



## wwPDB EM Validation Summary Report ⓘ

Sep 28, 2024 – 10:23 pm BST

PDB ID : 7NIK  
EMDB ID : EMD-12361  
Title : 1918 H1N1 Viral influenza polymerase heterotrimer with Nb8189 core  
Authors : Keown, J.R.; Carrique, L.; Fodor, E.; Grimes, J.M.  
Deposited on : 2021-02-12  
Resolution : 6.20 Å(reported)  
Based on initial model : 7HNA

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.dev113  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

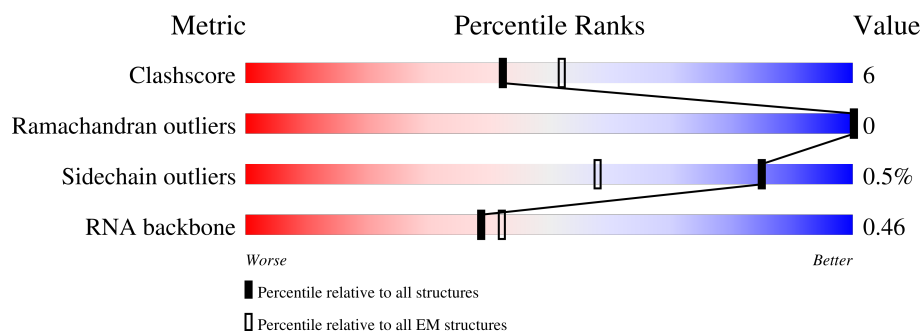
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 6.20 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	716	
2	B	757	
3	C	905	
4	E	17	
5	D	15	
6	F	128	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 25547 atoms, of which 12253 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	516	Total	C	H	N	O	S	0	0
			8256	2627	4111	701	788	29		

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	735	Total	C	H	N	O	S	0	0
			11704	3691	5833	1028	1106	46		

- Molecule 3 is a protein called Polymerase basic protein 2,Immunoglobulin G-binding protein A.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	C	249	Total	C	H	N	O	S	0	0
			4075	1271	2058	361	369	16		

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	760	GLU	-	linker	UNP Q3HM41
C	761	ASN	-	linker	UNP Q3HM41
C	762	LEU	-	linker	UNP Q3HM41
C	763	TYR	-	linker	UNP Q3HM41
C	764	PHE	-	linker	UNP Q3HM41
C	765	GLN	-	linker	UNP Q3HM41
C	766	GLY	-	linker	UNP Q3HM41
C	767	GLU	-	linker	UNP Q3HM41
C	768	LEU	-	linker	UNP Q3HM41
C	769	LYS	-	linker	UNP Q3HM41
C	770	THR	-	linker	UNP Q3HM41
C	771	ALA	-	linker	UNP Q3HM41
C	772	ALA	-	linker	UNP Q3HM41
C	773	LEU	-	linker	UNP Q3HM41

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	774	ALA	-	linker	UNP Q3HM41
C	775	GLN	-	linker	UNP Q3HM41
C	776	HIS	-	linker	UNP Q3HM41
C	777	ASP	-	linker	UNP Q3HM41
C	778	GLU	-	linker	UNP Q3HM41
C	779	ALA	-	linker	UNP Q3HM41
C	780	VAL	-	linker	UNP Q3HM41
C	781	ASP	-	linker	UNP Q3HM41
C	782	ASN	-	linker	UNP Q3HM41
C	783	LYS	-	linker	UNP Q3HM41
C	797	HIS	ASN	conflict	UNP P38507
C	798	LEU	MET	conflict	UNP P38507
C	808	ALA	GLY	conflict	UNP P38507
C	832	ASP	GLU	conflict	UNP P38507
C	833	ALA	SER	conflict	UNP P38507
C	838	VAL	ALA	conflict	UNP P38507
C	866	ALA	GLY	conflict	UNP P38507
C	873	ALA	ASP	conflict	UNP P38507
C	890	GLY	ASP	conflict	UNP P38507
C	896	VAL	ALA	conflict	UNP P38507
C	898	ALA	-	expression tag	UNP P38507
C	899	ASN	-	expression tag	UNP P38507
C	900	SER	-	expression tag	UNP P38507
C	901	ALA	-	expression tag	UNP P38507
C	902	GLY	-	expression tag	UNP P38507
C	903	LYS	-	expression tag	UNP P38507
C	904	SER	-	expression tag	UNP P38507
C	905	THR	-	expression tag	UNP P38507

- Molecule 4 is a RNA chain called RNA (5'-R(P\*GP\*GP\*CP\*CP\*UP\*GP\*CP\*U)-3').

Mol	Chain	Residues	Atoms						AltConf	Trace
4	E	8	Total	C	H	N	O	P	0	0
			255	75	86	28	58	8		

- Molecule 5 is a RNA chain called RNA (5'-R(P\*AP\*GP\*UP\*AP\*GP\*AP\*AP\*AP\*CP\*AP\*AP\*GP\*GP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace	
5	D	15	Total	C	H	N	O	P	0	0
			491	146	165	66	99	15		

