



# wwPDB EM Validation Summary Report ⓘ

Nov 13, 2022 – 02:25 PM EST

PDB ID : 6XNZ  
EMDB ID : EMD-22274  
Title : Structure of RAG1 (R848M/E649V)-RAG2-DNA Target Capture Complex  
Authors : Zhang, Y.; Corbett, E.; Wu, S.; Schatz, D.G.  
Deposited on : 2020-07-05  
Resolution : 3.80 Å(reported)  
Based on initial model : 5ZDZ

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

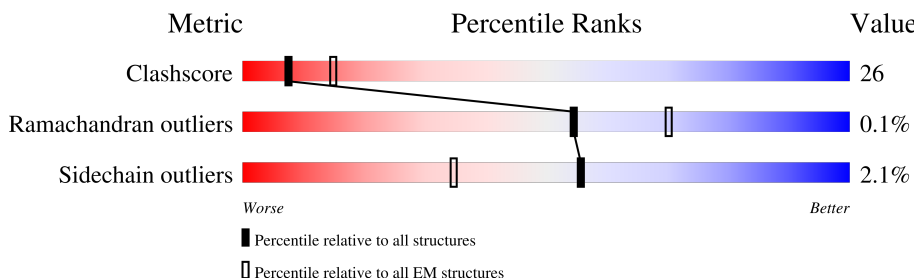
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.80 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	750	<div> <div>35%</div> <div>36%</div> <div>36%</div> <div>27%</div> </div>
1	C	750	<div> <div>32%</div> <div>40%</div> <div>32%</div> <div>27%</div> </div>
2	B	363	<div> <div>53%</div> <div>50%</div> <div>43%</div> <div>6%</div> </div>
2	D	363	<div> <div>58%</div> <div>50%</div> <div>43%</div> <div>6%</div> </div>
3	I	37	<div> <div>57%</div> <div>38%</div> <div>46%</div> <div>16%</div> </div>
4	J	37	<div> <div>65%</div> <div>38%</div> <div>46%</div> <div>16%</div> </div>
5	x	34	<div> <div>18%</div> <div>41%</div> <div>56%</div> </div>
6	M	34	<div> <div>18%</div> <div>29%</div> <div>15%</div> <div>56%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	y	45	<div><div><div></div><div></div><div></div></div><div>16%24%73%</div></div>
8	L	45	<div><div><div></div><div></div><div></div></div><div>11%7%20%73%</div></div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 16544 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	550	Total	C	N	O	S	0	0
			4411	2783	771	824	33		
1	C	549	Total	C	N	O	S	0	0
			4402	2779	769	821	33		

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	259	GLY	-	expression tag	UNP P15919
A	260	PRO	-	expression tag	UNP P15919
A	649	VAL	GLU	engineered mutation	UNP P15919
A	848	MET	ARG	engineered mutation	UNP P15919
C	259	GLY	-	expression tag	UNP P15919
C	260	PRO	-	expression tag	UNP P15919
C	649	VAL	GLU	engineered mutation	UNP P15919
C	848	MET	ARG	engineered mutation	UNP P15919

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	343	Total	C	N	O	S	0	0
			2680	1713	455	493	19		
2	D	343	Total	C	N	O	S	0	0
			2677	1711	454	493	19		

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP P21784
B	0	PRO	-	expression tag	UNP P21784
B	1	MET	-	expression tag	UNP P21784
B	2	ALA	-	expression tag	UNP P21784
D	-1	GLY	-	expression tag	UNP P21784

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	PRO	-	expression tag	UNP P21784
D	1	MET	-	expression tag	UNP P21784
D	2	ALA	-	expression tag	UNP P21784

- Molecule 3 is a DNA chain called Target DNA top strand (37-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
3	I	31	Total	C	N	O	P	0	0
			635	300	117	187	31		

- Molecule 4 is a DNA chain called Target DNA bottom strand (37-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	J	31	Total	C	N	O	P	0	0
			636	300	120	185	31		

- Molecule 5 is a DNA chain called 12RSS integration strand (34-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	x	15	Total	C	N	O	P	0	0
			301	144	51	91	15		

- Molecule 6 is a DNA chain called 12RSS non-integration strand (34-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	M	15	Total	C	N	O	P	0	0
			311	148	62	87	14		

- Molecule 7 is a DNA chain called 23RSS integration strand (45-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	y	12	Total	C	N	O	P	0	0
			243	116	43	72	12		

- Molecule 8 is a DNA chain called 23RSS non-integration strand (45-mer).

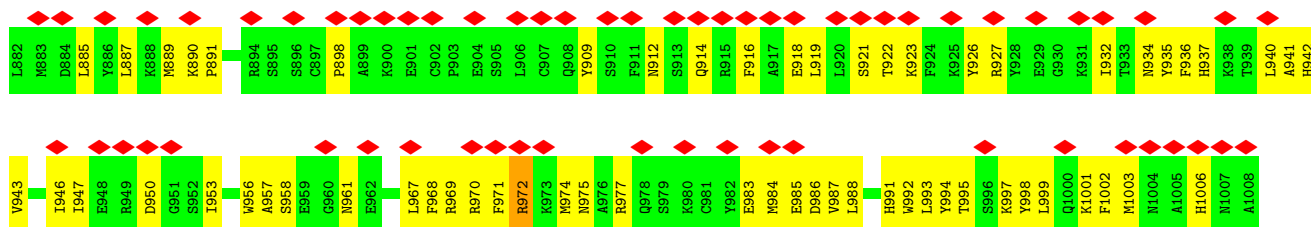
Mol	Chain	Residues	Atoms					AltConf	Trace
8	L	12	Total	C	N	O	P	0	0
			246	118	47	70	11		

