atom inclusion mapped to coordinate model



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 (0.025).

9.4 Atom inclusion (i)



At the recommended contour level, 79% of all backbone atoms, 60% 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 (0.025) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5997	0.2450
A	0.6131	0.2600
B	0.5741	0.2600
C	0.6150	0.2610
D	0.5722	0.2500
E	0.5710	0.1780
F	0.6446	0.2320
G	0.5820	0.1920
H	0.5158	0.1720
I	0.7055	0.2410
J	0.7239	0.2150