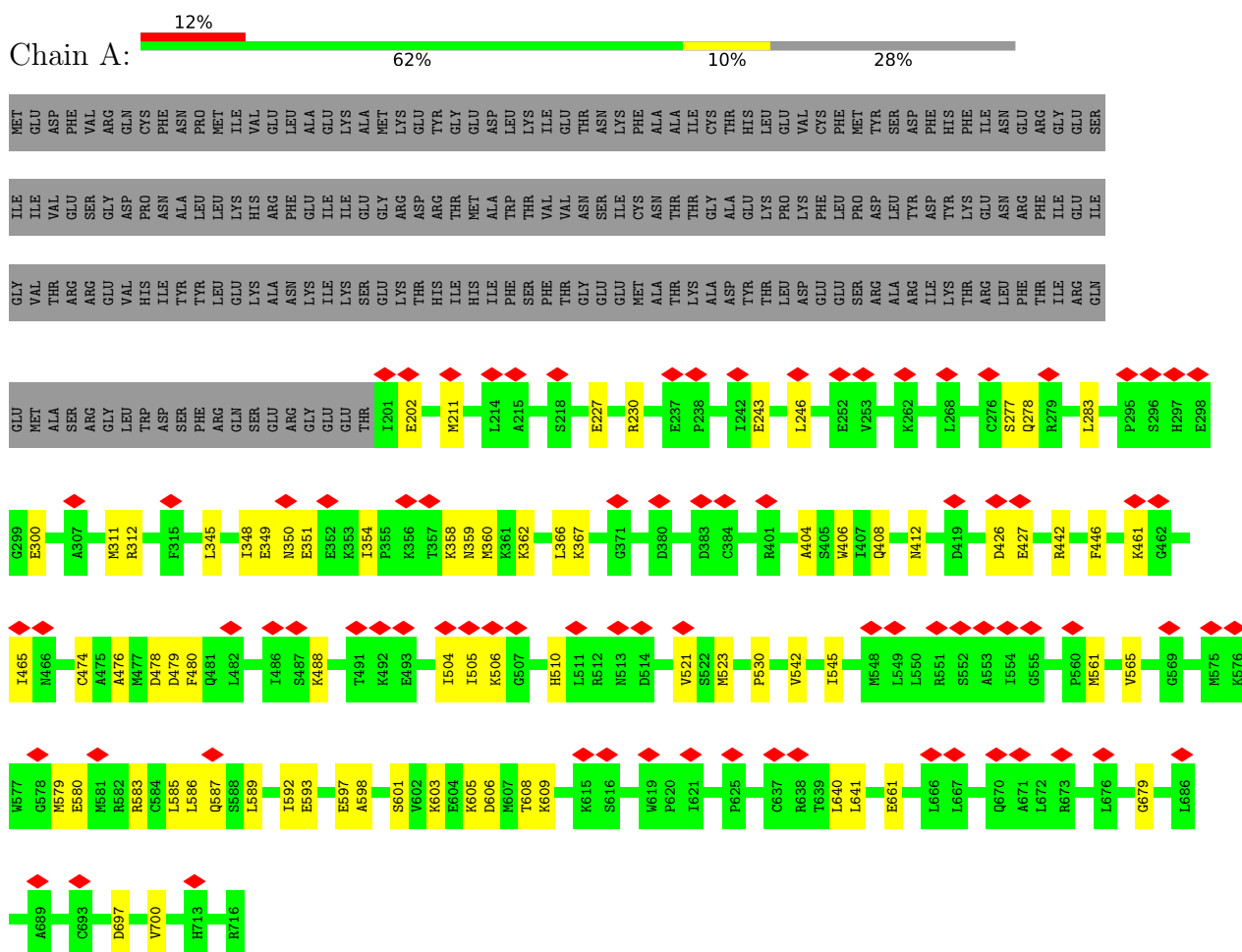
- Molecule 6 is a protein called Nanobody8189 core.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	99	766	476	139	147	4	0	0

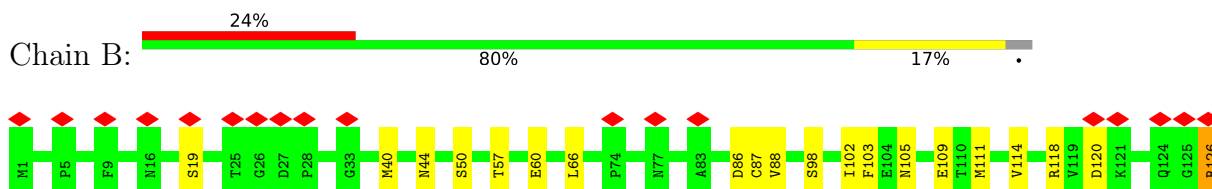
### 3 Residue-property plots

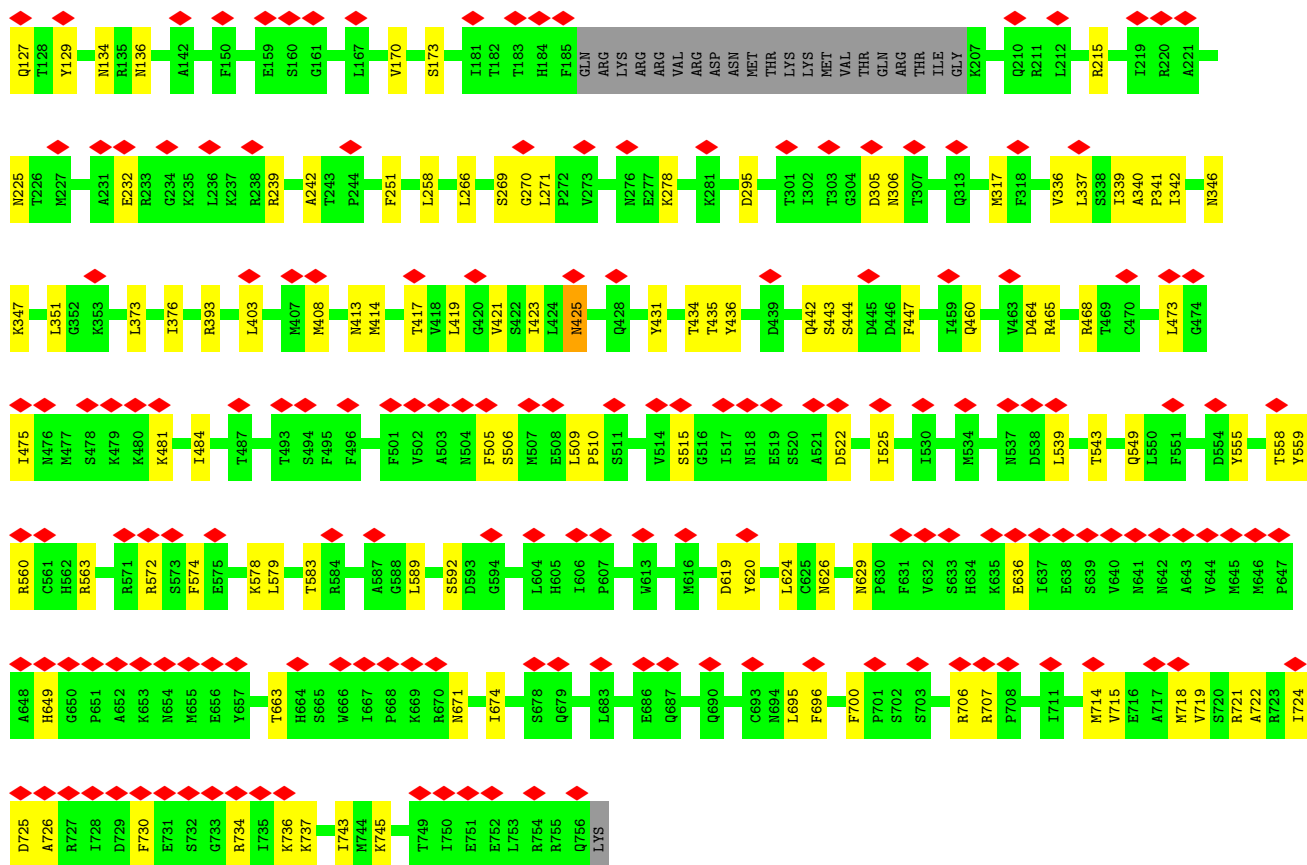
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Polymerase acidic protein

