



# wwPDB EM Validation Summary Report ⓘ

Jun 16, 2025 – 03:03 PM EDT

PDB ID : 9C29 / pdb\_00009c29  
EMDB ID : EMD-45151  
Title : Hexadecamer of NL4-3 WT HIV-1 intasome  
Authors : Lyumkis, D.; Jing, T.; Zhang, Z.  
Deposited on : 2024-05-30  
Resolution : 8.00 Å (reported)  
Based on initial models : 1EX4, 7Z1Z, 6PUT, 1IHV, 1K6Y

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.dev118  
MolProbity : 4-5-2 with Phenix2.0rc1  
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) : 2.44

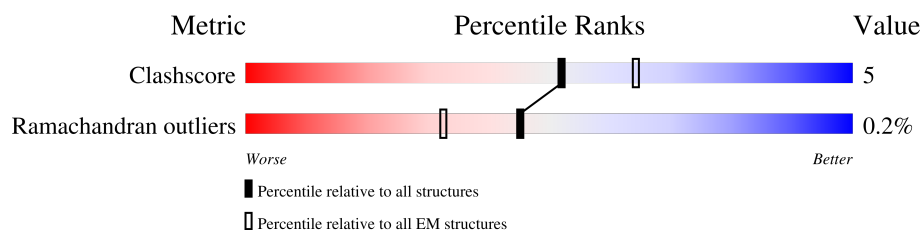
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.00 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

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	288	
1	B	288	
1	C	288	
1	D	288	
1	E	288	
1	F	288	
1	G	288	
1	H	288	
1	I	288	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	J	288	
1	K	288	
1	L	288	
1	M	288	
1	N	288	
1	O	288	
1	P	288	
2	Q	19	
2	S	19	
3	R	17	
3	T	17	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 21300 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 Integrase.

Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	261	Total	C	N	O	0	0
			1279	762	261	256		
1	B	242	Total	C	N	O	0	0
			1192	708	242	242		
1	C	261	Total	C	N	O	0	0
			1284	762	261	261		
1	D	250	Total	C	N	O	0	0
			1230	731	249	250		
1	E	255	Total	C	N	O	0	0
			1256	746	255	255		
1	F	252	Total	C	N	O	0	0
			1241	737	252	252		
1	G	242	Total	C	N	O	0	0
			1191	707	242	242		
1	H	252	Total	C	N	O	0	0
			1241	737	252	252		
1	I	261	Total	C	N	O	0	0
			1279	762	261	256		
1	J	242	Total	C	N	O	0	0
			1192	708	242	242		
1	K	261	Total	C	N	O	0	0
			1284	762	261	261		
1	L	250	Total	C	N	O	0	0
			1230	731	249	250		
1	M	255	Total	C	N	O	0	0
			1256	746	255	255		
1	N	252	Total	C	N	O	0	0
			1241	737	252	252		
1	O	242	Total	C	N	O	0	0
			1191	707	242	242		
1	P	252	Total	C	N	O	0	0
			1241	737	252	252		

- Molecule 2 is a DNA chain called DNA (5'-D(\*AP\*CP\*TP\*GP\*CP\*TP\*AP\*GP\*AP\*GP\*

AP\*TP\*TP\*TP\*TP\*CP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
2	Q	19	Total	C	N	O	P	0	0
			384	185	67	114	18		
2	S	19	Total	C	N	O	P	0	0
			384	185	67	114	18		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	17	Total	C	N	O	P	0	0
			350	166	68	99	17		
3	T	17	Total	C	N	O	P	0	0
			350	166	68	99	17		


- Molecule 4 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

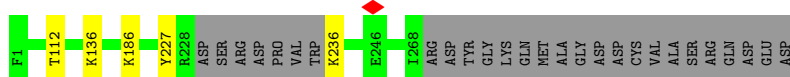
Mol	Chain	Residues	Atoms		AltConf
4	A	2	Total	Mg	0
			2	2	
4	I	2	Total	Mg	0
			2	2	

### 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.

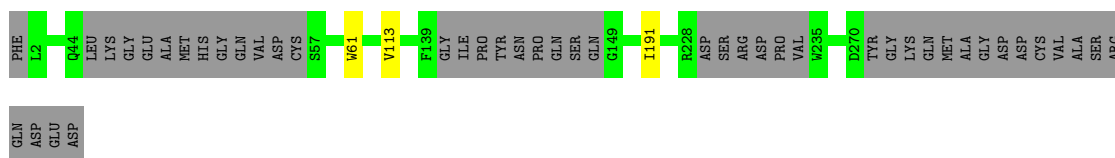
- Molecule 1: Integrase

Chain A: 




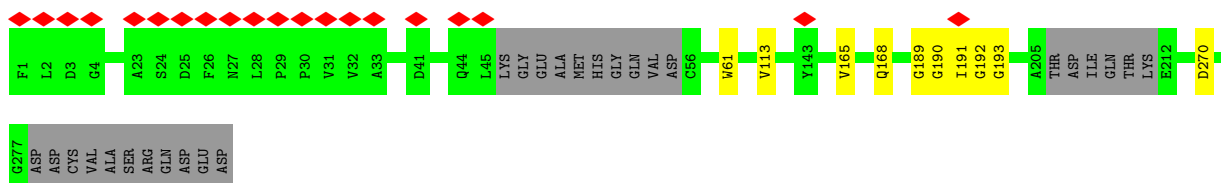
- Molecule 1: Integrase

Chain B: 




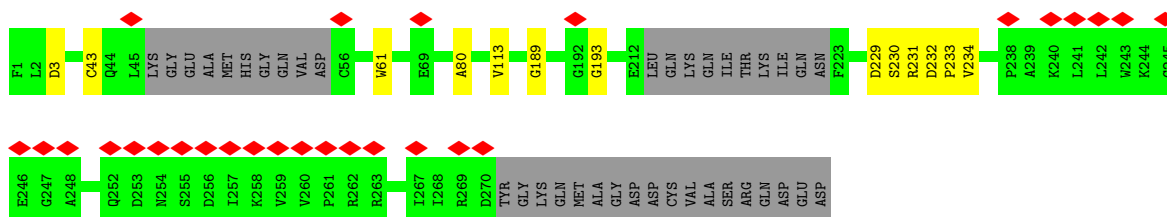
- Molecule 1: Integrase

Chain C: 