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Inter-

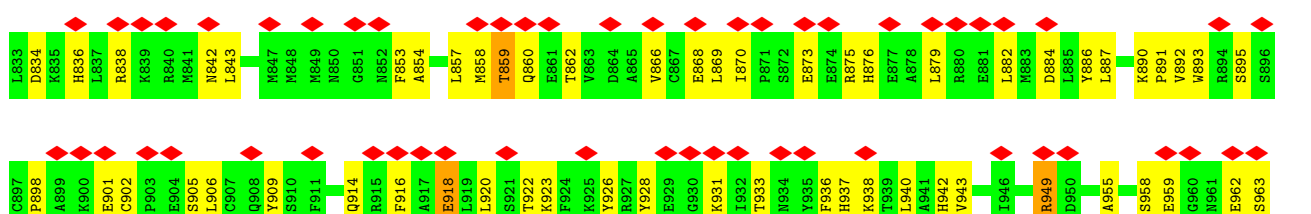
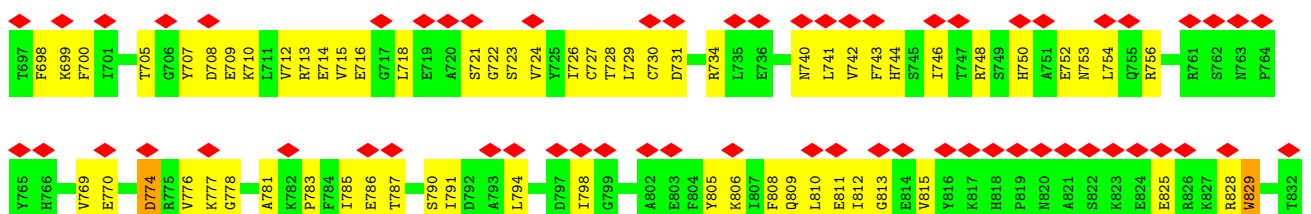
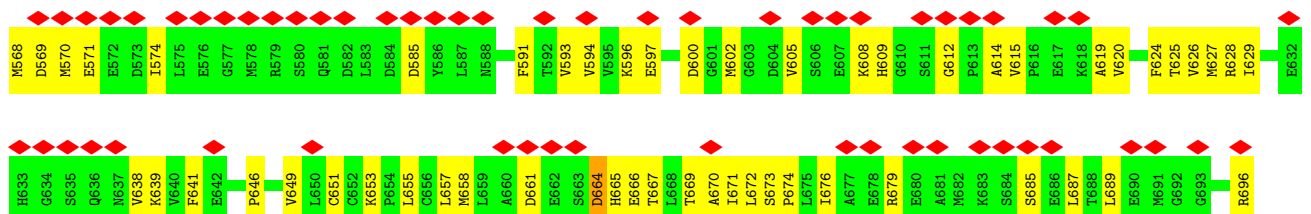
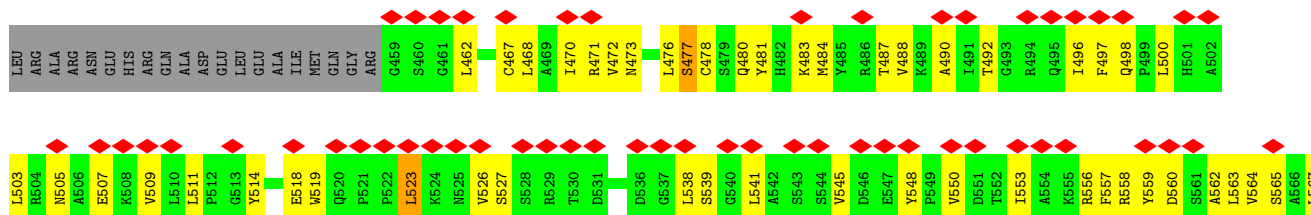
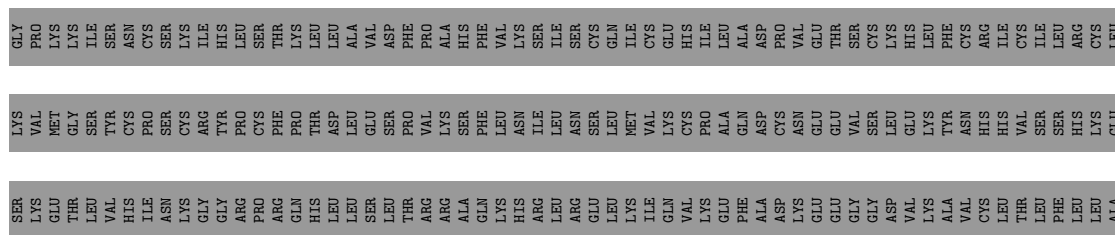
est" by depositor).