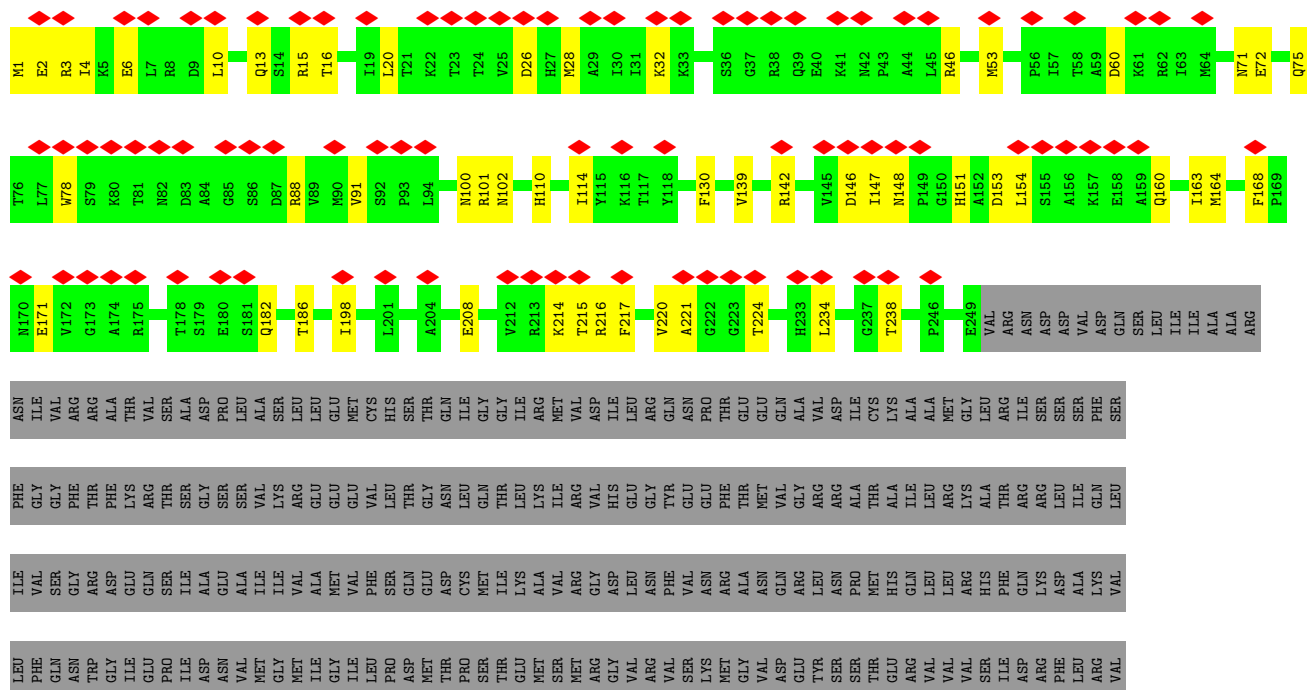


#### • Molecule 2: RNA-directed RNA polymerase catalytic subunit



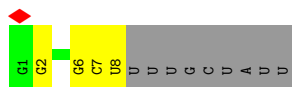


• Molecule 3: Polymerase basic protein 2,Immunoglobulin G-binding protein A

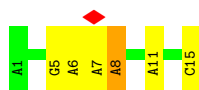


ASN	ASN	GLN	THR	GLY	PRO	SER	ARG
ALA	PHE	ARG	ASP	VAL	PRO	GLN	ASP
PHE	ASN	ASN	THR	GLU	GLN	ASN	GLN
ILE	ALA	ALA	GLN	SER	SER	THR	GLY
GLN	PHE	THR	ALA	VAL	ARG	MET	ASN
SER	ILE	ALA	THR	LEU	MET	LEU	VAL
LEU	GLN	GLN	THR	ARG	GLY	TYR	LEU
LYS	SER	LYS	LYS	GLN	LYS	TYR	LEU
ALA	ALA	LEU	ARG	GLY	PHE	ASN	SER
ASP	ASP	LEU	ILE	PHE	ASN	LYS	SER
PRO	ASP	ASP	MET	ILE	SER	MET	PRO
SER	ASN	ASP	ALA	LEU	LEU	GLU	GLU
GLN	PRO	PRO	ALA	LEU	THR	PHE	GLU
SER	ALA	ASN	ILE	GLY	VAL	GLY	VAL
ALA	GLN	GLN	ASN	LYS	ASN	PRO	SER
ASN	ASN	SER	GLU	GLU	VAL	PHE	GLU
LEU	LEU	ALA	ASN	ASP	ARG	GLN	THR
ALA	LEU	ASN	LEU	ARG	GLY	SER	GLN
ALA	LEU	LEU	TYR	ARG	SER	LEU	GLY
GLU	GLU	PHE	PHE	TYR	VAL	VAL	THR
ALA	ALA	GLN	GLN	GLY	MET	PRO	GLU
LYS	LYS	GLY	GLY	PRO	ARG	LYS	LYS
LEU	LYS	ALA	GLU	ALA	ILE	ALA	THR
LYS	LEU	LYS	LEU	LEU	LEU	ALA	ALA
ASN	ASN	LYS	LYS	SER	VAL	ARG	ILE
GLY	GLY	LEU	THR	ILE	ARG	GLY	THR
ALA	ALA	ASN	ALA	ASN	GLY	GLN	TYR
GLN	GLN	ALA	ALA	LEU	ASN	TYR	SER
ALA	ALA	ALA	LEU	GLU	SER	SER	SER
PRO	PRO	GLN	ALA	SER	PRO	GLY	SER
LYS	LYS	ALA	GLN	ASN	VAL	PHE	MET
VAL	VAL	PRO	HTS	LEU	VAL	VAL	TRP
ASP	ASP	LYS	ASP	ALA	ASN	ARG	GLU
ALA	ALA	VAL	GLU	LYS	TYR	THR	GLU
ASN	ASN	ASP	ALA	GLY	ASN	LEU	VAL
SER	SER	ASN	VAL	GLU	LYS	PHE	ASN
ALA	ALA	LYS	ASP	LYS	ALA	GLN	GLY
GLY	GLY	PHE	ASN	ALA	THR	GLN	PRO
LYS	LYS	ASN	ASN	ASN	LYS	MET	GLU
SER	SER	LYS	PHE	VAL	ARG	ARG	SER
THR	THR	GLU	ASN	LEU	LEU	ASP	VAL
		GLN	LYS	ILE	THR	VAL	LEU
		ASN	GLN	GLN	VAL	LEU	VAL
		PHE	GLN	GLY	GLY	GLY	ASN
		ALA	GLN	GLY	LYS	PHE	THR
		THR	ALA	VAL	ASP	ASP	TRP
		ILE	PHE	VAL	ALA	THR	ILE
		LEU	TYR	LEU	GLY	VAL	GLN
		GLU	GLU	VAL	ILE	GLN	ARG
		HIS	ILE	MET	LEU	ILE	ASN
		LEU	LEU	LYS	THR	ILE	TRP
		PRO	HIS	ARG	GLU	LYS	GLU
		ASN	LEU	LYS	ASP	LEU	THR
		LEU	PRO	ARG	PRO	PRO	VAL
		ASN	ASN	ASP	ASP	PHE	LYS
		GLU	LEU	SER	GLY	ALA	ILE
		GLU	ASN	SER	THR	GLN	THR
		ARG	GLU	ILE	ALA	ALA	THR

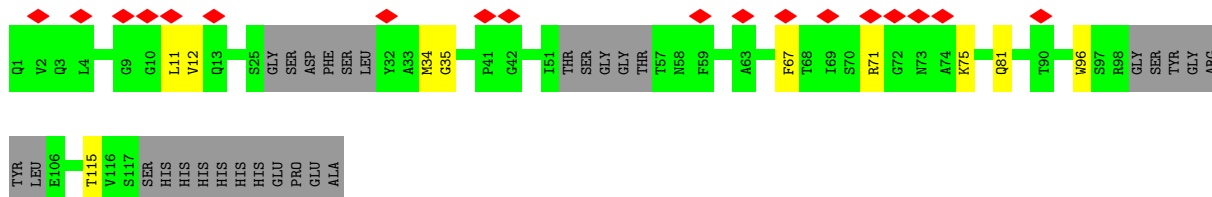
- Molecule 4: RNA (5'-R(P\*GP\*GP\*CP\*CP\*UP\*GP\*CP\*U)-3')



- Molecule 5: RNA (5'-R(P\*AP\*GP\*UP\*AP\*GP\*AP\*AP\*AP\*CP\*AP\*AP\*GP\*GP\*CP\*C)-3')



- Molecule 6: Nanobody8189 core





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	53970	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	64	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	8.013	Depositor