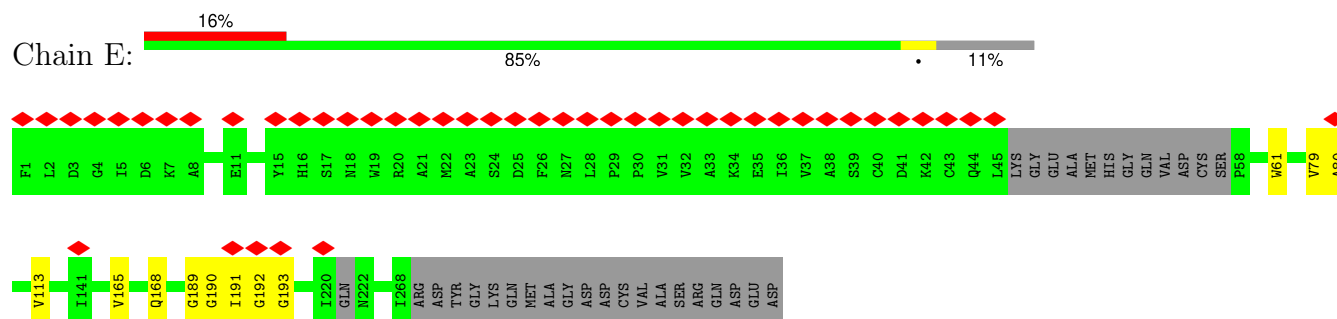


- Molecule 1: Integrase

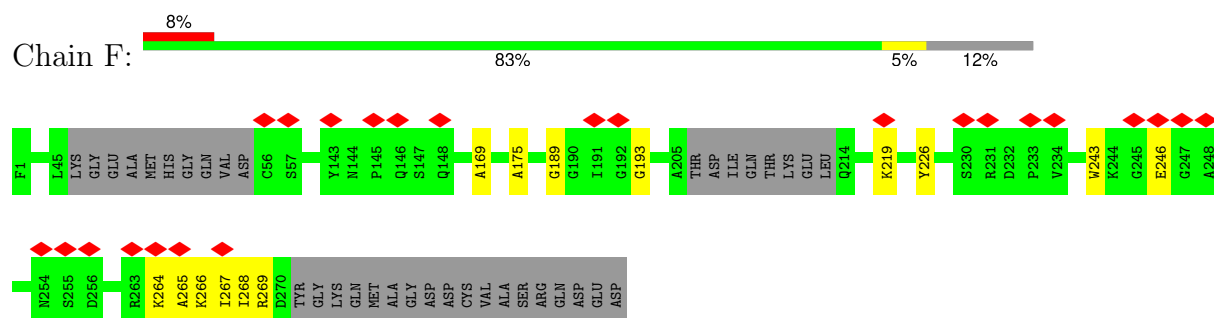
Chain D: 