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

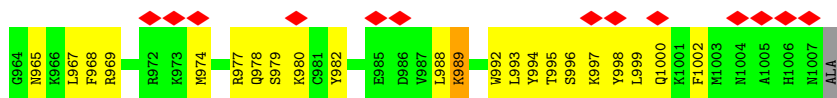




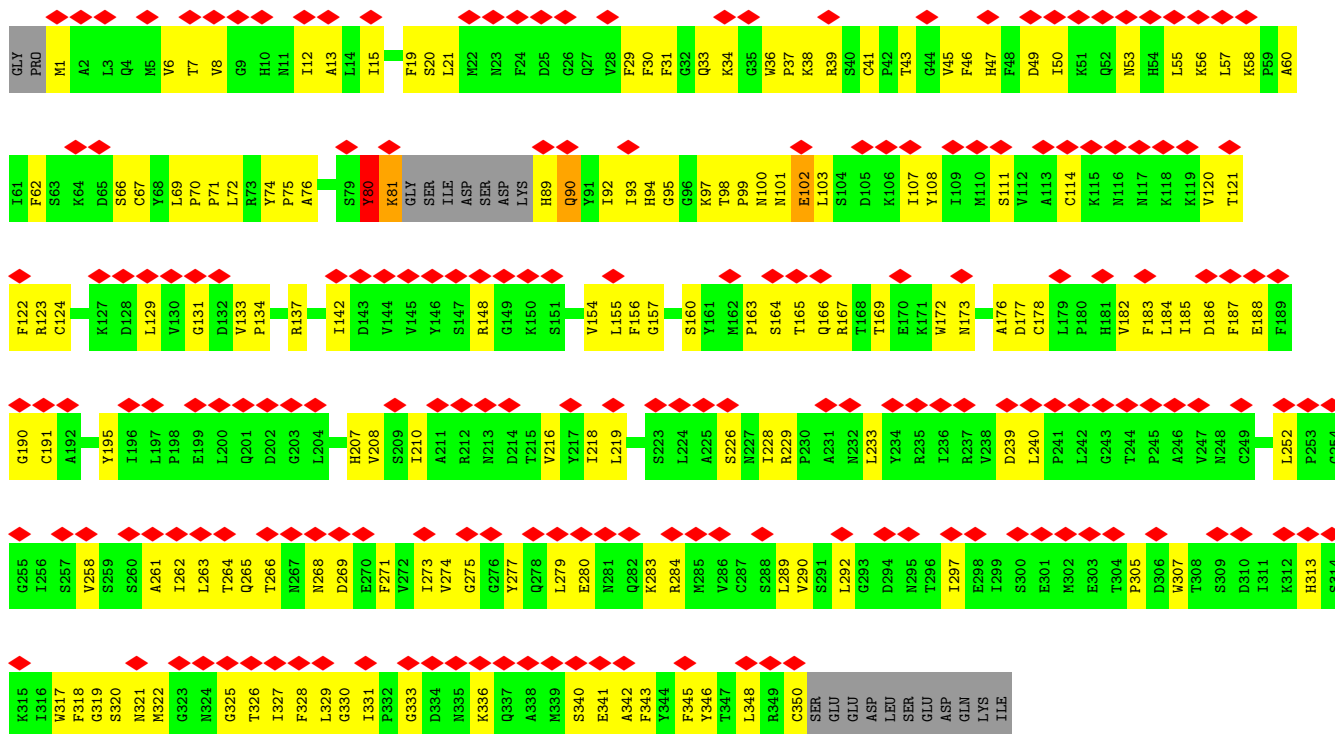
• Molecule 1: V(D)J recombination-activating protein 1



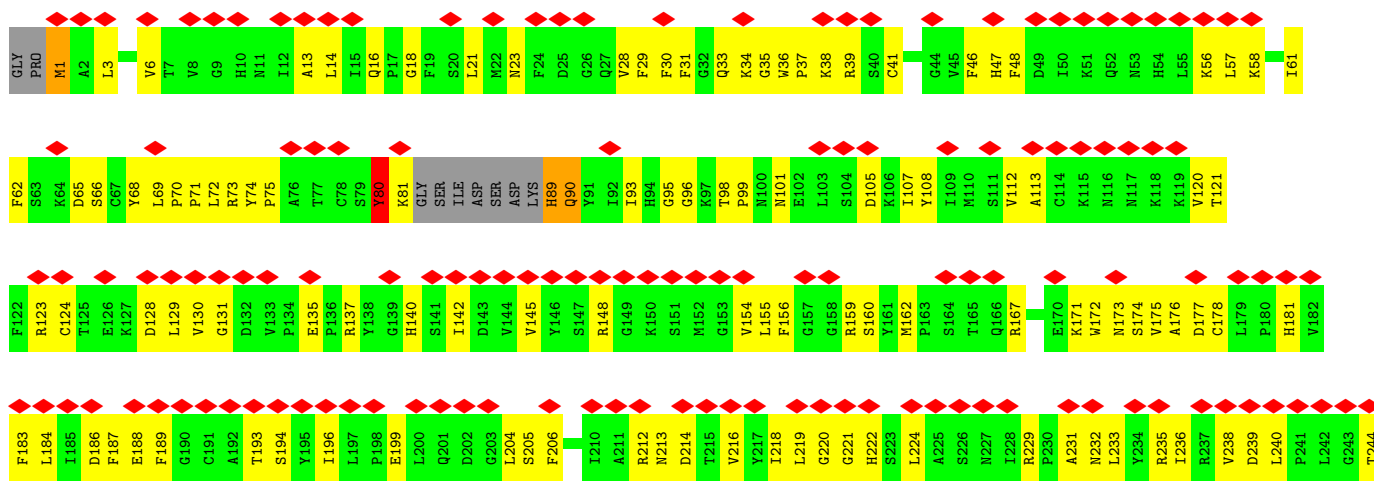




• Molecule 2: V(D)J recombination-activating protein 2



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







## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	22270	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 ( $e^-/\text{\AA}^2$ )	72.8	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2300	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.101	Depositor
Minimum map value	-0.059	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.03	Depositor
Map size (Å)	256.19998, 256.19998, 256.19998	wwPDB
Map dimensions	244, 244, 244	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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	0.35	0/4508	0.49	0/6079
1	C	0.33	0/4499	0.48	0/6068
2	B	0.36	0/2748	0.60	4/3725 (0.1%)
2	D	0.34	0/2744	0.52	1/3719 (0.0%)
3	I	0.61	0/711	0.88	0/1095
4	J	0.60	0/713	0.91	0/1098
5	x	0.67	0/335	1.04	1/513 (0.2%)
6	M	0.48	0/350	0.90	0/540
7	y	0.66	0/271	0.98	1/415 (0.2%)
8	L	0.66	0/276	0.99	0/425
All	All	0.40	0/17155	0.61	7/23677 (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
2	B	0	1
2	D	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	80	TYR	CB-CA-C	-13.18	84.04	110.40
2	B	80	TYR	CA-CB-CG	8.42	129.40	113.40
5	x	45	DG	C1'-O4'-C4'	-7.95	102.15	110.10
2	B	80	TYR	N-CA-CB	-5.66	100.41	110.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	y	35	DG	O5'-P-OP1	-5.17	101.05	105.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	80	TYR	Mainchain
2	D	80	TYR	Mainchain

## 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	4411	0	4366	271	0
1	C	4402	0	4357	235	0
2	B	2680	0	2638	130	0
2	D	2677	0	2629	157	0
3	I	635	0	348	21	0
4	J	636	0	347	19	0
5	x	301	0	170	0	0
6	M	311	0	170	5	0
7	y	243	0	136	0	0
8	L	246	0	137	18	0
9	A	1	0	0	0	0
9	C	1	0	0	0	0
All	All	16544	0	15298	791	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 26.