Minimum map value	-3.806	Depositor
Average map value	0.026	Depositor
Map value standard deviation	0.226	Depositor
Recommended contour level	1.33	Depositor
Map size (Å)	280.0, 280.0, 280.0	wwPDB
Map dimensions	140, 140, 140	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.0, 2.0, 2.0	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.25	0/4234	0.48	0/5712
2	B	0.25	0/5987	0.49	0/8075
3	C	0.26	0/2057	0.51	0/2773
4	E	0.18	0/187	0.75	0/289
5	D	0.20	0/366	0.71	0/569
6	F	0.73	0/778	1.09	4/1049 (0.4%)
All	All	0.30	0/13609	0.55	4/18467 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	F	67	PHE	CB-CG-CD1	7.06	125.74	120.80
6	F	67	PHE	CB-CG-CD2	-5.51	116.94	120.80
6	F	71	ARG	NE-CZ-NH1	5.39	123.00	120.30
6	F	34	MET	CG-SD-CE	-5.21	91.86	100.20

There are no chirality outliers.

There are no planarity outliers.

### 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	4145	4111	4116	51	0
2	B	5871	5833	5847	93	0
3	C	2017	2058	2058	44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	169	86	87	3	0
5	D	326	165	165	2	0
6	F	766	0	741	2	0
All	All	13294	12253	13014	171	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 171 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:522:ASP:OD2	2:B:559:TYR:OH	1.91	0.88
5:D:5:G:OP2	5:D:5:G:N2	2.10	0.84
1:A:300:GLU:O	1:A:488:LYS:NZ	2.17	0.77
1:A:476:ALA:O	1:A:506:LYS:NZ	2.19	0.76
2:B:636:GLU:OE2	3:C:32:LYS:NZ	2.19	0.76

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	514/716 (72%)	494 (96%)	20 (4%)	0	100	100
2	B	731/757 (97%)	709 (97%)	22 (3%)	0	100	100
3	C	247/905 (27%)	233 (94%)	14 (6%)	0	100	100
6	F	91/128 (71%)	88 (97%)	3 (3%)	0	100	100
All	All	1583/2506 (63%)	1524 (96%)	59 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	461/640 (72%)	461 (100%)	0	100	100
2	B	648/669 (97%)	644 (99%)	4 (1%)	84	88
3	C	223/788 (28%)	223 (100%)	0	100	100
6	F	80/103 (78%)	77 (96%)	3 (4%)	28	49
All	All	1412/2200 (64%)	1405 (100%)	7 (0%)	85	89

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	745	LYS
6	F	12	VAL
6	F	81	GLN
6	F	75	LYS
2	B	734	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	E	7/17 (41%)	1 (14%)	0
5	D	14/15 (93%)	4 (28%)	0
All	All	21/32 (65%)	5 (23%)	0

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	E	2	G
5	D	6	A
5	D	8	A
5	D	11	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	D	15	C

There are no RNA pucker outliers to report.