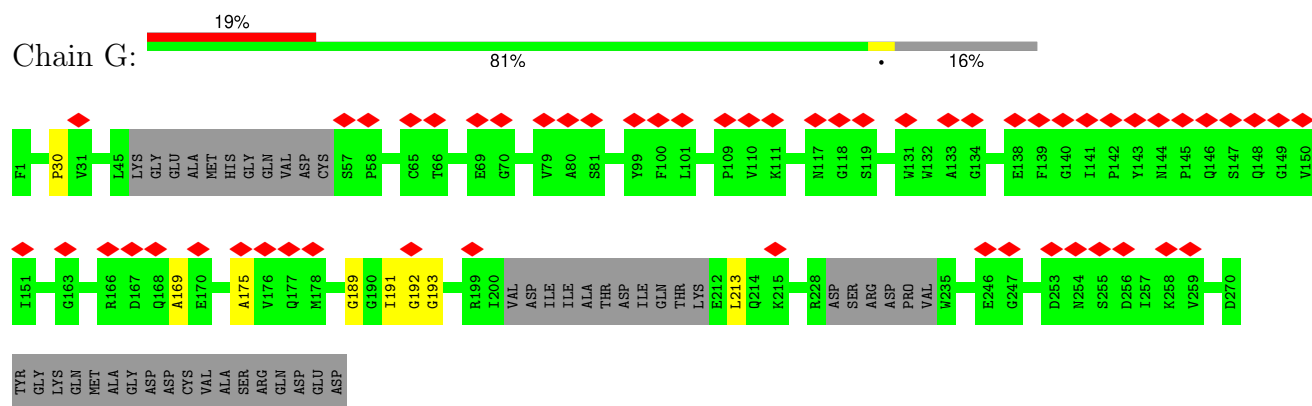
- Molecule 1: Integrase



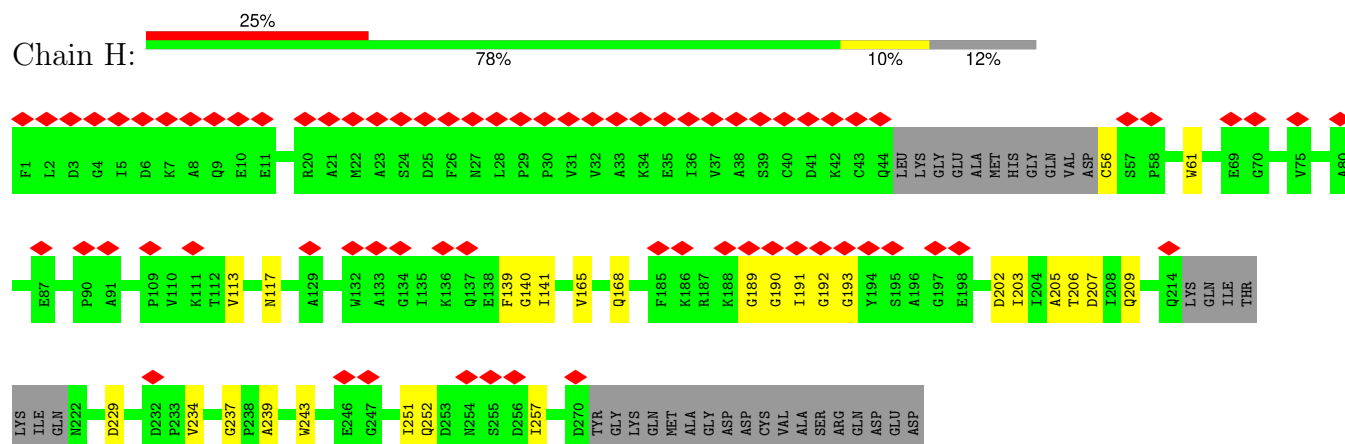
- Molecule 1: Integrase




- Molecule 1: Integrase

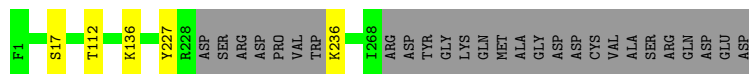


- Molecule 1: Integrase




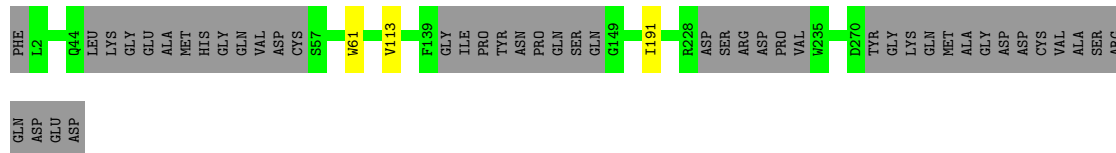
- Molecule 1: Integrase

Chain I:  89% 9%

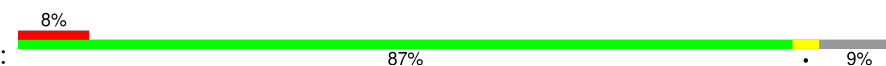


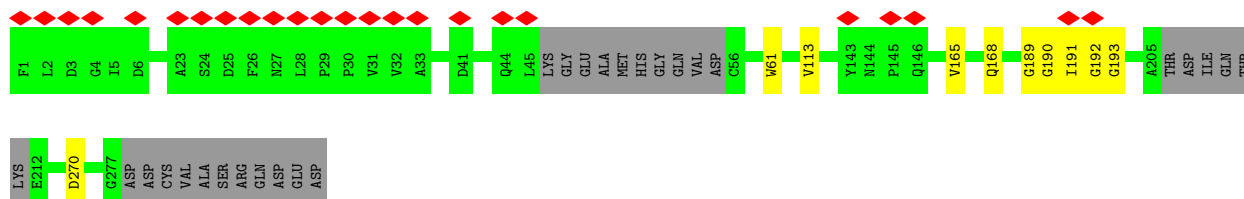
- Molecule 1: Integrase

Chain J:  83% 16%




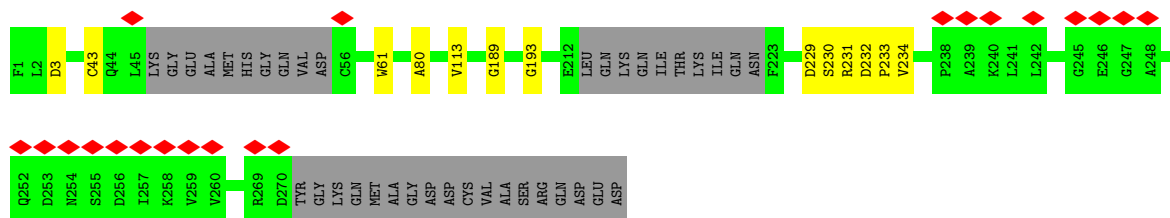
- Molecule 1: Integrase

Chain K:  87% 9%




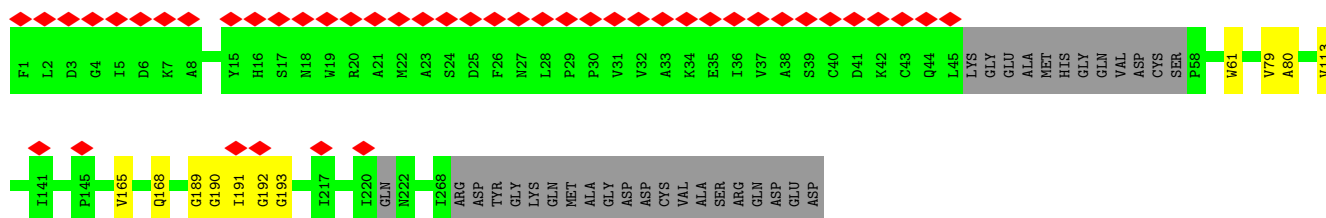
- Molecule 1: Integrase

Chain L:  82% 5% 13%



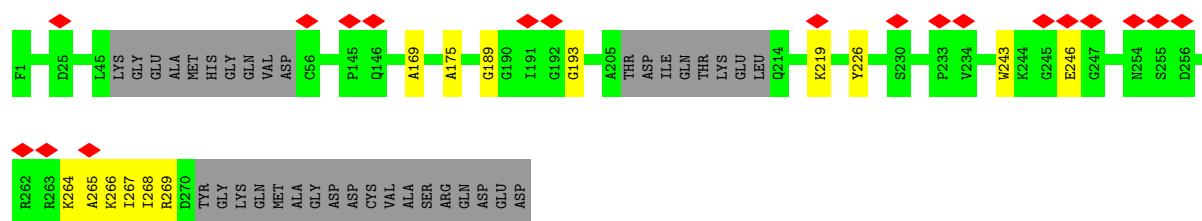
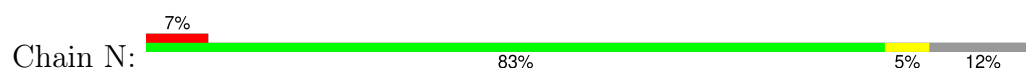
- Molecule 1: Integrase

Chain M:  85% 11%

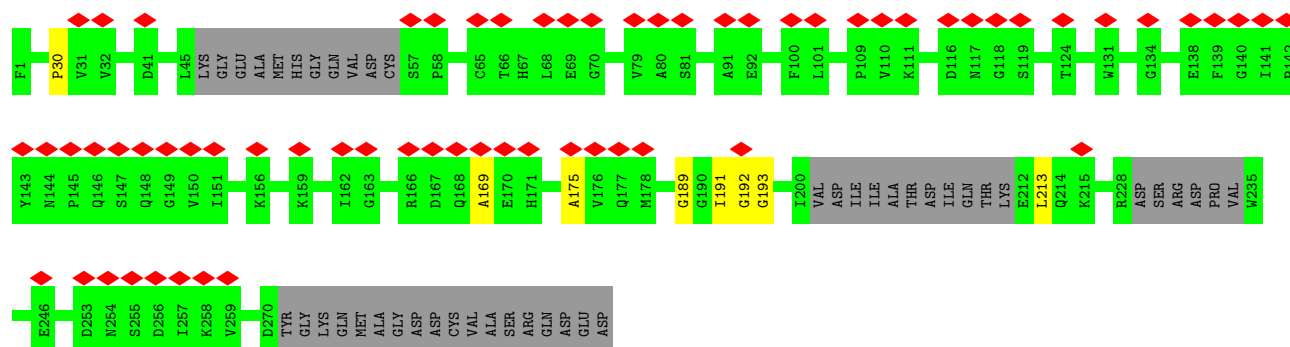
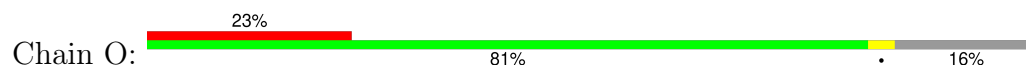


- Molecule 1: Integrase

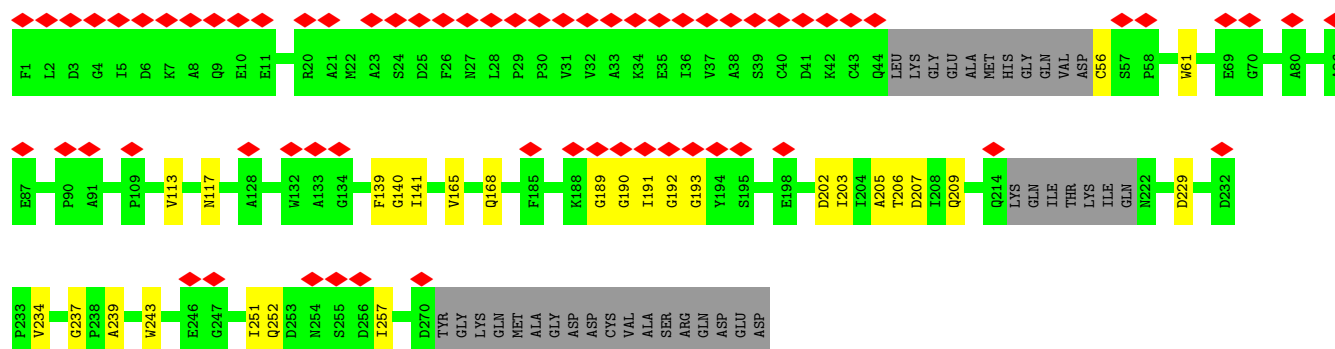
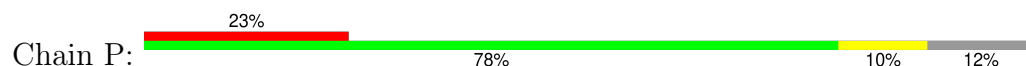




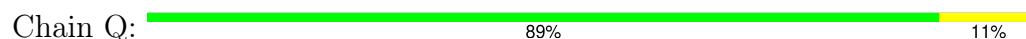
• Molecule 1: Integrase




• Molecule 1: Integrase



• Molecule 2: DNA (5'-D(\*AP\*CP\*TP\*GP\*CP\*TP\*AP\*GP\*AP\*GP\*AP\*TP\*TP\*TP\*TP\*CP\*CP\*CP\*G)-3')




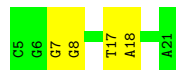
• Molecule 2: DNA (5'-D(\*AP\*CP\*TP\*GP\*CP\*TP\*AP\*GP\*AP\*GP\*AP\*TP\*TP\*TP\*TP\*CP\*CP\*CP\*G)-3')

Chain S:  84% 16%




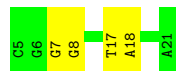
● Molecule 3: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*AP\*AP\*AP\*TP\*CP\*TP\*CP\*TP\*AP\*GP\*C  
P\*A)-3')