The worst 5 of 791 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)
2:B:1:MET:CE	2:B:305:PRO:HB3	1.70	1.20
2:B:1:MET:HE2	2:B:305:PRO:HB3	1.22	1.13

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:283:LYS:NZ	2:D:315:LYS:H	1.48	1.08
2:D:283:LYS:HZ3	2:D:315:LYS:H	1.06	1.01
1:A:1003:MET:SD	1:C:487:THR:OG1	2.22	0.95

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	548/750 (73%)	471 (86%)	76 (14%)	1 (0%)	47	79
1	C	547/750 (73%)	478 (87%)	68 (12%)	1 (0%)	47	79
2	B	339/363 (93%)	280 (83%)	59 (17%)	0	100	100
2	D	337/363 (93%)	286 (85%)	51 (15%)	0	100	100
All	All	1771/2226 (80%)	1515 (86%)	254 (14%)	2 (0%)	54	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	638	VAL
1	C	829	TRP

### 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	486/668 (73%)	473 (97%)	13 (3%)	44	69
1	C	485/668 (73%)	475 (98%)	10 (2%)	53	74
2	B	298/318 (94%)	292 (98%)	6 (2%)	55	75
2	D	297/318 (93%)	293 (99%)	4 (1%)	69	82
All	All	1566/1972 (79%)	1533 (98%)	33 (2%)	56	74

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	989	LYS
2	D	1	MET
2	D	283	LYS
1	A	972	ARG
1	A	872	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	173	ASN
2	D	90	GLN
2	B	101	ASN
2	D	27	GLN
2	B	16	GLN

### 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

There are no chain breaks in this entry.

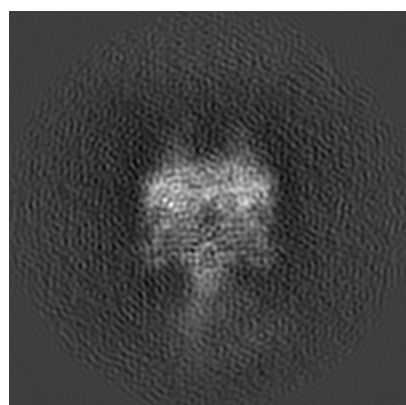
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22274. These allow visual inspection of the internal detail of the map and identification of artifacts.

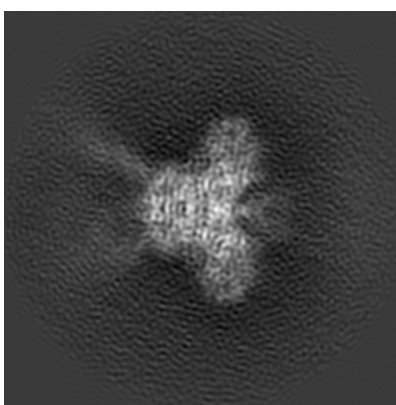
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

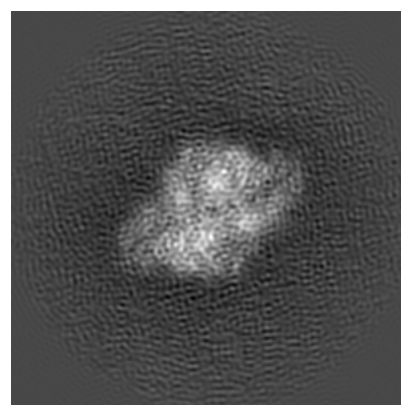
#### 6.1.1 Primary map



X



Y

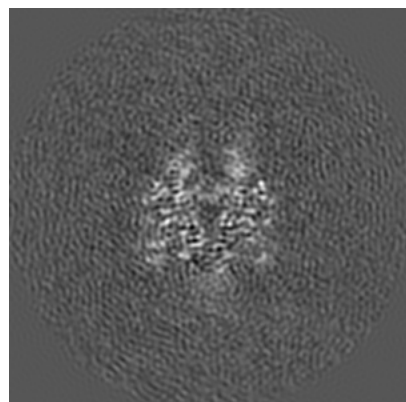


Z

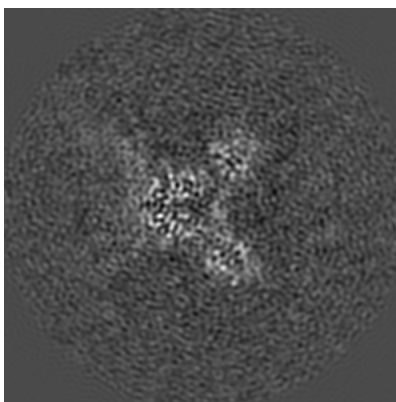
The images above show the map projected in three orthogonal directions.