#### 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 oligosaccharides in this entry.

#### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

#### 5.7 Other polymers [i](#)

There are no such residues in this entry.

#### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-12361. These allow visual inspection of the internal detail of the map and identification of artifacts.

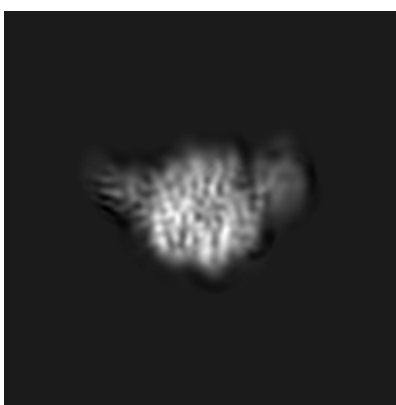
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

#### 6.1.1 Primary map



X



Y



Z

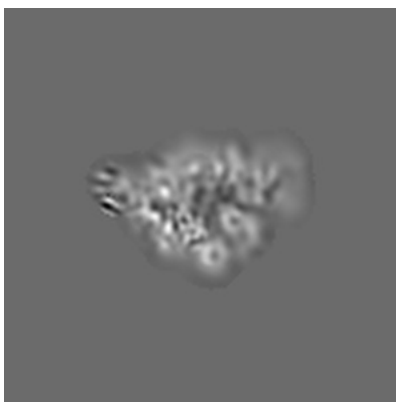
The images above show the map projected in three orthogonal directions.

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

#### 6.2.1 Primary map



X Index: 70



Y Index: 70

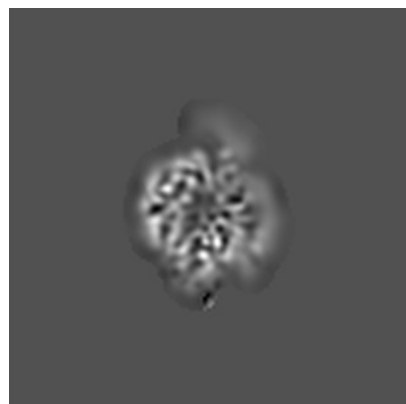


Z Index: 70

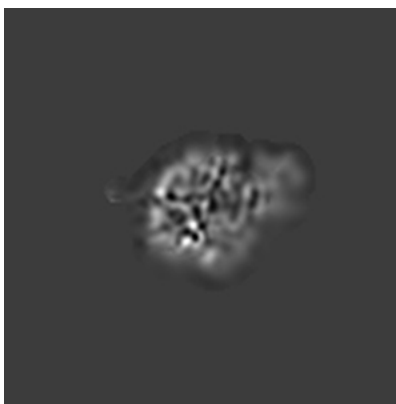
The images above show central slices of the map in three orthogonal directions.

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

### 6.3.1 Primary map



X Index: 68



Y Index: 76



Z Index: 77

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

## 6.4 Orthogonal standard-deviation projections (False-color) [i](#)

### 6.4.1 Primary map



X



Y

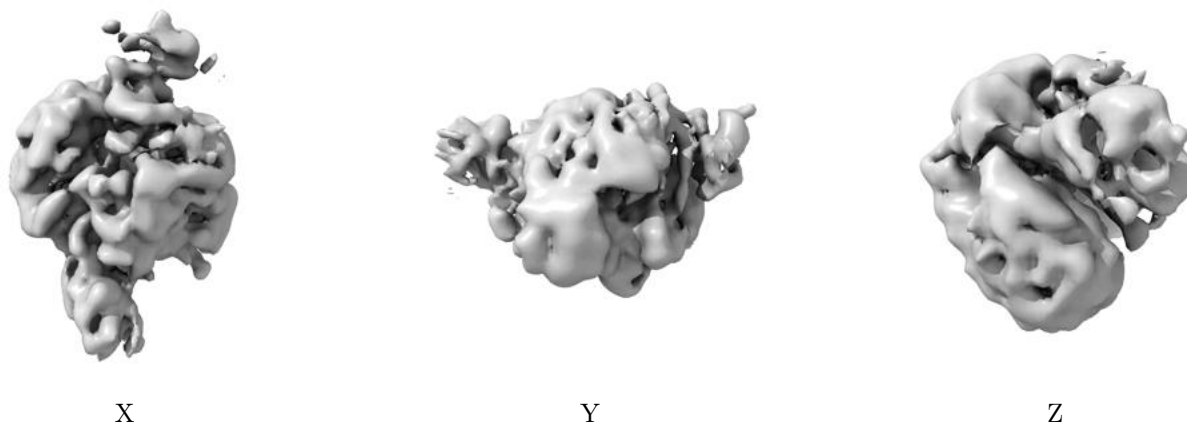


Z

The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

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

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 1.33. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

