



## wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 11:16 AM EST

PDB ID : 3J0K  
EMDB ID : EMD-5343  
Title : Orientation of RNA polymerase II within the human VP16-Mediator-pol II-TFIIF assembly  
Authors : Bernecky, C.; Grob, P.; Ebmeier, C.C.; Nogales, E.; Taatjes, D.J.  
Deposited on : 2011-10-04  
Resolution : 36.00 Å(reported)  
Based on initial model : 1Y1V

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) : 2.31.3

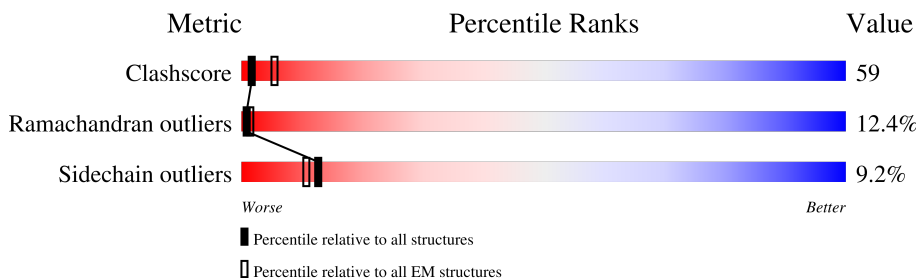
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 36.00 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	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	1455	9% (Upper red bar) 26% (Red) 55% (Yellow) 16% (Orange) .. (Grey)
2	B	1224	21% (Upper red bar) 25% (Red) 52% (Yellow) 13% (Orange) 9% (Grey)
3	C	268	60% (Upper red bar) 26% (Red) 56% (Yellow) 15% (Orange) .. (Grey)
4	D	221	18% (Upper red bar) 31% (Red) 42% (Yellow) 7% (Orange) 20% (Grey)
5	E	215	53% (Upper red bar) 35% (Red) 54% (Yellow) 11% (Orange)
6	F	84	27% (Upper red bar) 29% (Red) 56% (Yellow) 15% (Orange)
7	G	171	21% (Upper red bar) 33% (Red) 59% (Yellow) 8% (Orange)
8	H	146	47% (Upper red bar) 29% (Red) 53% (Yellow) 9% (Orange) 9% (Grey)

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
9	I	122	
10	J	70	
11	K	120	
12	L	46	

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 31137 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1426	11214	7069	1959	2124	62	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1112	8837	5594	1548	1640	55	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	266	2095	1317	348	417	13	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase II 32 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	177	1356	840	241	273	2	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	214	1752	1111	309	321	11	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	84	679	434	115	127	3	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	171	1340	861	222	249	8	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	133	1068	673	180	211	4	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit 9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	I	119	971	596	179	186	10	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	J	65	532	339	93	94	6	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	K	114	919	590	156	171	2	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	L	46	364	224	72	64	4	0	0

- Molecule 13 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
13	A	1	Total 1	Mg 1	0

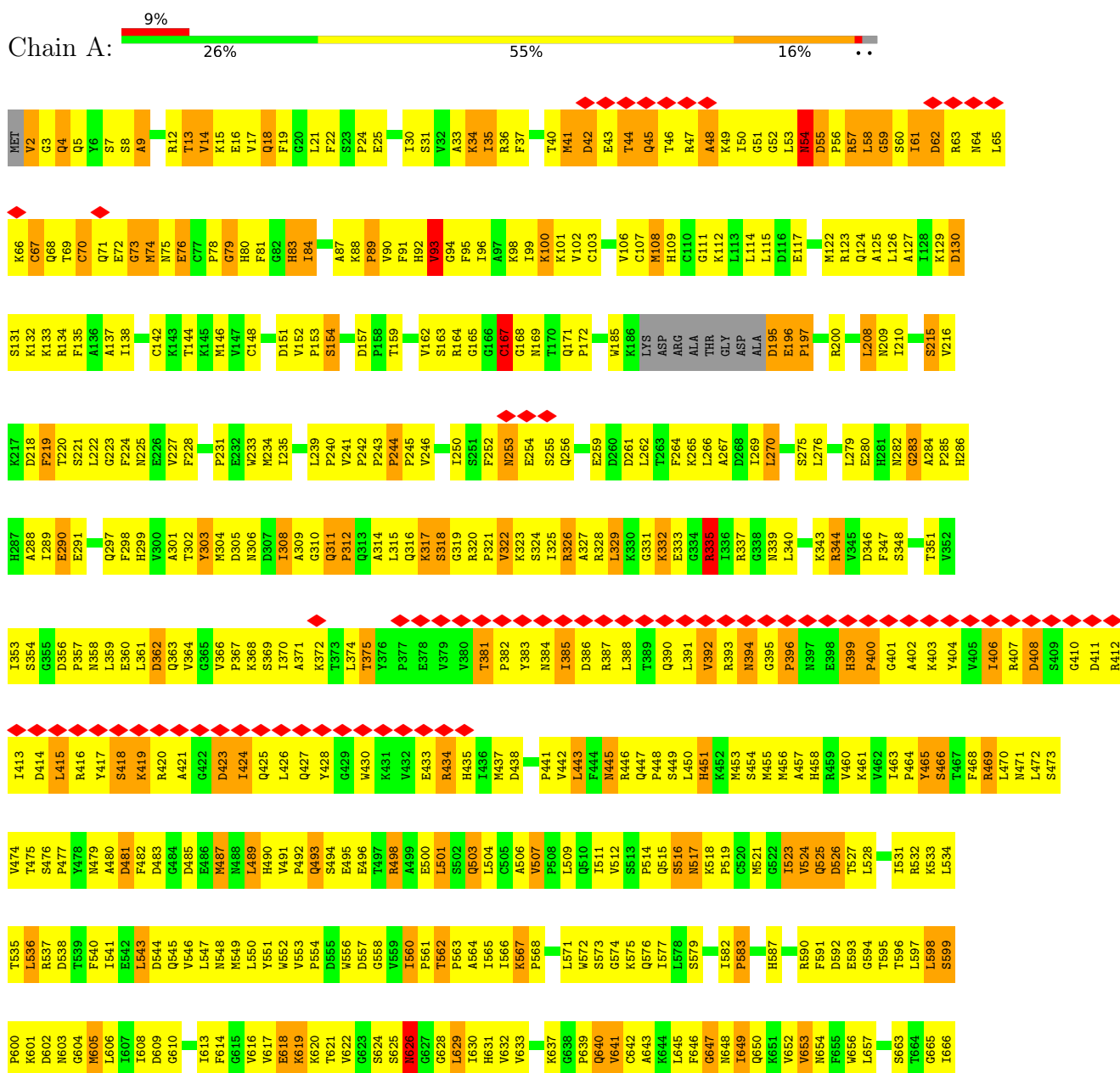
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