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

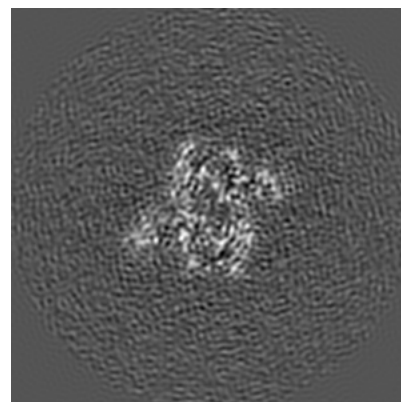
#### 6.2.1 Primary map



X Index: 122



Y Index: 122

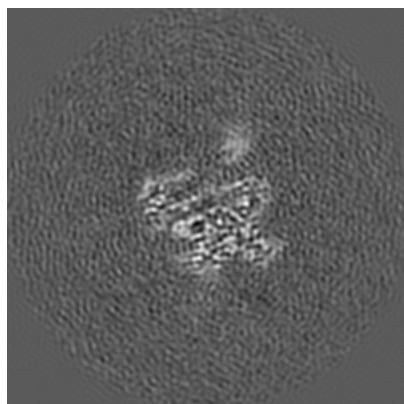


Z Index: 122

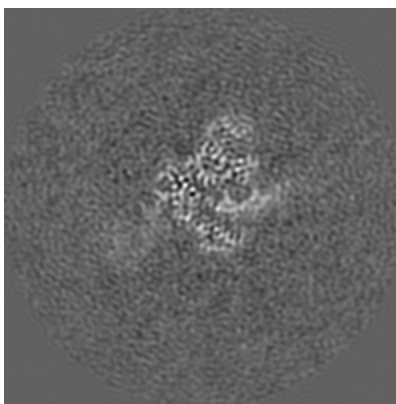
The images above show central slices of the map in three orthogonal directions.

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

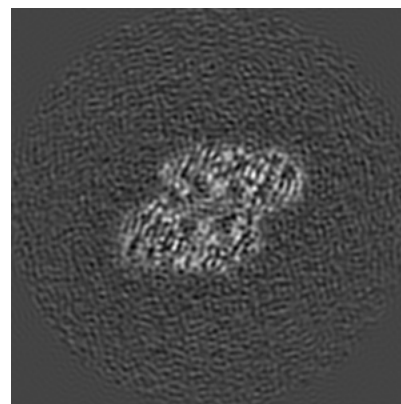
### 6.3.1 Primary map



X Index: 129



Y Index: 138



Z Index: 134

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

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

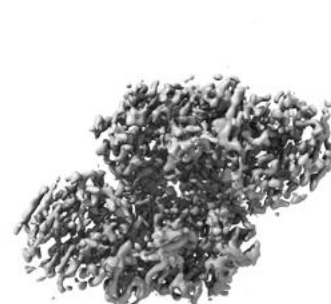
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.03. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

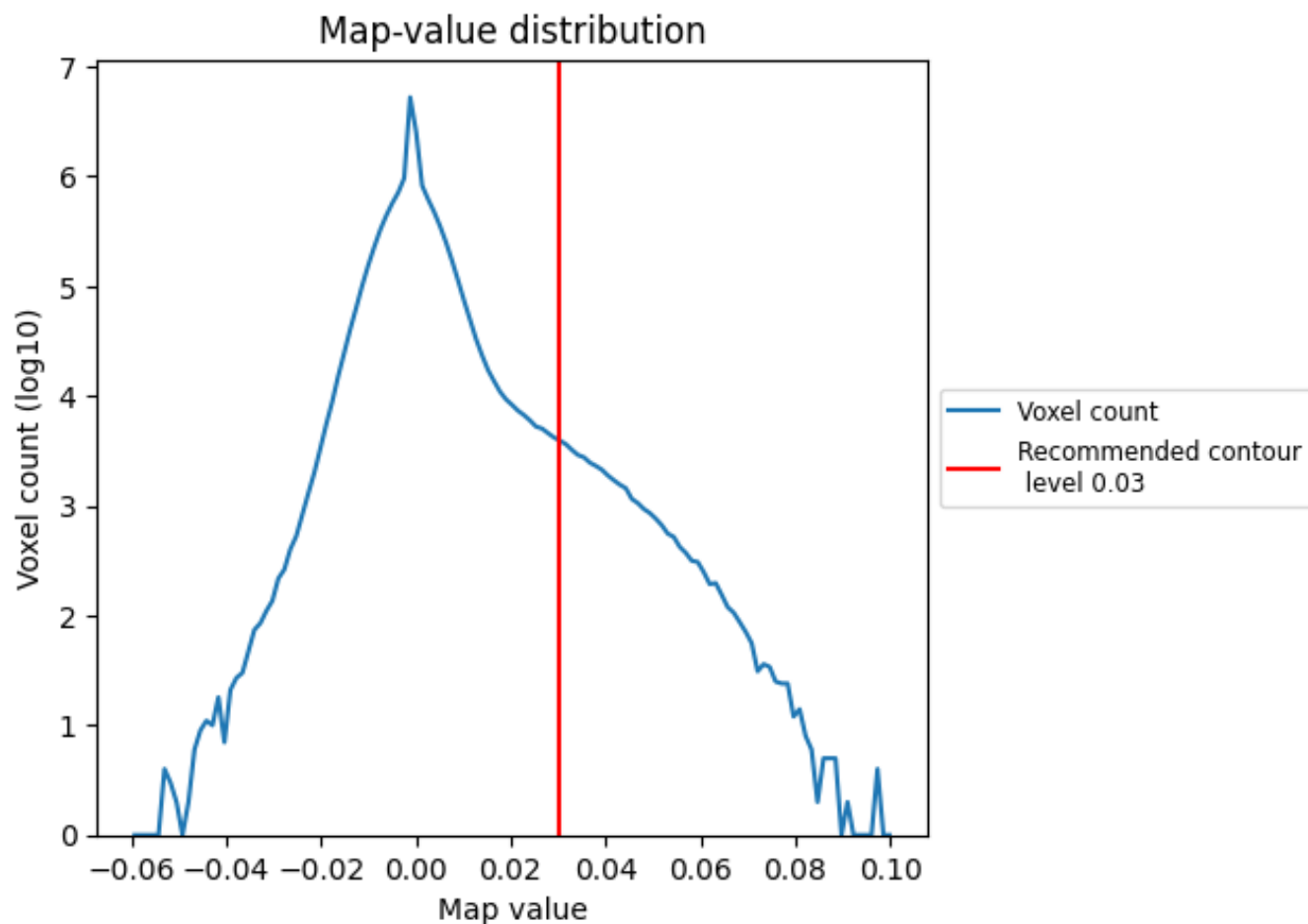
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