Chain R:  76% 24%



● Molecule 3: DNA (5'-D(P\*CP\*GP\*GP\*GP\*AP\*AP\*AP\*AP\*TP\*CP\*TP\*CP\*TP\*AP\*GP\*C  
P\*A)-3')

Chain T:  76% 24%



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	22780	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	29000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	1.820	Depositor
Minimum map value	-0.282	Depositor
Average map value	0.015	Depositor
Map value standard deviation	0.094	Depositor
Recommended contour level	0.4	Depositor
Map size (Å)	324.8, 324.8, 324.8	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.015, 1.015, 1.015	Depositor

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.31	0/1277	0.54	0/1769
1	B	0.23	0/1188	0.45	0/1646
1	C	0.24	0/1281	0.47	0/1776
1	D	0.32	0/1227	0.70	2/1701 (0.1%)
1	E	0.25	0/1253	0.48	0/1738
1	F	0.21	0/1238	0.46	0/1717
1	G	0.22	0/1187	0.47	0/1644
1	H	0.24	0/1238	0.48	0/1717
1	I	0.31	0/1277	0.54	0/1769
1	J	0.23	0/1188	0.45	0/1646
1	K	0.24	0/1281	0.47	0/1776
1	L	0.32	0/1227	0.70	2/1701 (0.1%)
1	M	0.25	0/1253	0.48	0/1738
1	N	0.21	0/1238	0.46	0/1717
1	O	0.22	0/1187	0.47	0/1644
1	P	0.25	0/1238	0.48	0/1717
2	Q	0.55	0/429	0.57	0/660
2	S	0.56	0/429	0.57	0/660
3	R	0.44	0/393	0.50	0/604
3	T	0.44	0/393	0.50	0/604
All	All	0.28	0/21422	0.51	4/29944 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	43	CYS	N-CA-C	6.42	119.67	109.65
1	D	43	CYS	N-CA-C	6.42	119.66	109.65
1	D	3	ASP	N-CA-C	-5.76	106.20	113.23
1	L	3	ASP	N-CA-C	-5.74	106.23	113.23

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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	1279	0	604	3	0
1	B	1192	0	557	2	0
1	C	1284	0	602	4	0
1	D	1230	0	571	11	0
1	E	1256	0	586	9	0
1	F	1241	0	579	42	0
1	G	1191	0	556	8	0
1	H	1241	0	579	52	0
1	I	1279	0	604	3	0
1	J	1192	0	557	2	0
1	K	1284	0	602	4	0
1	L	1230	0	571	11	0
1	M	1256	0	586	9	0
1	N	1241	0	579	42	0
1	O	1191	0	556	8	0
1	P	1241	0	579	52	0
2	Q	384	0	217	1	0
2	S	384	0	217	2	0
3	R	350	0	191	2	0
3	T	350	0	191	2	0
4	A	2	0	0	0	0
4	I	2	0	0	0	0
All	All	21300	0	10084	164	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 5.

The worst 5 of 164 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:N:265:ALA:N	1:P:56:CYS:N	1.69	1.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:L:232:ASP:O	1:P:117:ASN:CB	1.72	1.38
1:N:266:LYS:CB	1:P:56:CYS:O	1.71	1.38
1:F:266:LYS:CB	1:H:56:CYS:O	1.72	1.37
1:F:265:ALA:N	1:H:56:CYS:N	1.69	1.37

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	257/288 (89%)	251 (98%)	6 (2%)	0	100	100
1	B	234/288 (81%)	223 (95%)	11 (5%)	0	100	100
1	C	255/288 (88%)	240 (94%)	13 (5%)	2 (1%)	16	55
1	D	244/288 (85%)	231 (95%)	13 (5%)	0	100	100
1	E	249/288 (86%)	239 (96%)	9 (4%)	1 (0%)	30	68
1	F	246/288 (85%)	233 (95%)	13 (5%)	0	100	100
1	G	234/288 (81%)	220 (94%)	14 (6%)	0	100	100
1	H	246/288 (85%)	233 (95%)	12 (5%)	1 (0%)	30	68
1	I	257/288 (89%)	251 (98%)	6 (2%)	0	100	100
1	J	234/288 (81%)	223 (95%)	11 (5%)	0	100	100
1	K	255/288 (88%)	240 (94%)	13 (5%)	2 (1%)	16	55
1	L	244/288 (85%)	231 (95%)	13 (5%)	0	100	100
1	M	249/288 (86%)	239 (96%)	9 (4%)	1 (0%)	30	68
1	N	246/288 (85%)	233 (95%)	13 (5%)	0	100	100
1	O	234/288 (81%)	220 (94%)	14 (6%)	0	100	100
1	P	246/288 (85%)	233 (95%)	12 (5%)	1 (0%)	30	68

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3930/4608 (85%)	3740 (95%)	182 (5%)	8 (0%)	45 78

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	191	ILE
1	K	191	ILE
1	E	191	ILE
1	H	191	ILE
1	M	191	ILE

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.



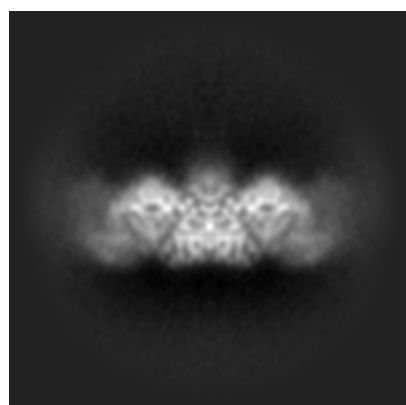
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-45151. These allow visual inspection of the internal detail of the map and identification of artifacts.

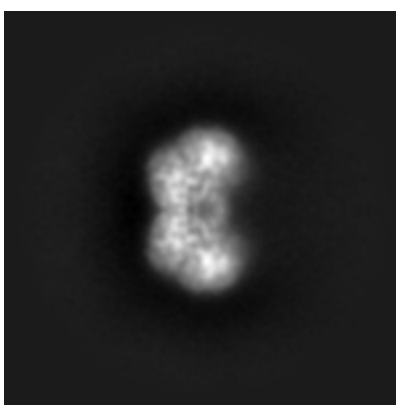
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

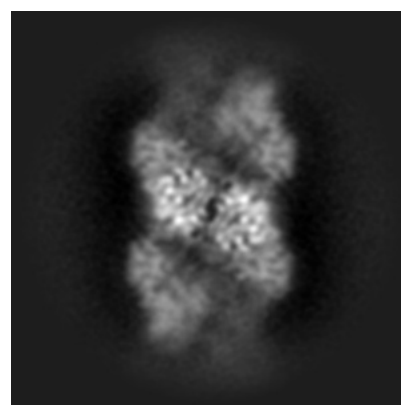
#### 6.1.1 Primary map



X



Y

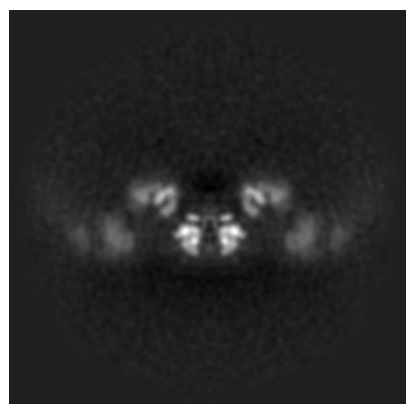


Z

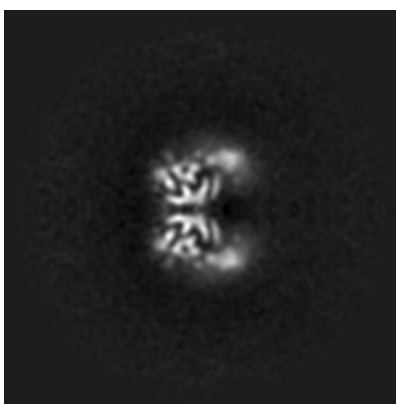
The images above show the map projected in three orthogonal directions.