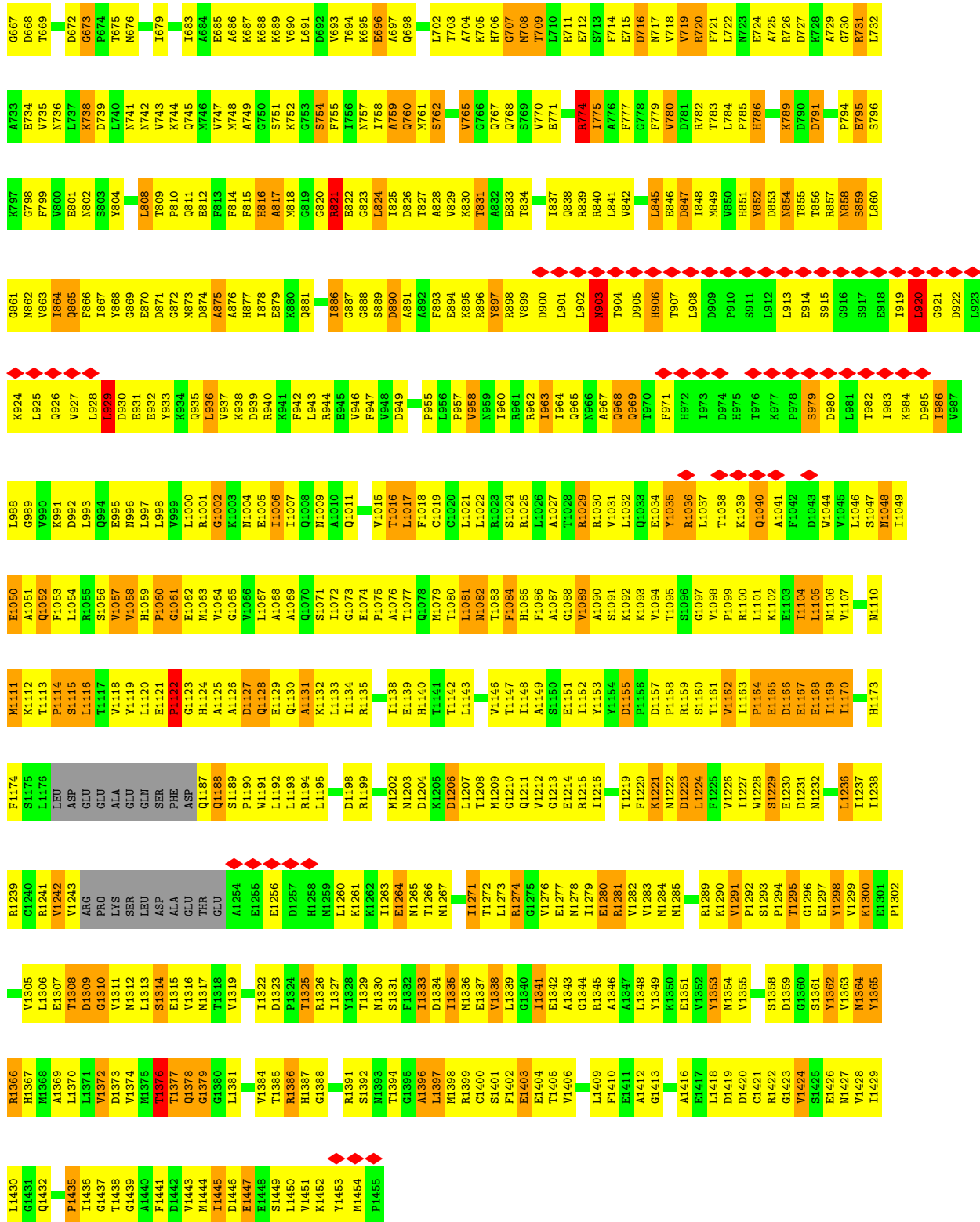
Mol	Chain	Residues	Atoms		AltConf
14	A	3	Total 3	Zn 3	0
14	B	1	Total 1	Zn 1	0
14	C	1	Total 1	Zn 1	0
14	I	2	Total 2	Zn 2	0
14	J	1	Total 1	Zn 1	0
14	L	1	Total 1	Zn 1	0