## 6.6 Mask visualisation [i](#)

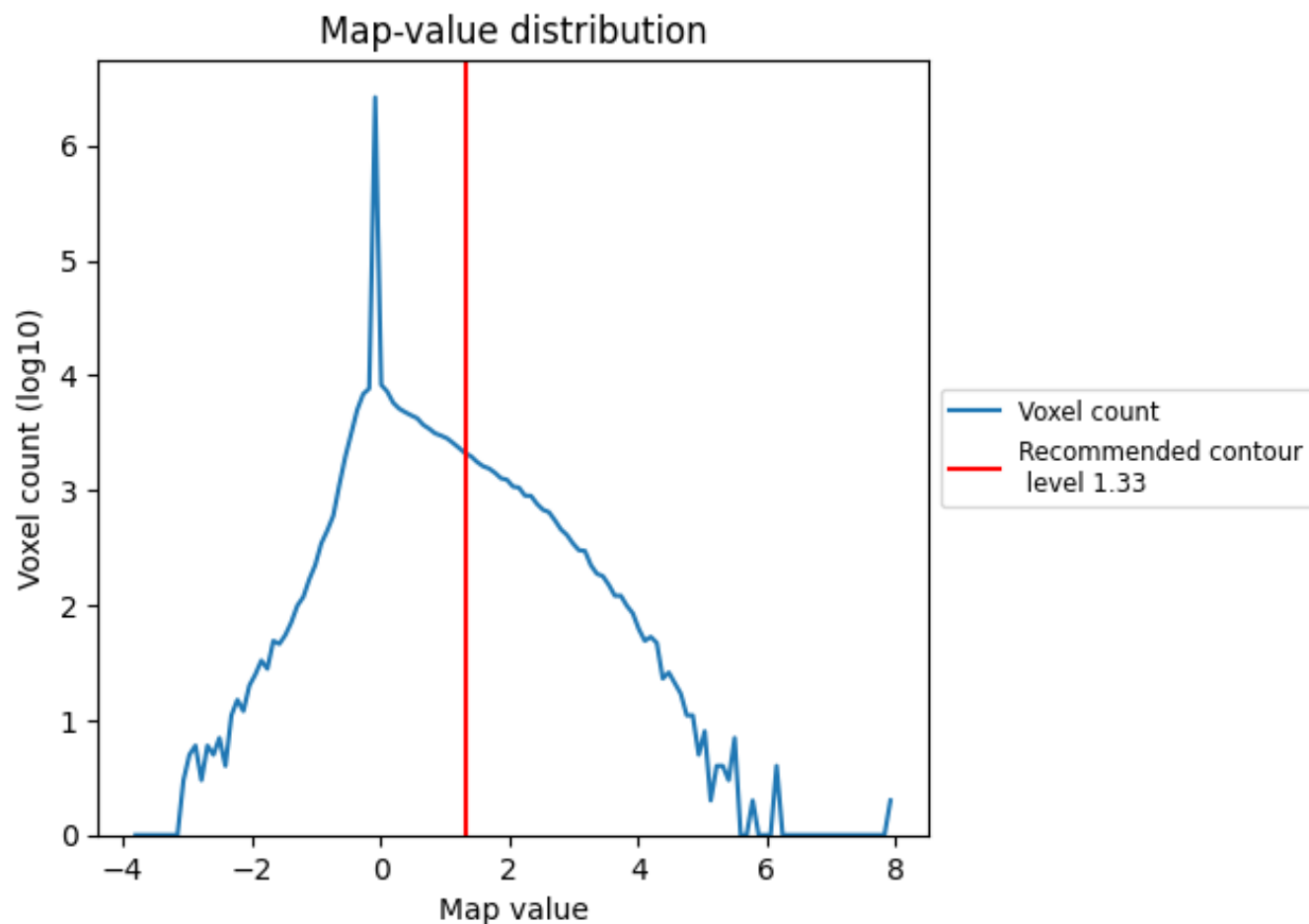
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

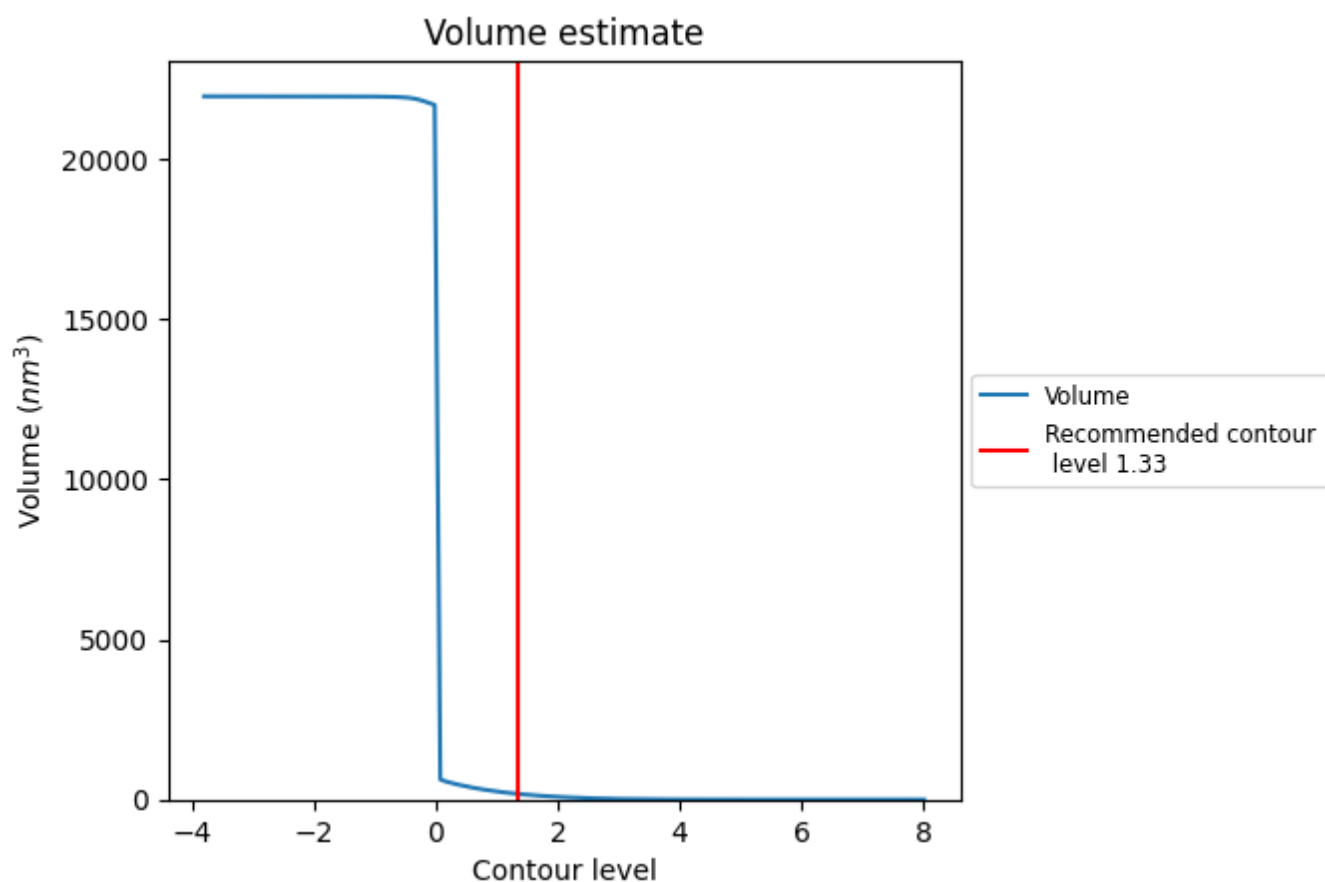
This section contains the results of statistical analysis of the map.