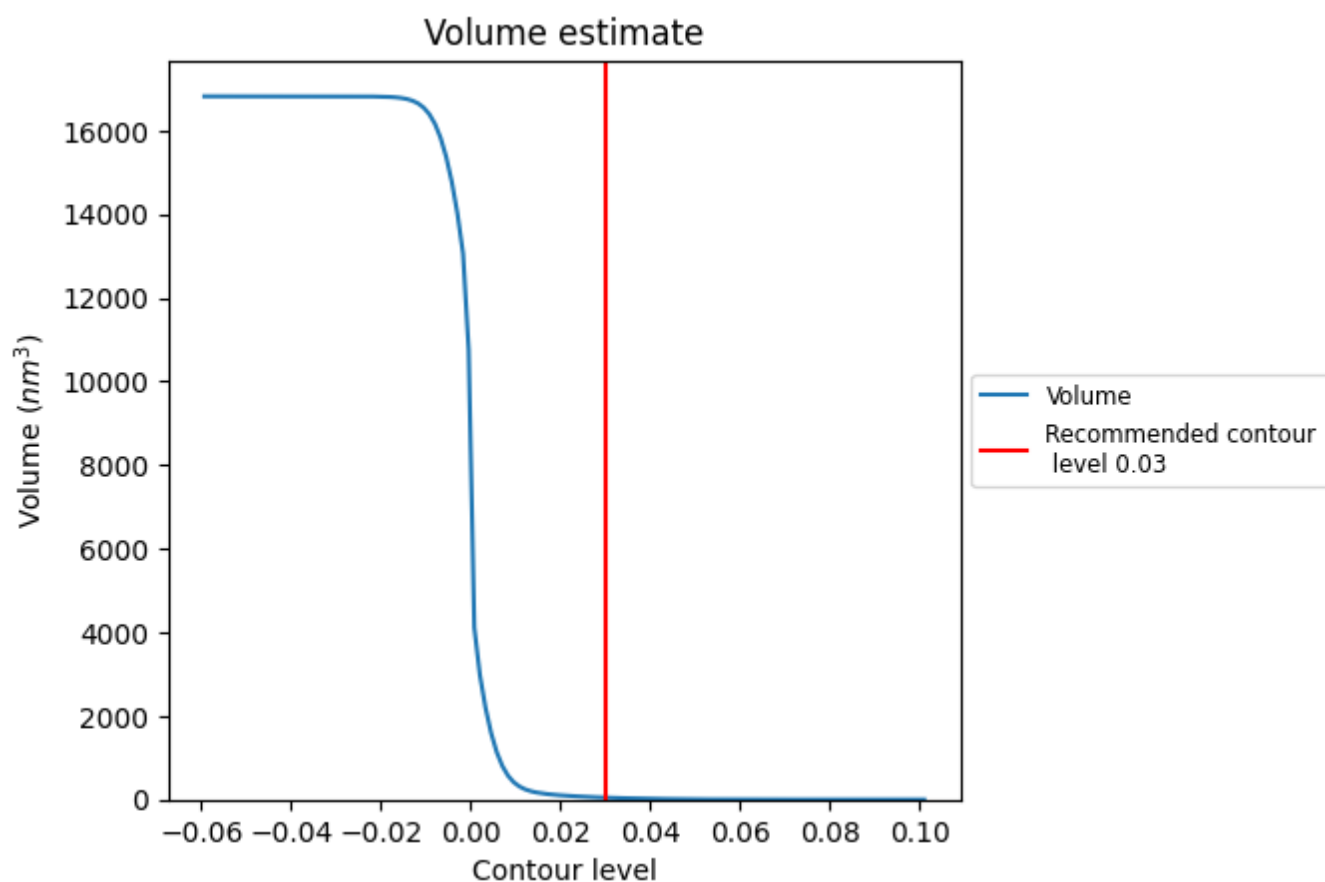
This section contains the results of statistical analysis of the map.

### 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.

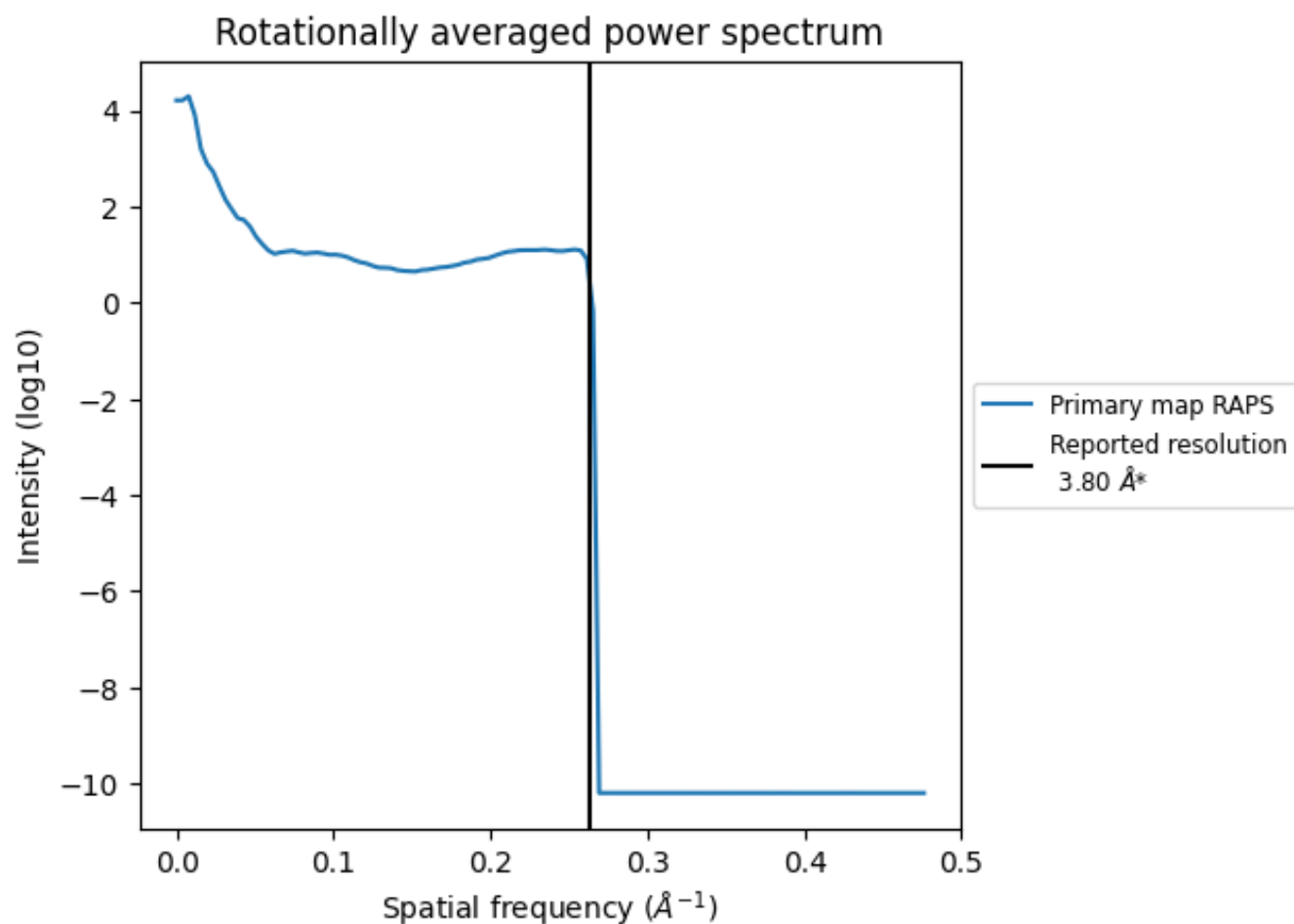
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 47  $\text{nm}^3$ ; this corresponds to an approximate mass of 42 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 0.263 Å<sup>-1</sup>