### 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: DNA-directed RNA polymerase II largest subunit

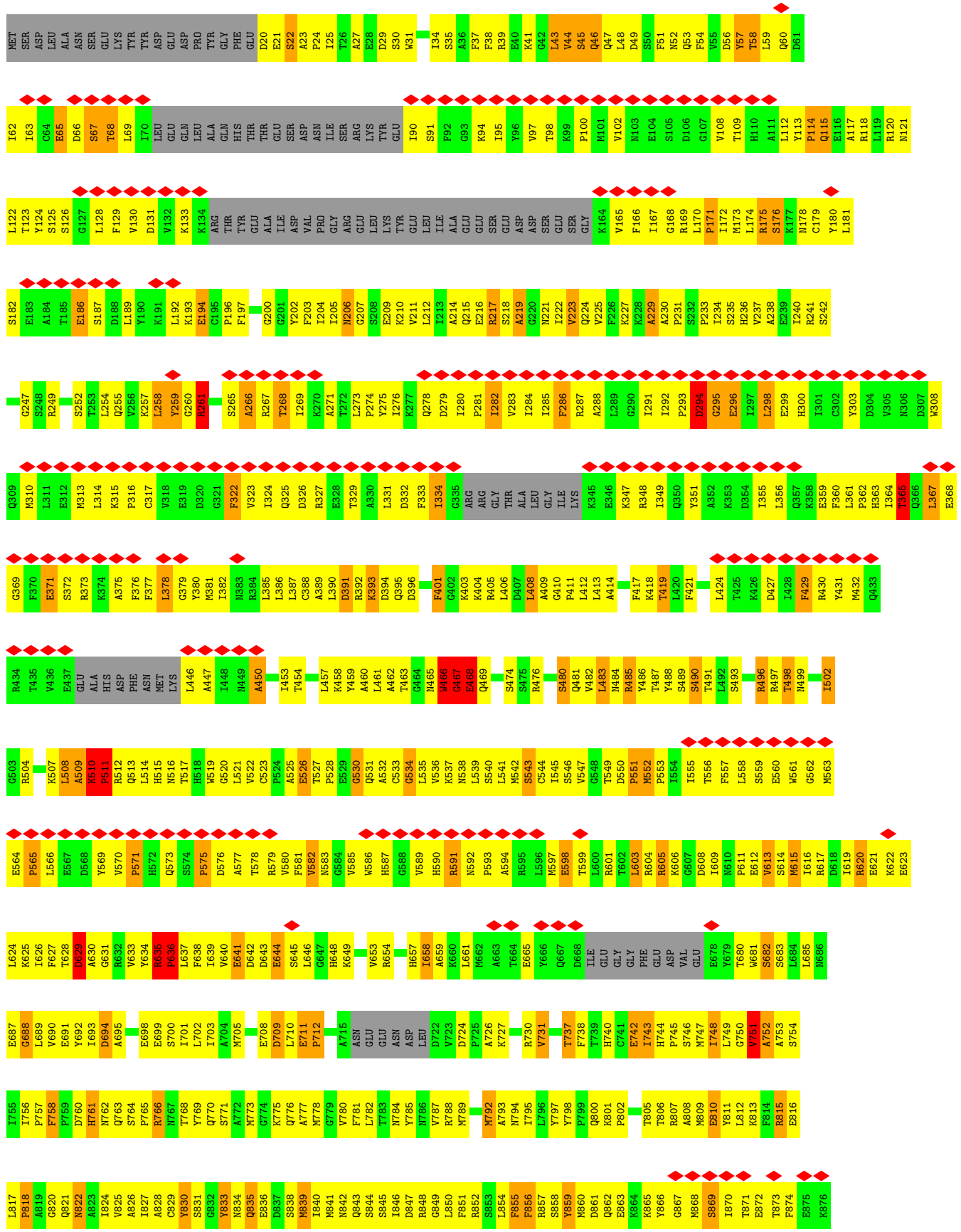


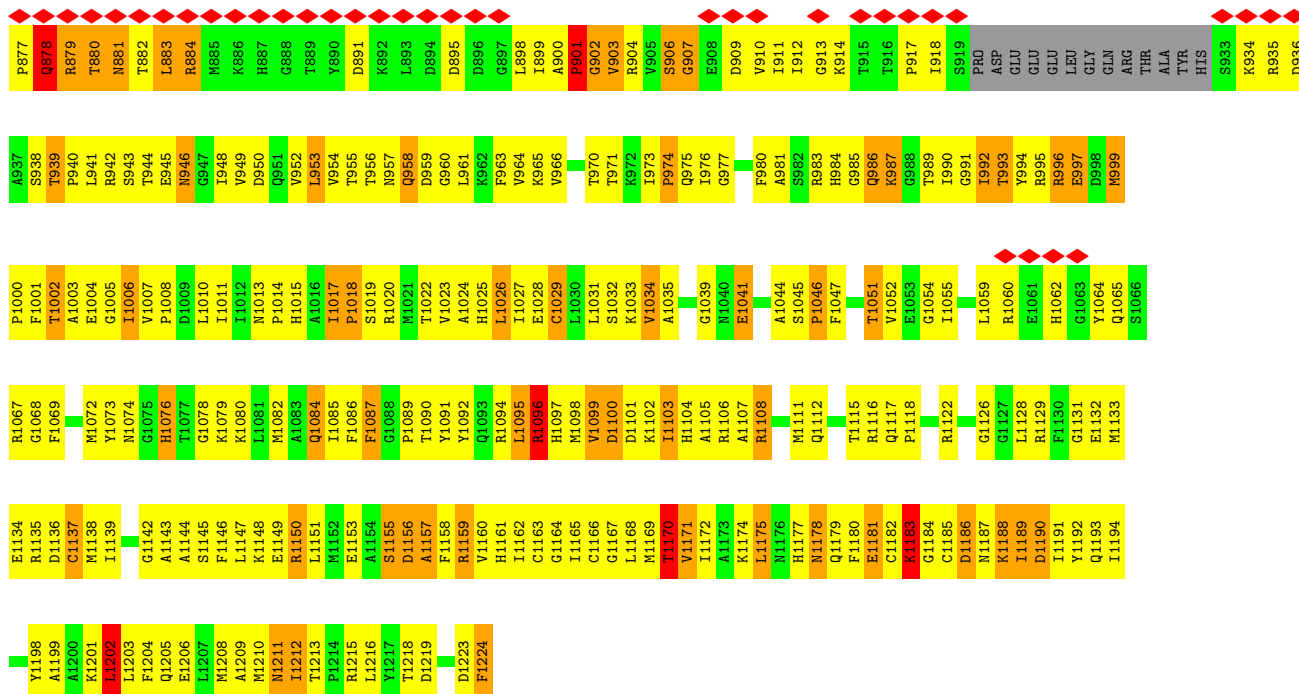


• Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide

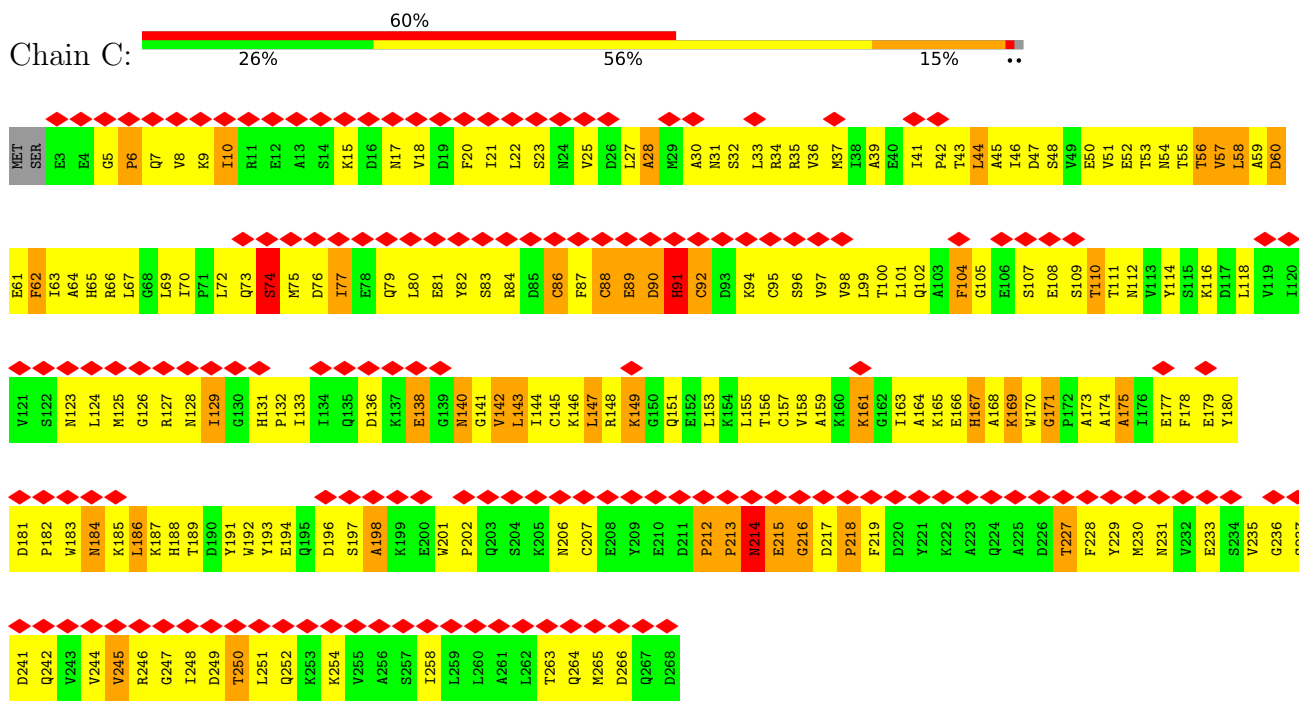




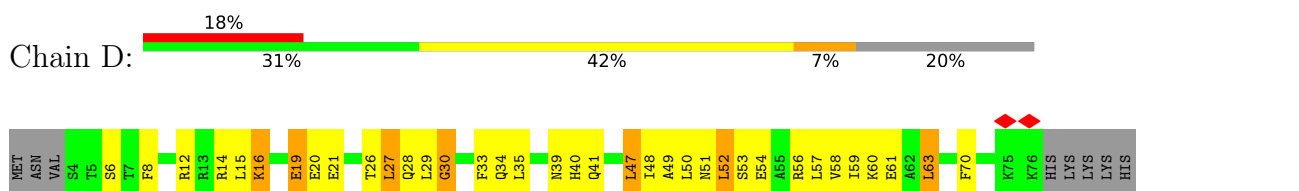