### 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.

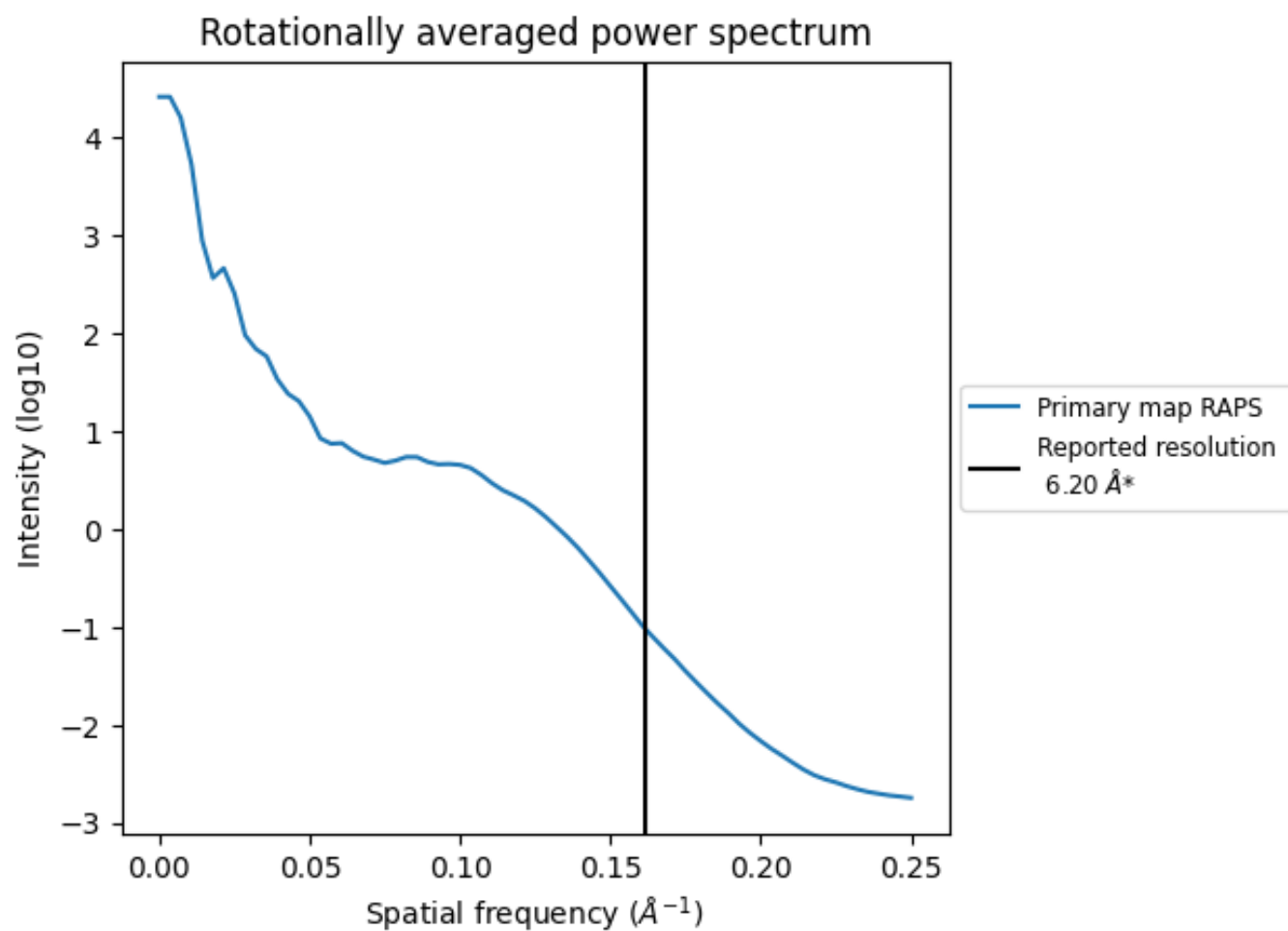
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 180  $\text{nm}^3$ ; this corresponds to an approximate mass of 162 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 0.161 Å<sup>-1</sup>

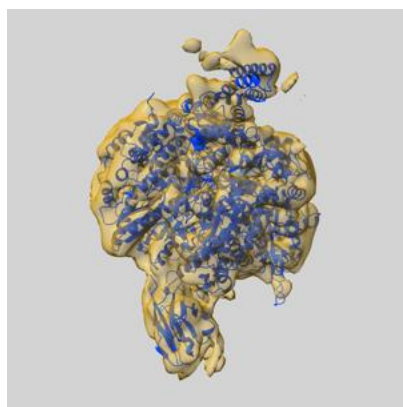
## 8 Fourier-Shell correlation

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

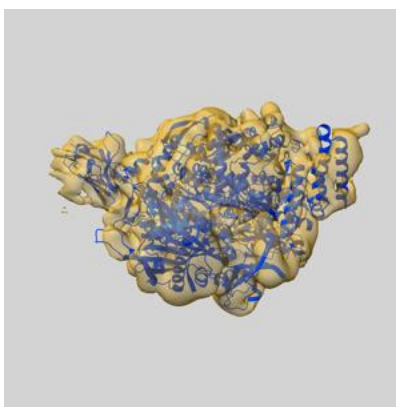
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-12361 and PDB model 7NIK. Per-residue inclusion information can be found in section [3](#) on page [6](#).