## 8 Fourier-Shell correlation

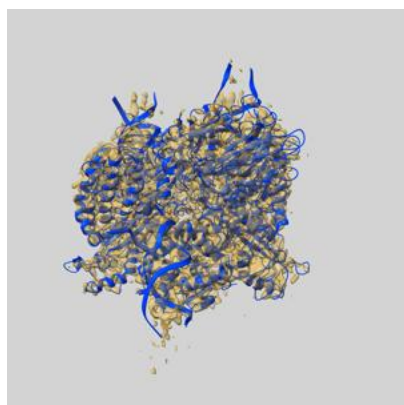
This section was not generated. No FSC curve or half-maps provided.



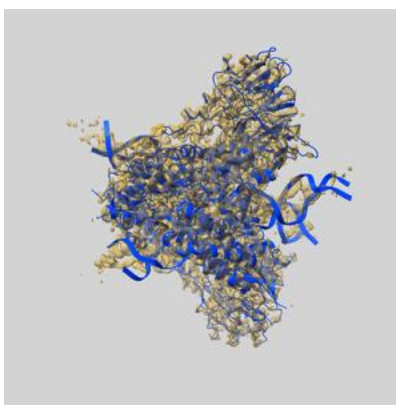
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-22274 and PDB model 6XNZ. Per-residue inclusion information can be found in section [3](#) on page [7](#).

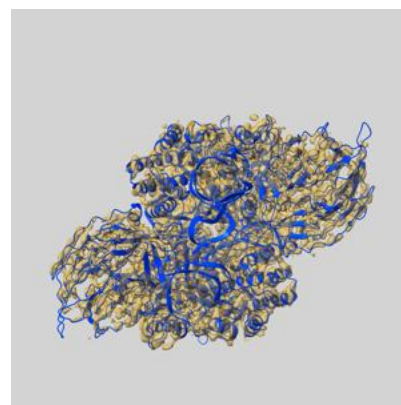
### 9.1 Map-model overlay [i](#)