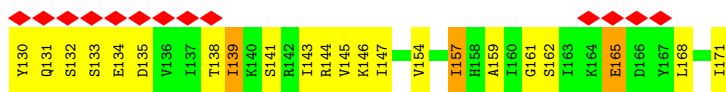
• Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide



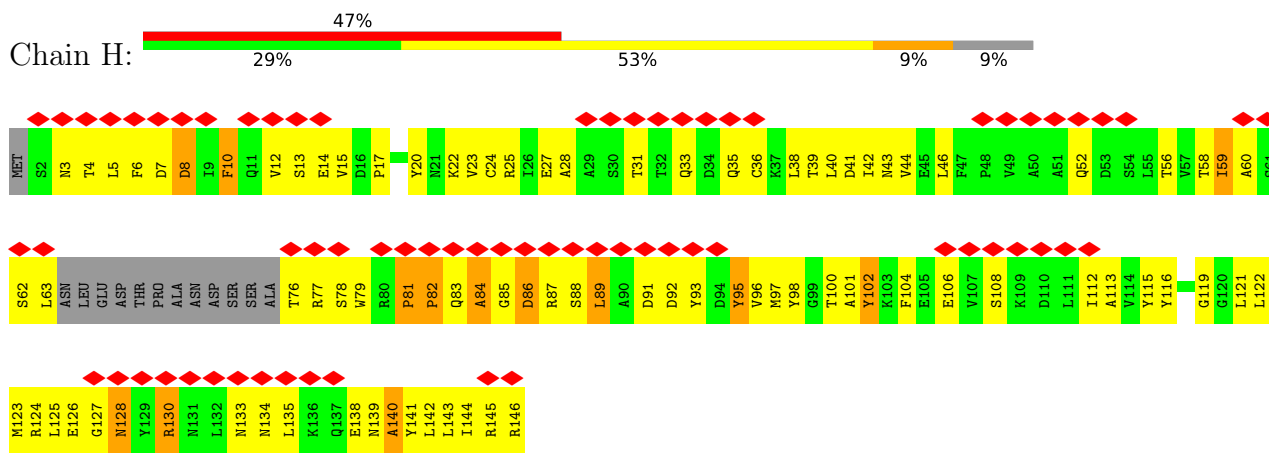
• Molecule 4: DNA-directed RNA polymerase II 32 kDa polypeptide



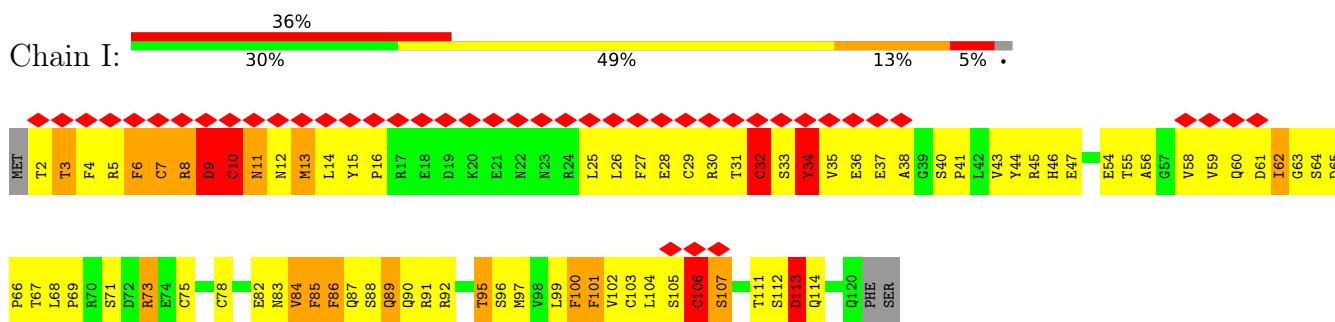




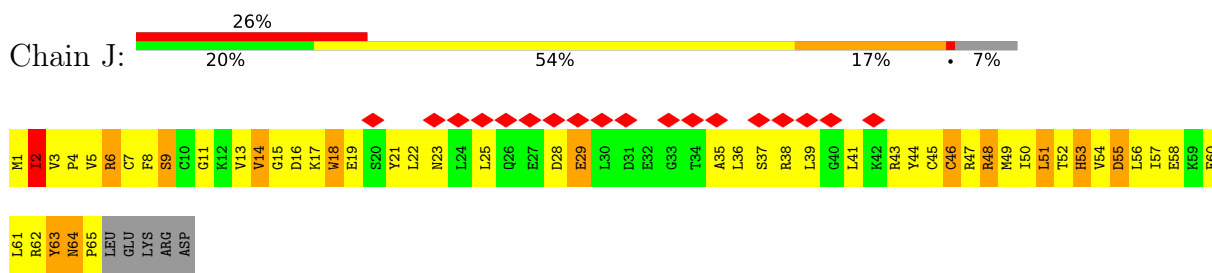
• Molecule 8: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