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

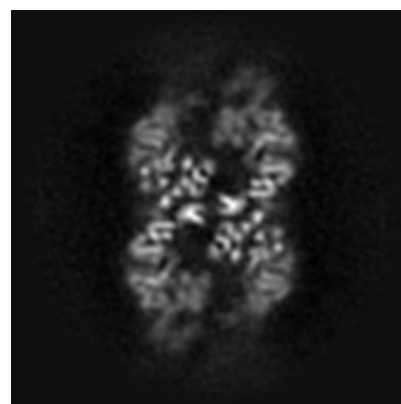
#### 6.2.1 Primary map



X Index: 160



Y Index: 160

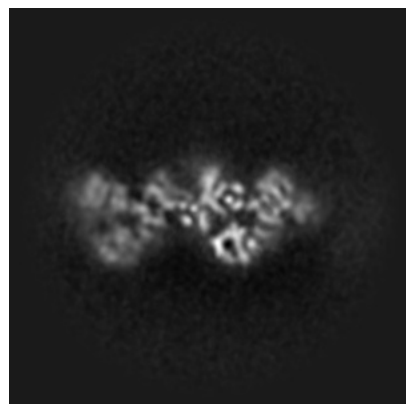


Z Index: 160

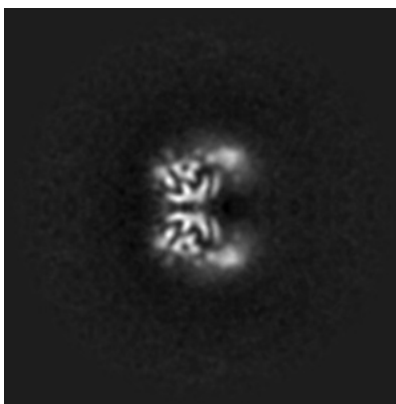
The images above show central slices of the map in three orthogonal directions.

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

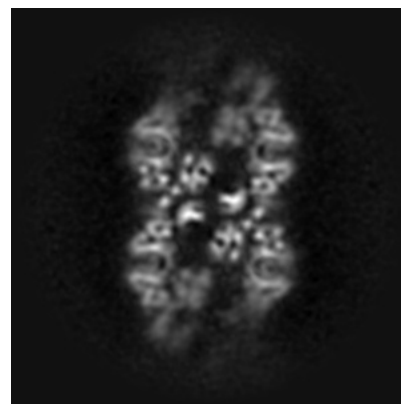
### 6.3.1 Primary map



X Index: 120



Y Index: 160

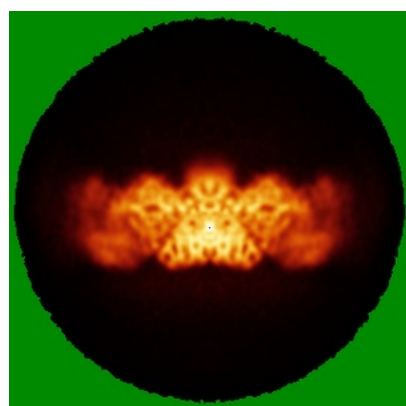


Z Index: 162

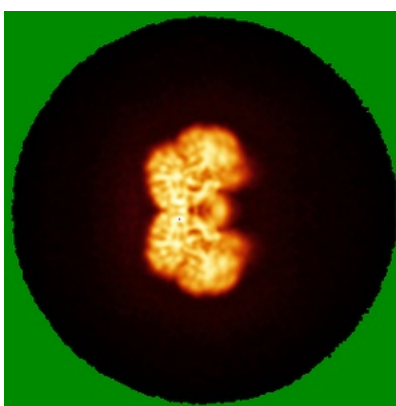
The images above show the largest variance slices of the map in three orthogonal directions.

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

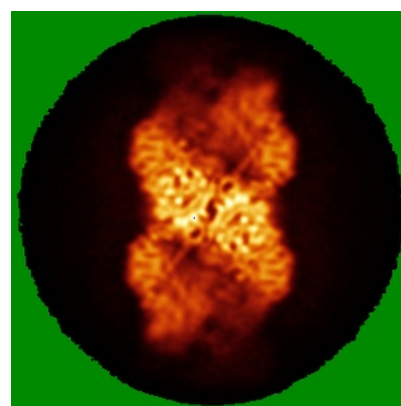
### 6.4.1 Primary map



X



Y

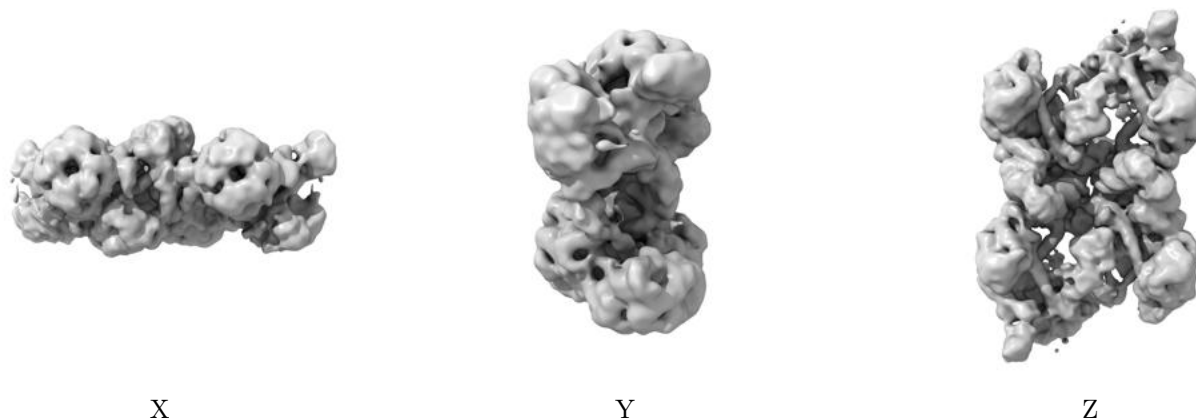


Z

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

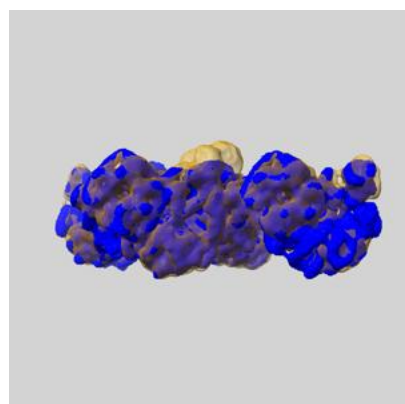
## 6.6 Mask visualisation [i](#)

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

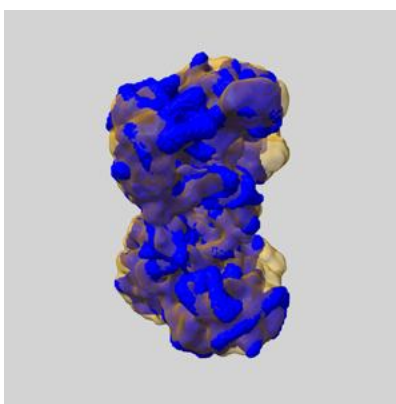
A mask typically either:

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

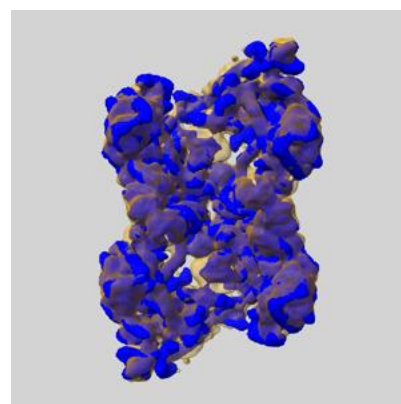
### 6.6.1 emd\_45151\_msk\_1.map [i](#)



X



Y

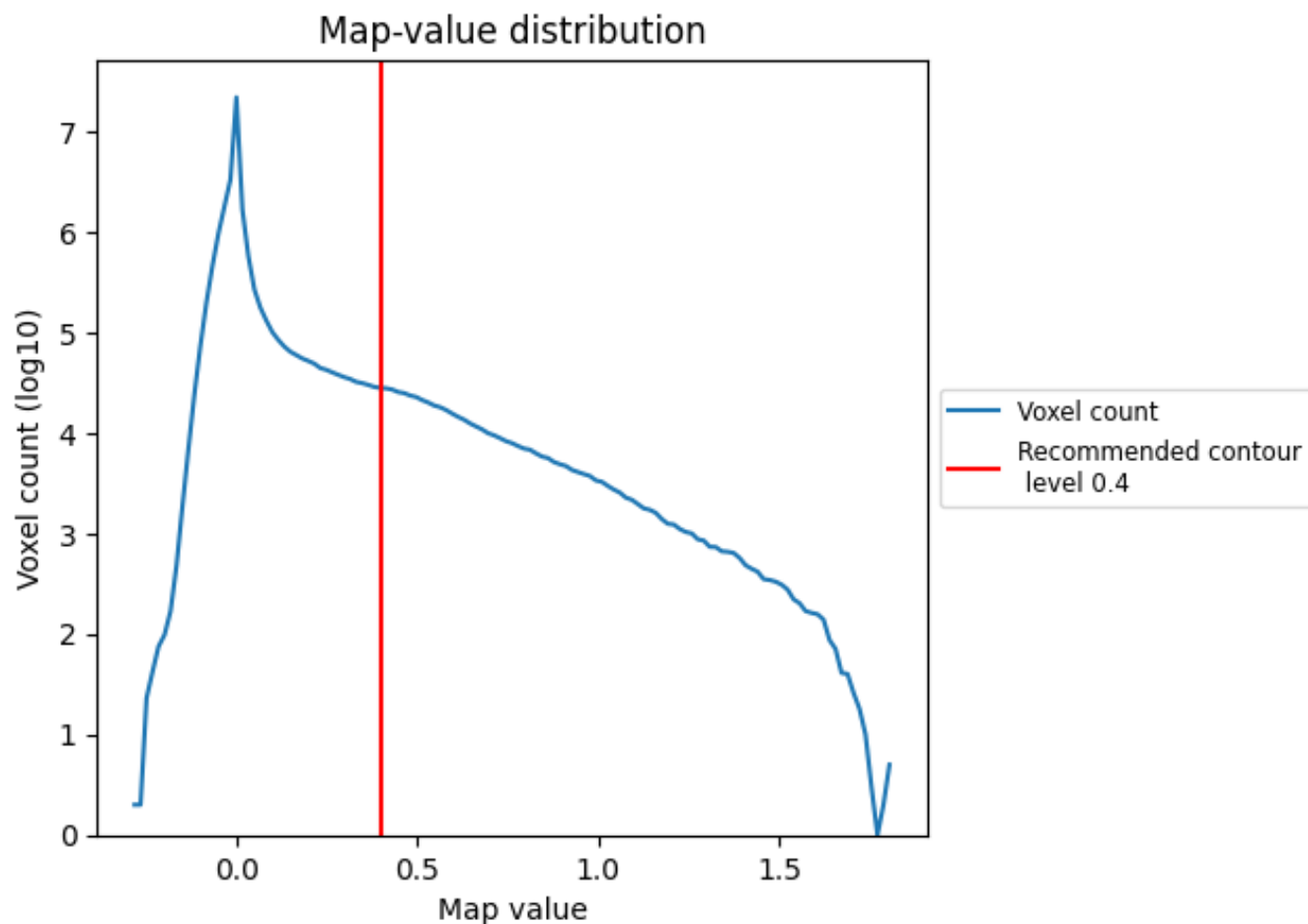


Z

## 7 Map analysis [i](#)

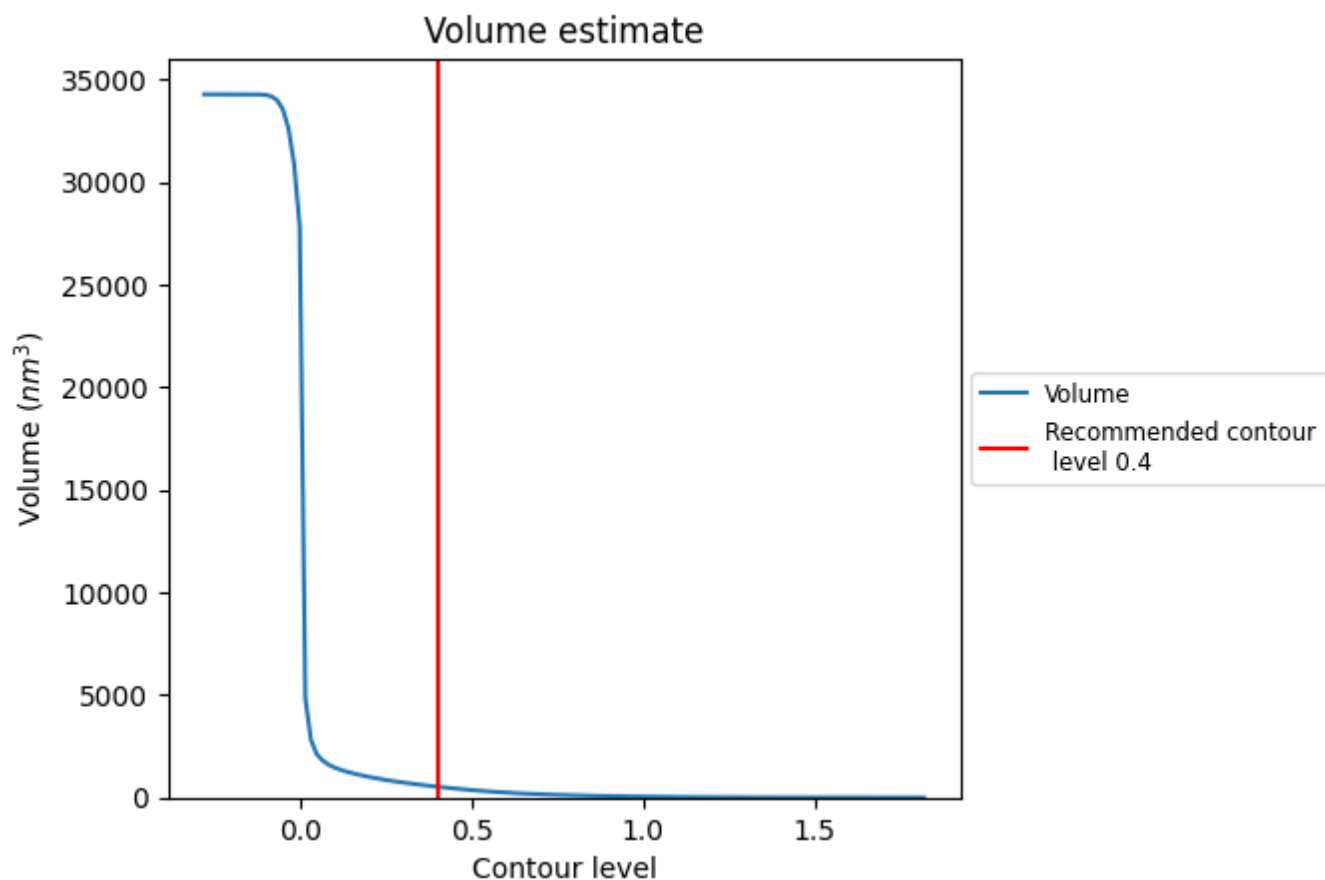
This section contains the results of statistical analysis of the map.