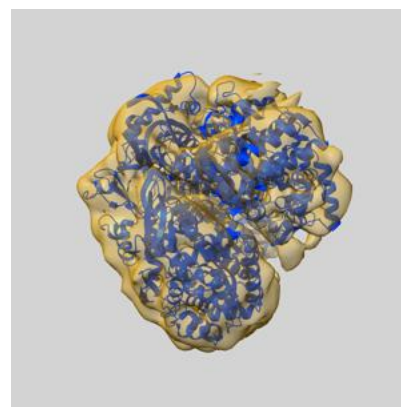
### 9.1 Map-model overlay [i](#)



X



Y



Z

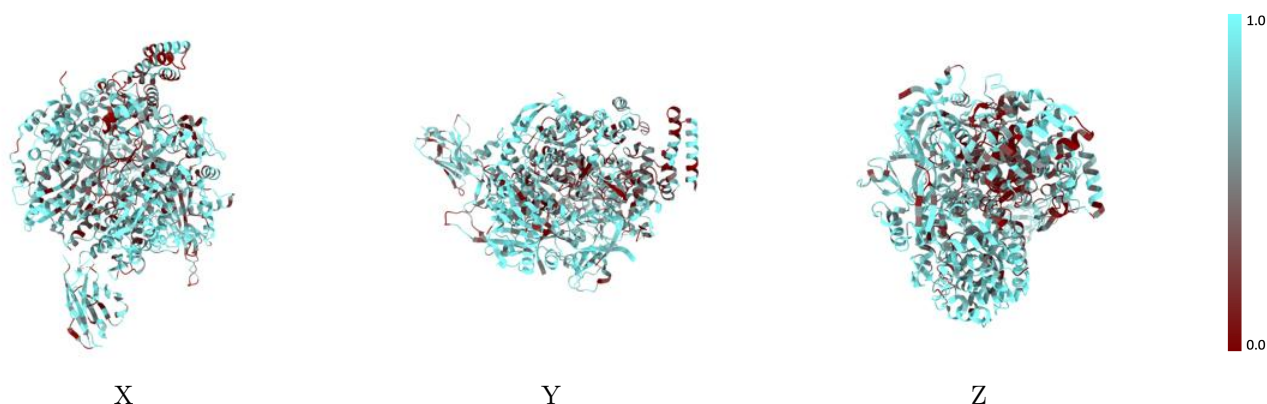
The images above show the 3D surface view of the map at the recommended contour level 1.33 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



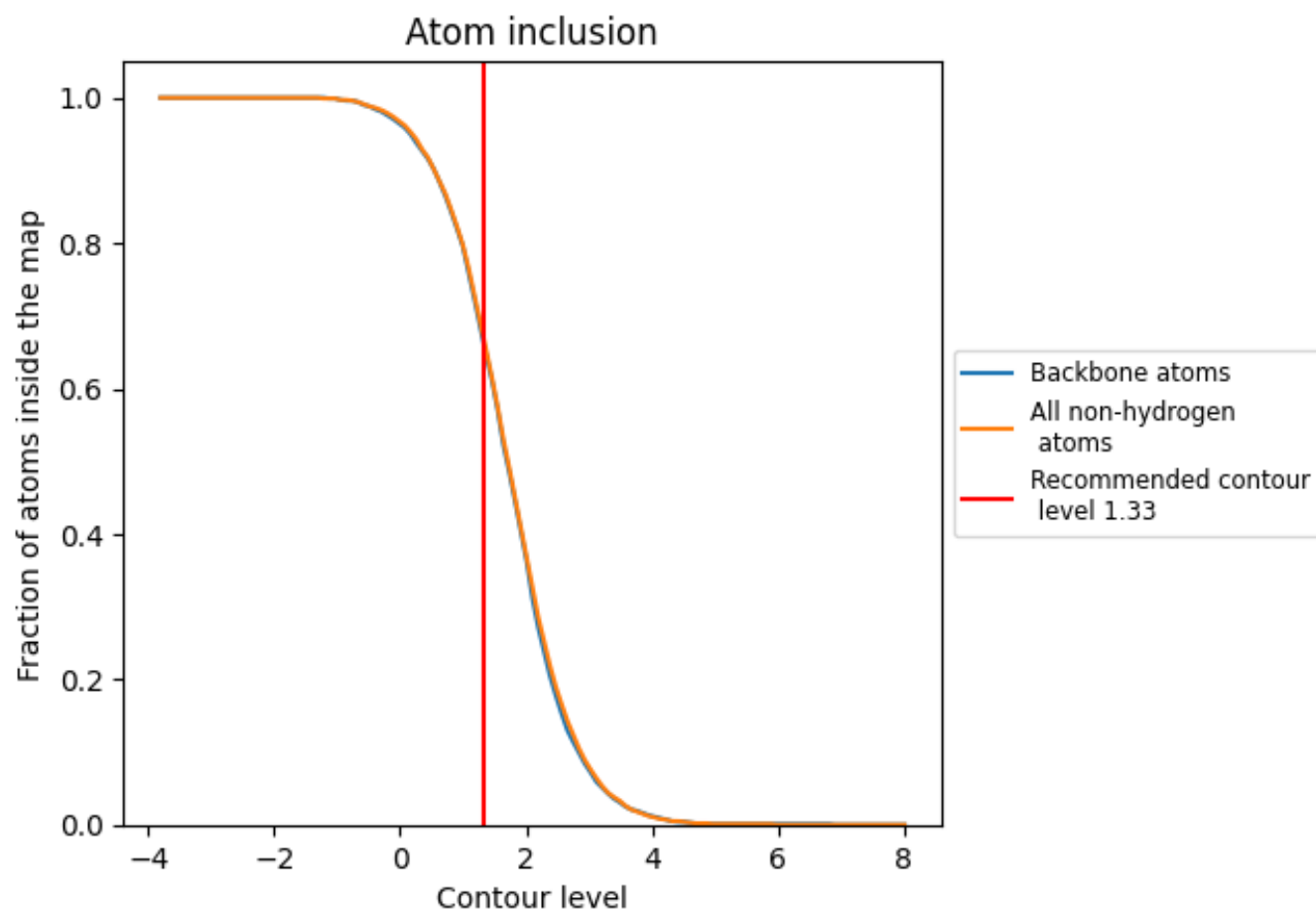
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (1.33).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 66% of all backbone atoms, 67% 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 (1.33) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div><div></div>0.6680</div>	<div><div></div>0.1030</div>
A	<div><div></div>0.7220</div>	<div><div></div>0.0990</div>
B	<div><div></div>0.6420</div>	<div><div></div>0.1090</div>
C	<div><div></div>0.5730</div>	<div><div></div>0.0920</div>
D	<div><div></div>0.8220</div>	<div><div></div>0.0850</div>
E	<div><div></div>0.7510</div>	<div><div></div>0.1470</div>
F	<div><div></div>0.7160</div>	<div><div></div>0.1100</div>

1.0

0.0

<0.0