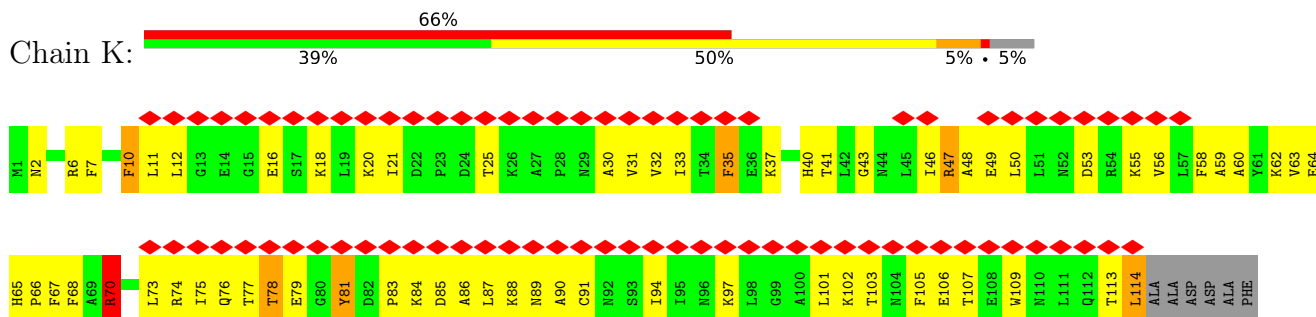
• Molecule 9: DNA-directed RNA polymerase II subunit 9



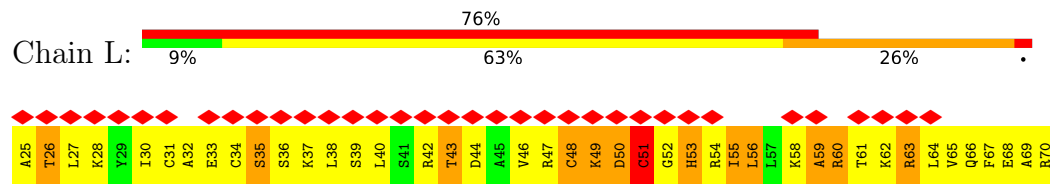
• Molecule 10: DNA-directed RNA polymerases I/II/III subunit 10



• Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	3146	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	each micrograph	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	15	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	29000	Depositor
Image detector	KODAK SO-163 FILM	Depositor
Maximum map value	0.073	Depositor
Minimum map value	-0.009	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.0284	Depositor
Map size ( $\text{\AA}$ )	690.69, 690.69, 690.69	wwPDB
Map dimensions	161, 161, 161	wwPDB
Map angles ( $^\circ$ )	90, 90, 90	wwPDB
Pixel spacing ( $\text{\AA}$ )	4.29, 4.29, 4.29	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, 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	0.50	2/11417 (0.0%)	0.78	4/15442 (0.0%)
2	B	0.52	4/9009 (0.0%)	0.76	8/12146 (0.1%)
3	C	0.48	0/2133	0.77	1/2891 (0.0%)
4	D	0.41	0/1365	0.64	0/1837
5	E	0.43	0/1788	0.66	0/2406
6	F	0.52	0/691	0.77	0/933
7	G	0.49	0/1368	0.72	0/1844
8	H	0.38	0/1086	0.65	1/1470 (0.1%)
9	I	0.46	0/989	0.77	1/1331 (0.1%)
10	J	0.48	0/541	0.75	0/727
11	K	0.45	0/937	0.67	0/1265
12	L	0.54	0/366	0.79	0/485
All	All	0.49	6/31690 (0.0%)	0.75	15/42777 (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	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	467	GLY	C-O	-11.91	1.04	1.23
2	B	468	GLU	CB-CG	8.45	1.68	1.52
2	B	510	LYS	CB-CG	5.58	1.67	1.52
1	A	195	ASP	N-CA	5.49	1.57	1.46
2	B	468	GLU	CG-CD	5.21	1.59	1.51

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	195	ASP	N-CA-C	9.36	136.28	111.00
2	B	510	LYS	CB-CA-C	-7.65	95.10	110.40
2	B	510	LYS	C-N-CD	-7.45	104.21	120.60
3	C	92	CYS	CA-CB-SG	-6.78	101.80	114.00
1	A	1310	GLY	N-CA-C	-6.49	96.88	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	510	LYS	Mainchain
2	B	833	TYR	Sidechain

## 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	11214	0	11281	1475	0
2	B	8837	0	8871	1203	0
3	C	2095	0	2052	263	0
4	D	1356	0	1319	106	0
5	E	1752	0	1776	189	0
6	F	679	0	701	84	0
7	G	1340	0	1357	161	0
8	H	1068	0	1040	120	0
9	I	971	0	929	113	0
10	J	532	0	542	110	0
11	K	919	0	929	97	0
12	L	364	0	387	68	0
13	A	1	0	0	0	0
14	A	3	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	31137	0	31184	3684	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 59.

The worst 5 of 3684 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)
3:C:45:ALA:HA	3:C:72:LEU:HD12	1.28	1.13
4:D:40:HIS:HB3	7:G:73:LYS:HZ3	1.10	1.12
7:G:138:THR:HG22	7:G:139:ILE:H	1.11	1.11
1:A:913:LEU:HD12	1:A:914:GLU:H	1.15	1.11
6:F:82:THR:HG22	6:F:84:TYR:H	1.15	1.11

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	1418/1455 (98%)	914 (64%)	316 (22%)	188 (13%)	0 5
2	B	1096/1224 (90%)	725 (66%)	223 (20%)	148 (14%)	0 4
3	C	264/268 (98%)	169 (64%)	62 (24%)	33 (12%)	0 5
4	D	173/221 (78%)	129 (75%)	27 (16%)	17 (10%)	0 10
5	E	212/215 (99%)	141 (66%)	50 (24%)	21 (10%)	0 9
6	F	82/84 (98%)	60 (73%)	15 (18%)	7 (8%)	1 12
7	G	169/171 (99%)	123 (73%)	34 (20%)	12 (7%)	1 14
8	H	129/146 (88%)	93 (72%)	25 (19%)	11 (8%)	1 12
9	I	117/122 (96%)	80 (68%)	22 (19%)	15 (13%)	0 5

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	J	63/70 (90%)	36 (57%)	14 (22%)	13 (21%)	0	2
11	K	112/120 (93%)	82 (73%)	25 (22%)	5 (4%)	2	22
12	L	44/46 (96%)	18 (41%)	14 (32%)	12 (27%)	0	0
All	All	3879/4142 (94%)	2570 (66%)	827 (21%)	482 (12%)	1	5

5 of 482 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	48	ALA
1	A	55	ASP
1	A	58	LEU
1	A	62	ASP