X



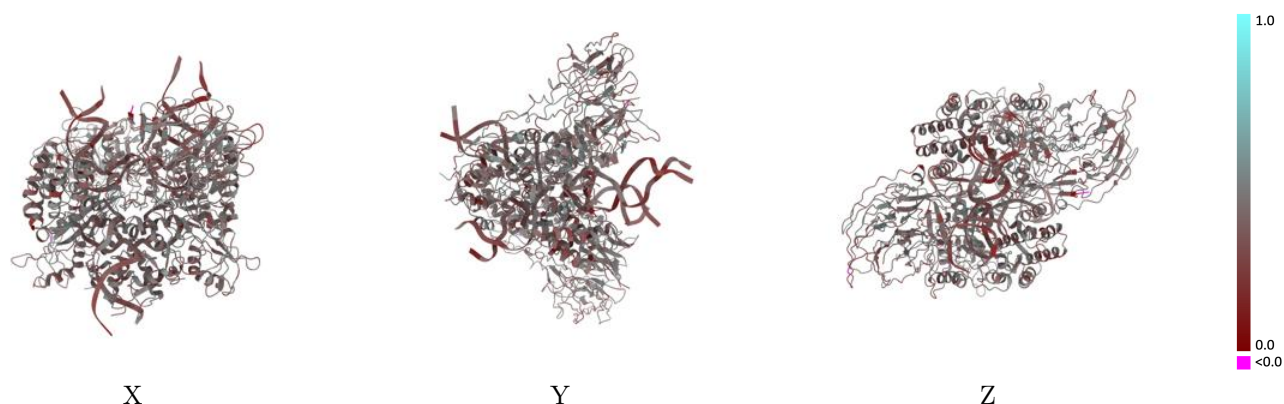
Y



Z

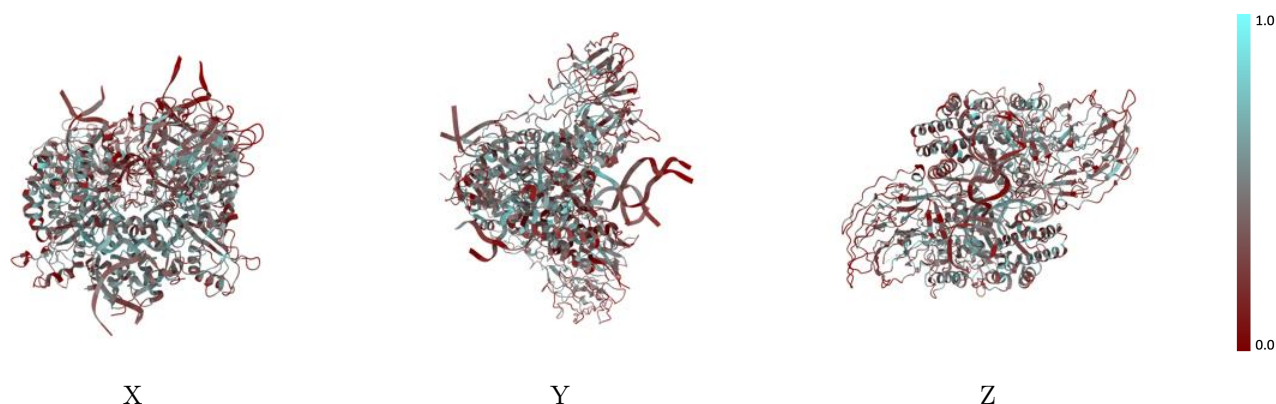
The images above show the 3D surface view of the map at the recommended contour level 0.03 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



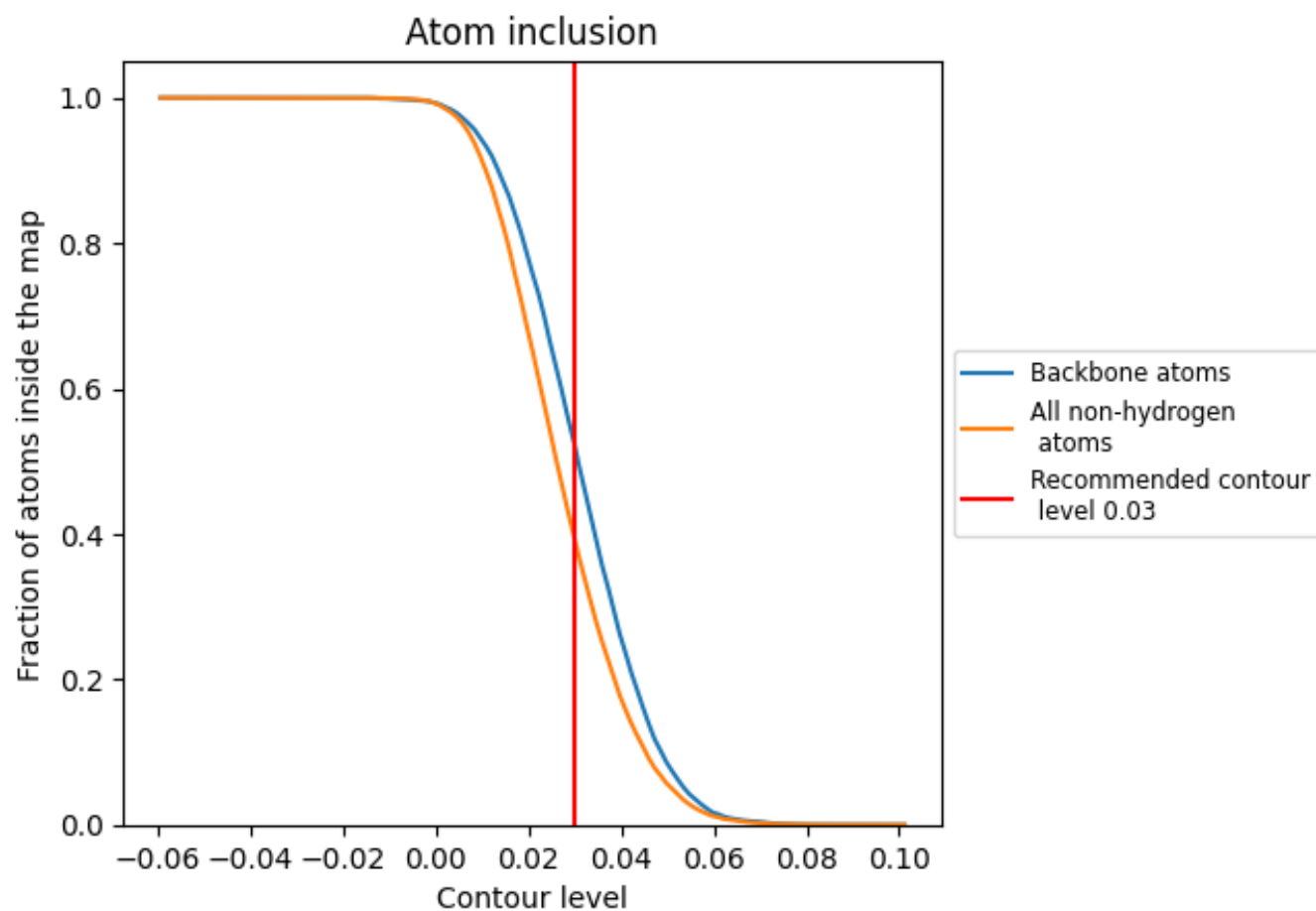
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.03).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 52% of all backbone atoms, 39% 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 (0.03) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.3938	<div></div> 0.3930
A	<div></div> 0.4196	<div></div> 0.4060
B	<div></div> 0.3615	<div></div> 0.4080
C	<div></div> 0.4381	<div></div> 0.4100
D	<div></div> 0.3376	<div></div> 0.4010
I	<div></div> 0.3055	<div></div> 0.3060
J	<div></div> 0.2752	<div></div> 0.2700
L	<div></div> 0.4675	<div></div> 0.3140
M	<div></div> 0.4887	<div></div> 0.3210
x	<div></div> 0.4718	<div></div> 0.3740
y	<div></div> 0.3539	<div></div> 0.3220