### 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.

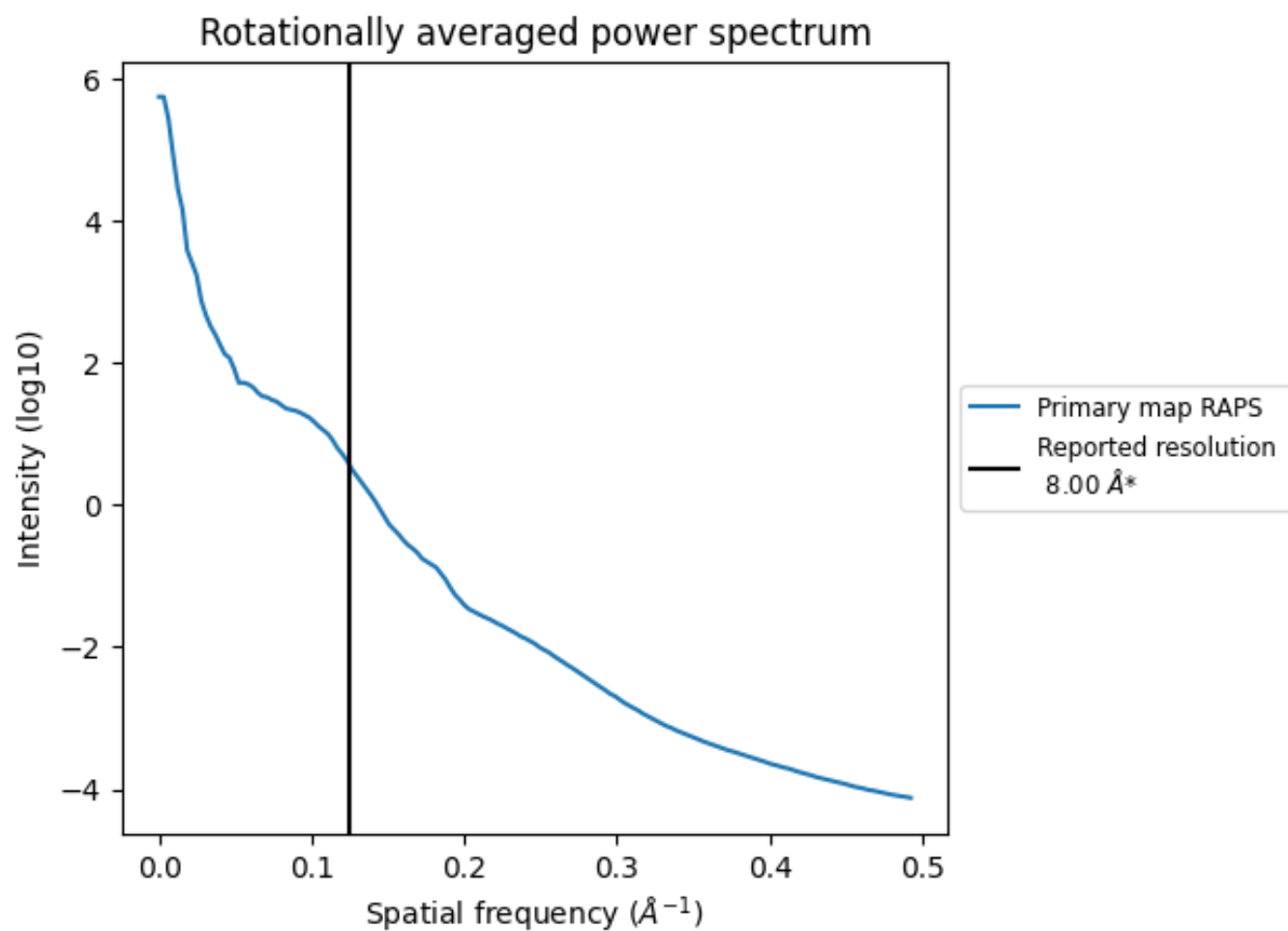
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 527 nm<sup>3</sup>; this corresponds to an approximate mass of 476 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 0.125 Å<sup>-1</sup>

## 8 Fourier-Shell correlation

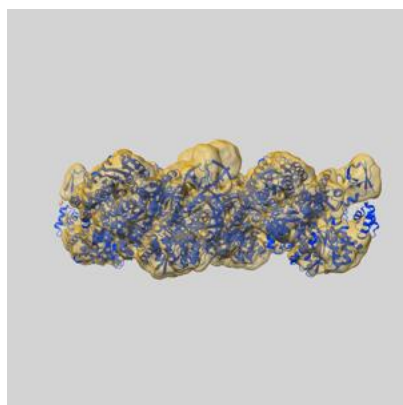
This section was not generated. No FSC curve or half-maps provided.



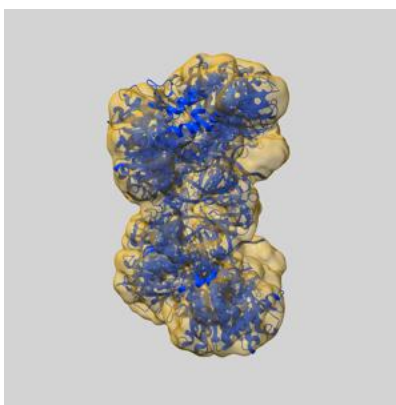
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-45151 and PDB model 9C29. Per-residue inclusion information can be found in [section 3](#) on [page 6](#).

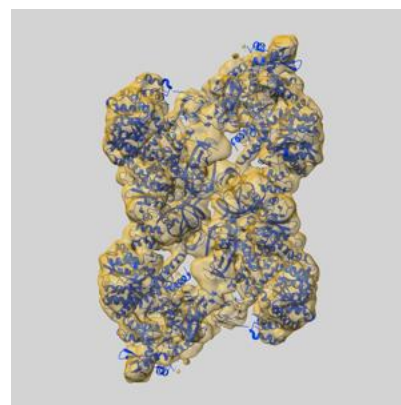
### 9.1 Map-model overlay [i](#)