### 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	1246/1270 (98%)	1133 (91%)	113 (9%)	9	29
2	B	964/1061 (91%)	880 (91%)	84 (9%)	10	31
3	C	234/236 (99%)	205 (88%)	29 (12%)	4	19
4	D	140/200 (70%)	126 (90%)	14 (10%)	7	26
5	E	196/197 (100%)	184 (94%)	12 (6%)	18	44
6	F	74/74 (100%)	63 (85%)	11 (15%)	3	15
7	G	152/152 (100%)	143 (94%)	9 (6%)	19	45
8	H	117/128 (91%)	110 (94%)	7 (6%)	19	44
9	I	113/116 (97%)	97 (86%)	16 (14%)	3	16
10	J	60/65 (92%)	55 (92%)	5 (8%)	11	34
11	K	99/102 (97%)	91 (92%)	8 (8%)	11	35
12	L	40/40 (100%)	33 (82%)	7 (18%)	2	11
All	All	3435/3641 (94%)	3120 (91%)	315 (9%)	13	29

5 of 315 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	148	LEU
9	I	46	HIS
4	D	197	SER
6	F	143	PHE
11	K	10	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 88 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1179	GLN
5	E	101	GLN
3	C	65	HIS
3	C	252	GLN
5	E	147	HIS

### 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 [i](#)

Of 10 ligands modelled in this entry, 10 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-5343. 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](#)

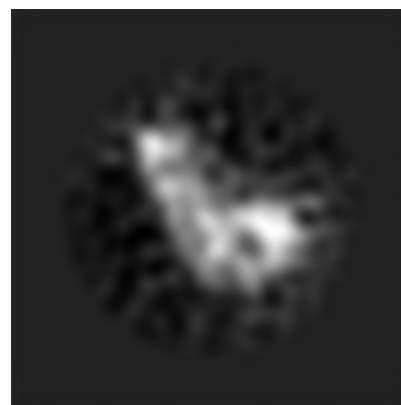
#### 6.1.1 Primary map



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: 80



Y Index: 80



Z Index: 80

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

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

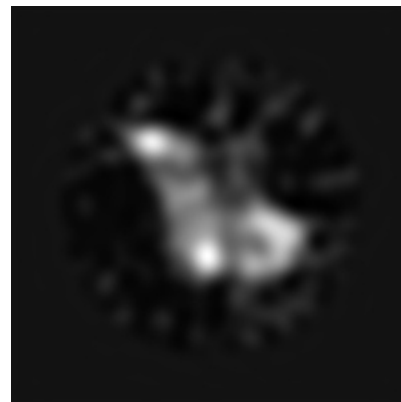
### 6.3.1 Primary map



X Index: 104



Y Index: 71



Z Index: 76

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

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

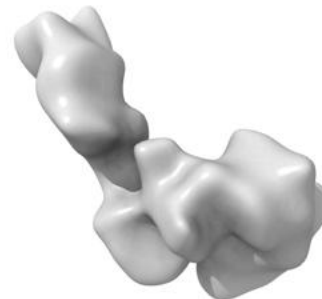
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0284. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

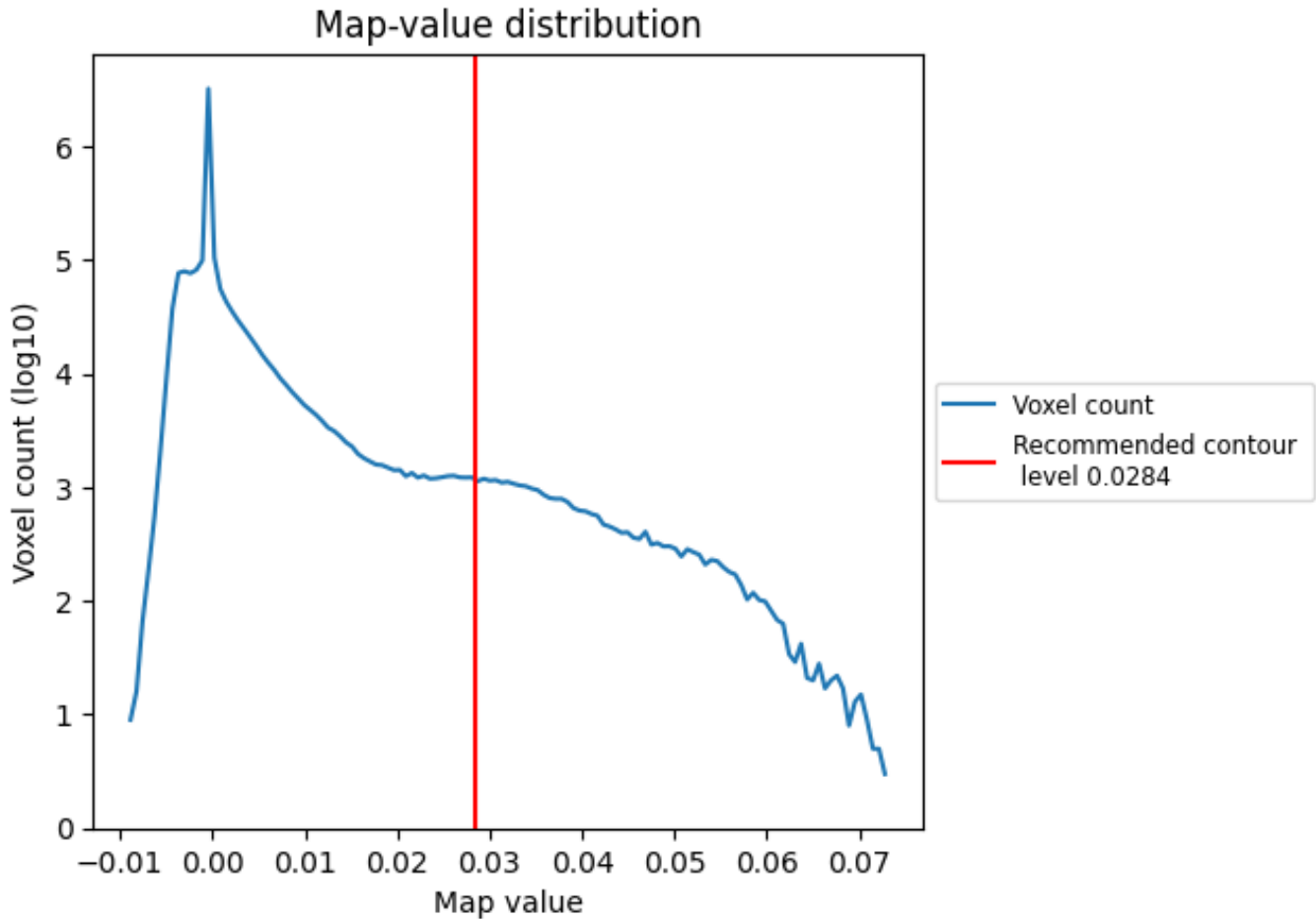
## 6.5 Mask visualisation

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.

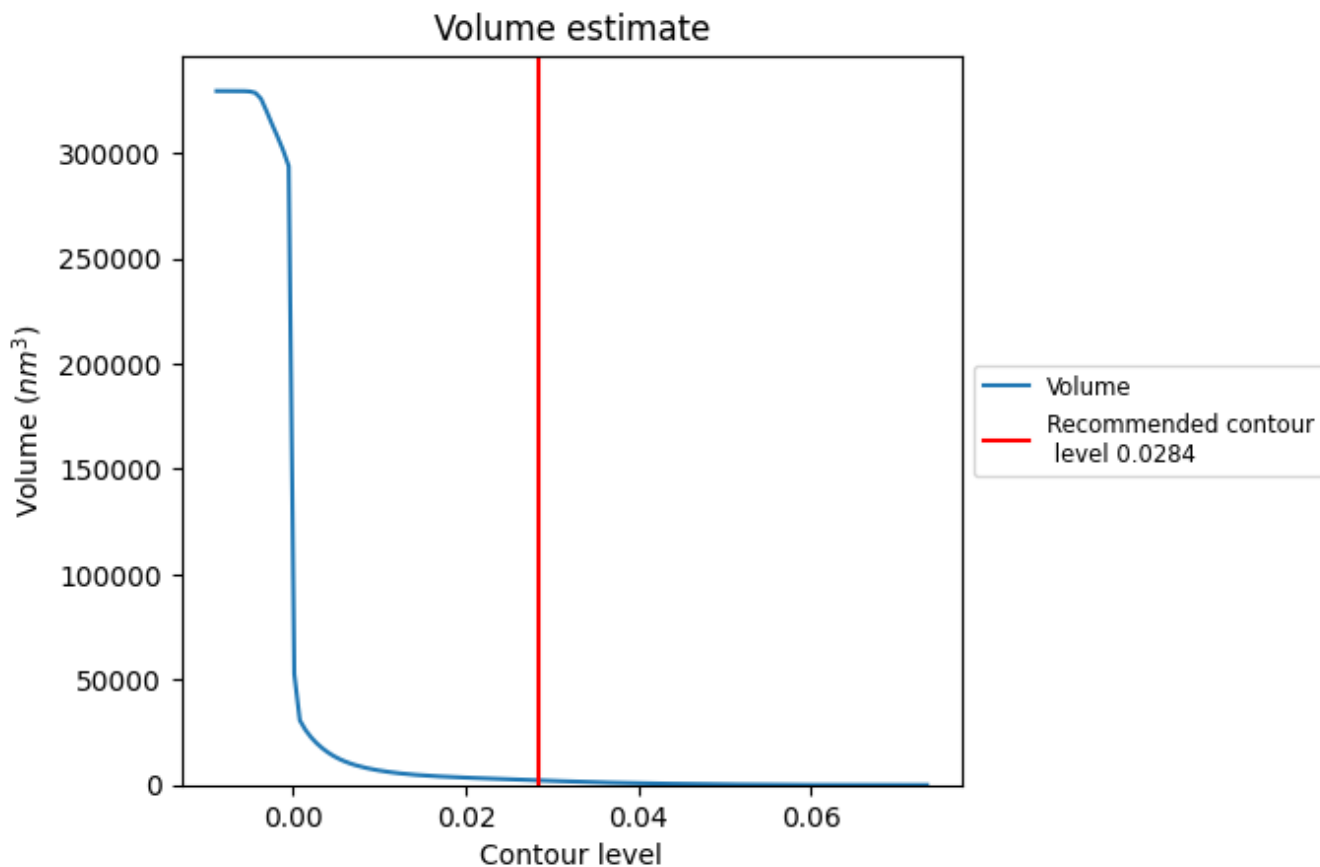
### 7.1 Map-value distribution [i](#)



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.



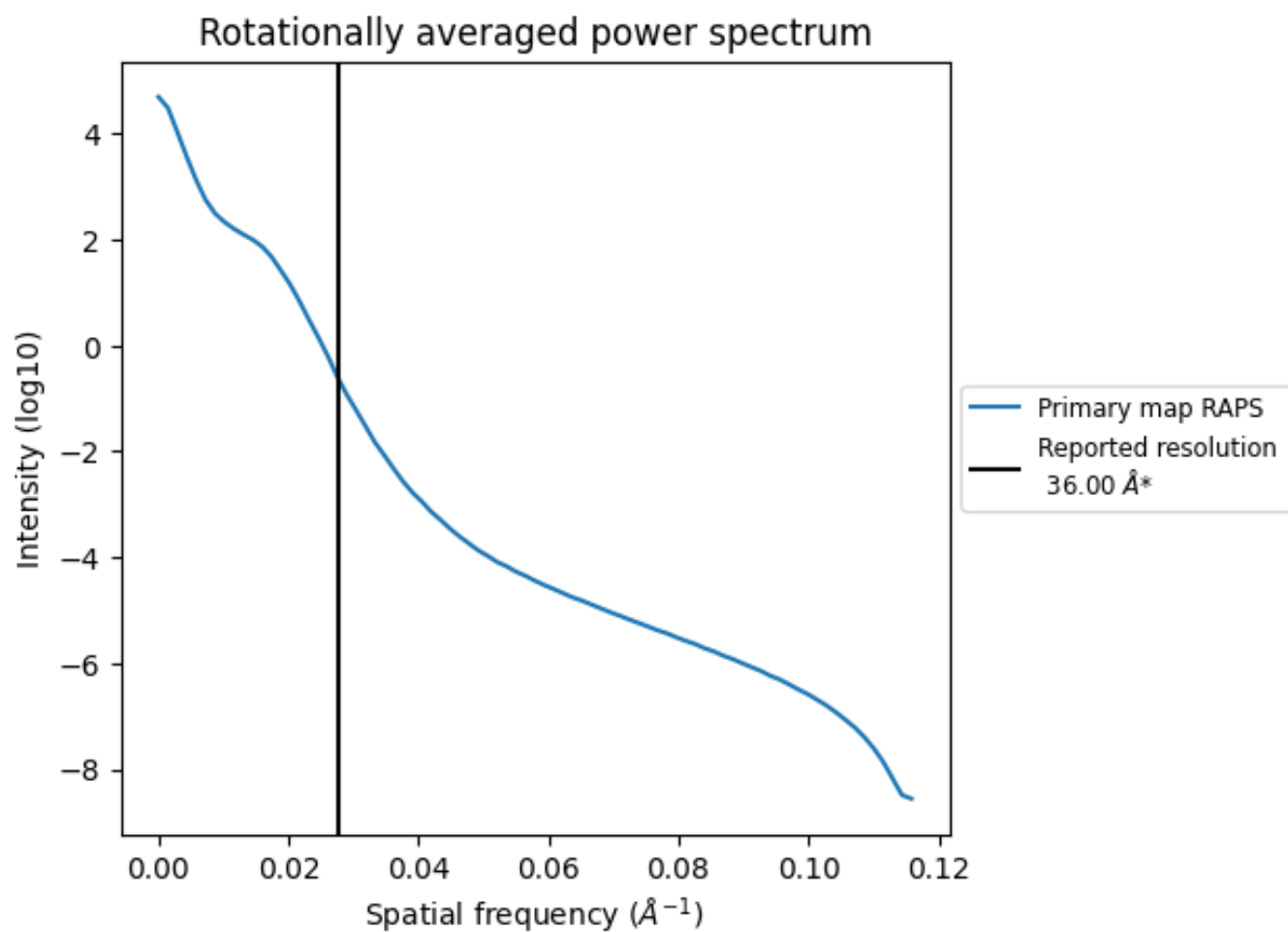
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 2188  $\text{nm}^3$ ; this corresponds to an approximate mass of 1977 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 0.028 Å<sup>-1</sup>