X



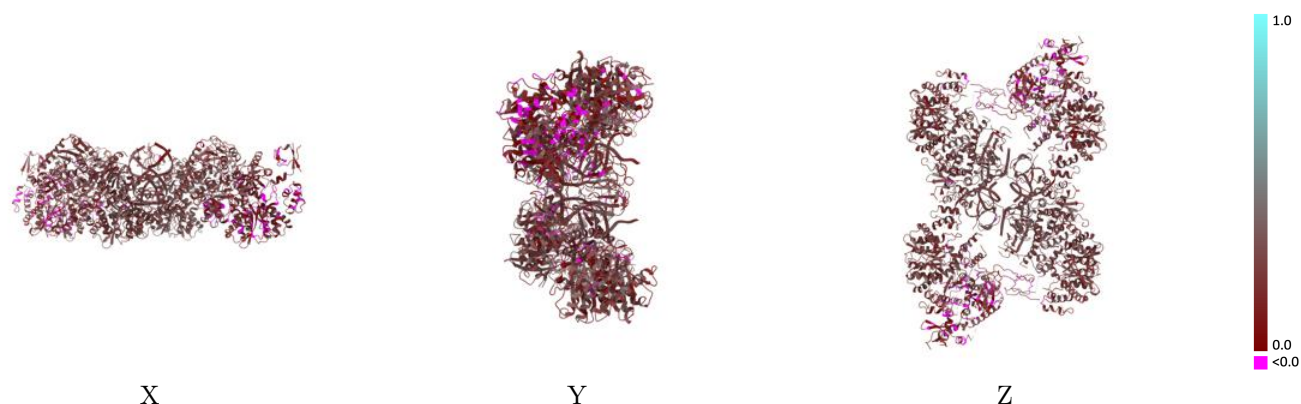
Y



Z

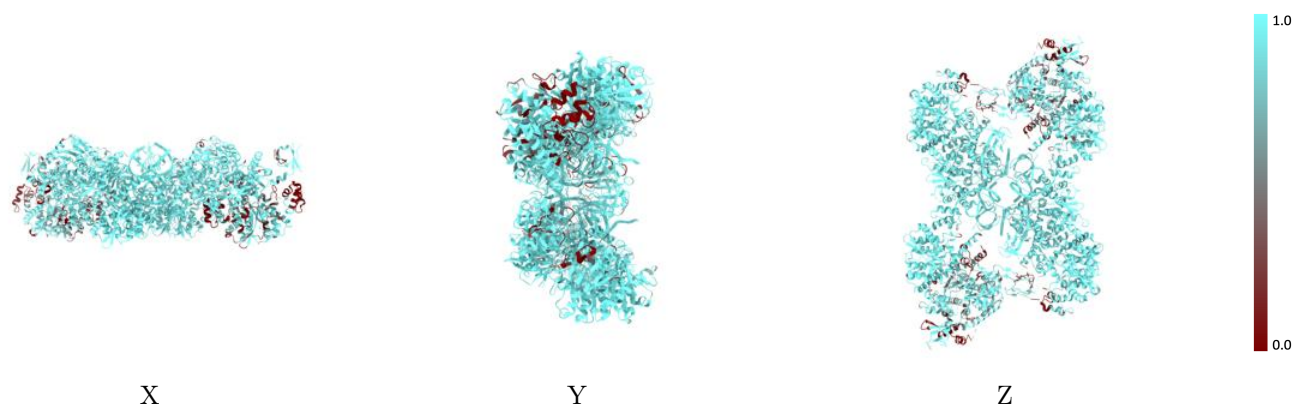
The images above show the 3D surface view of the map at the recommended contour level 0.4 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



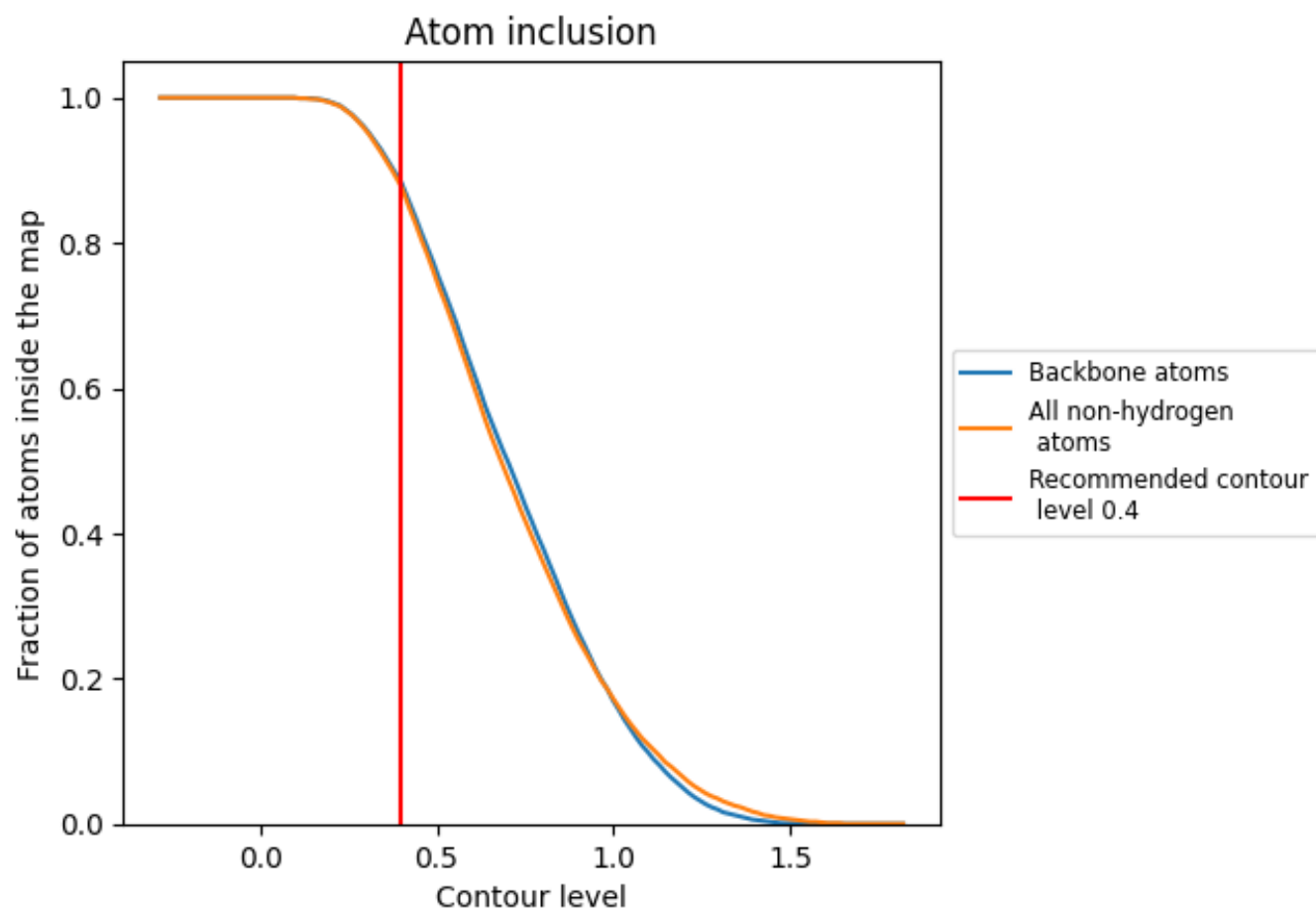
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.4).











































## 9.4 Atom inclusion ⓘ



At the recommended contour level, 88% of all backbone atoms, 88% 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 (0.4) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8770	 0.2190
A	 0.9910	 0.2750
B	 0.9950	 0.2650
C	 0.9080	 0.2430
D	 0.8760	 0.2120
E	 0.8100	 0.2190
F	 0.8940	 0.2030
G	 0.7520	 0.1540
H	 0.6980	 0.1370
I	 0.9960	 0.2820
J	 0.9990	 0.2730
K	 0.9020	 0.2530
L	 0.8940	 0.2230
M	 0.8130	 0.2250
N	 0.9080	 0.2070
O	 0.7370	 0.1430
P	 0.7090	 0.1370
Q	 0.9970	 0.2650
R	 0.9910	 0.2570
S	 0.9920	 0.2500
T	 0.9910	 0.2580