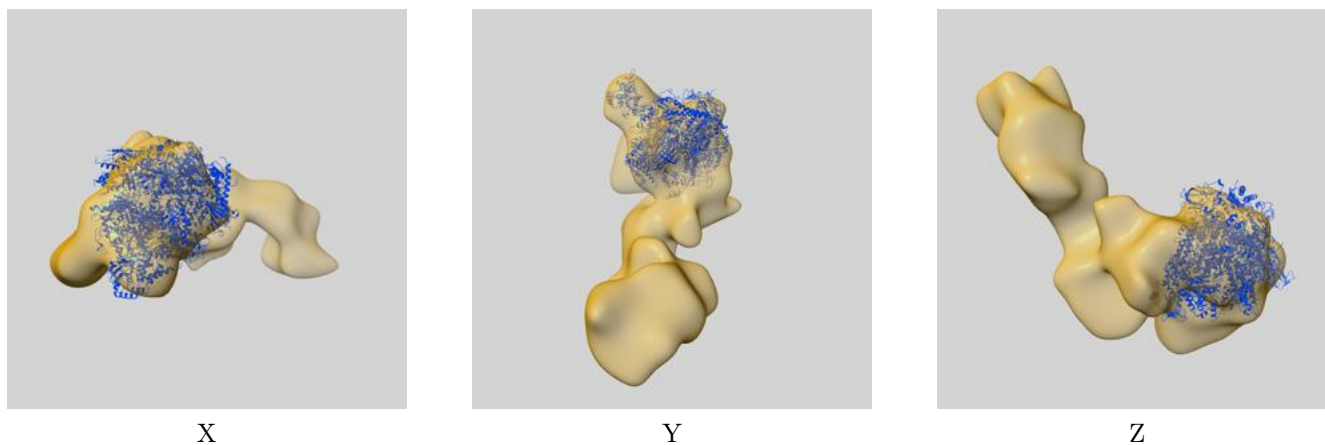
## 8 Fourier-Shell correlation

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

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

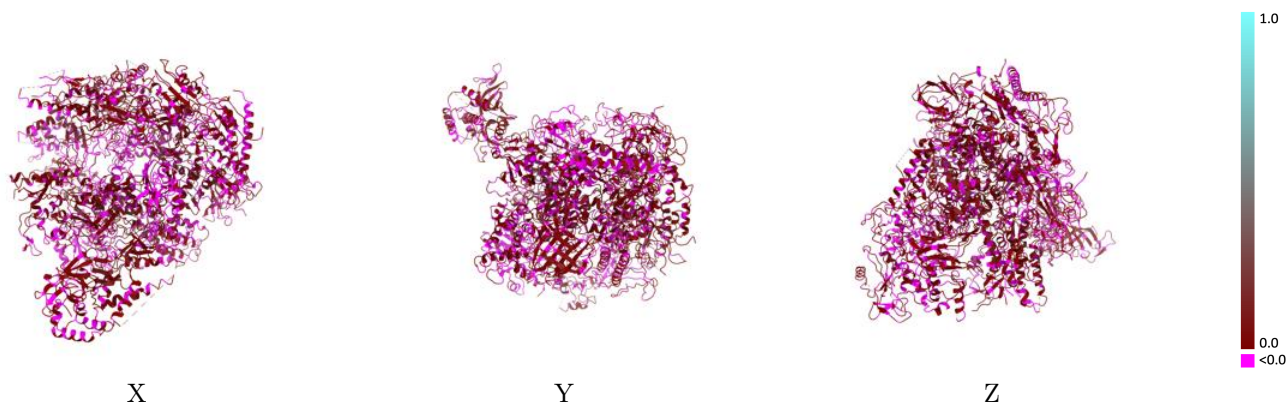
This section contains information regarding the fit between EMDB map EMD-5343 and PDB model 3J0K. Per-residue inclusion information can be found in section 3 on page 7.

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



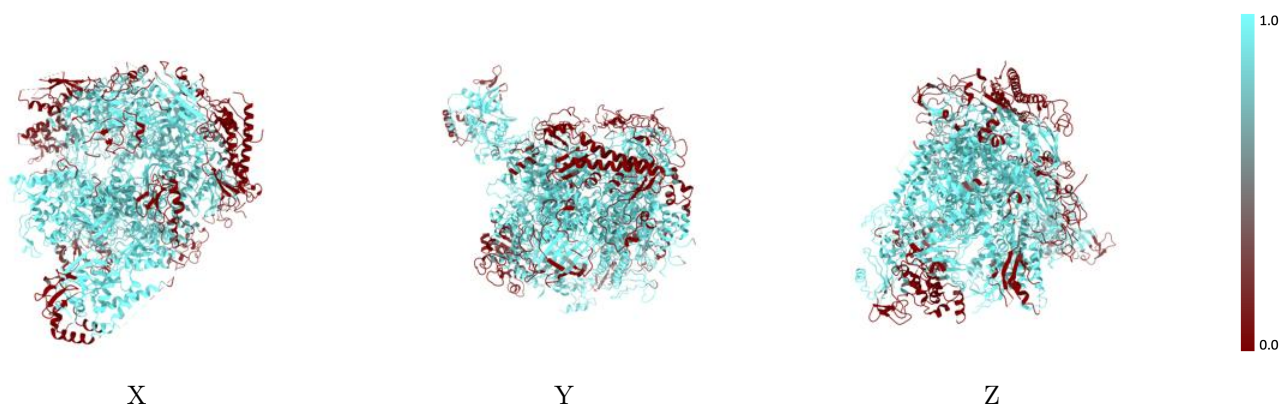
The images above show the 3D surface view of the map at the recommended contour level 0.0284 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



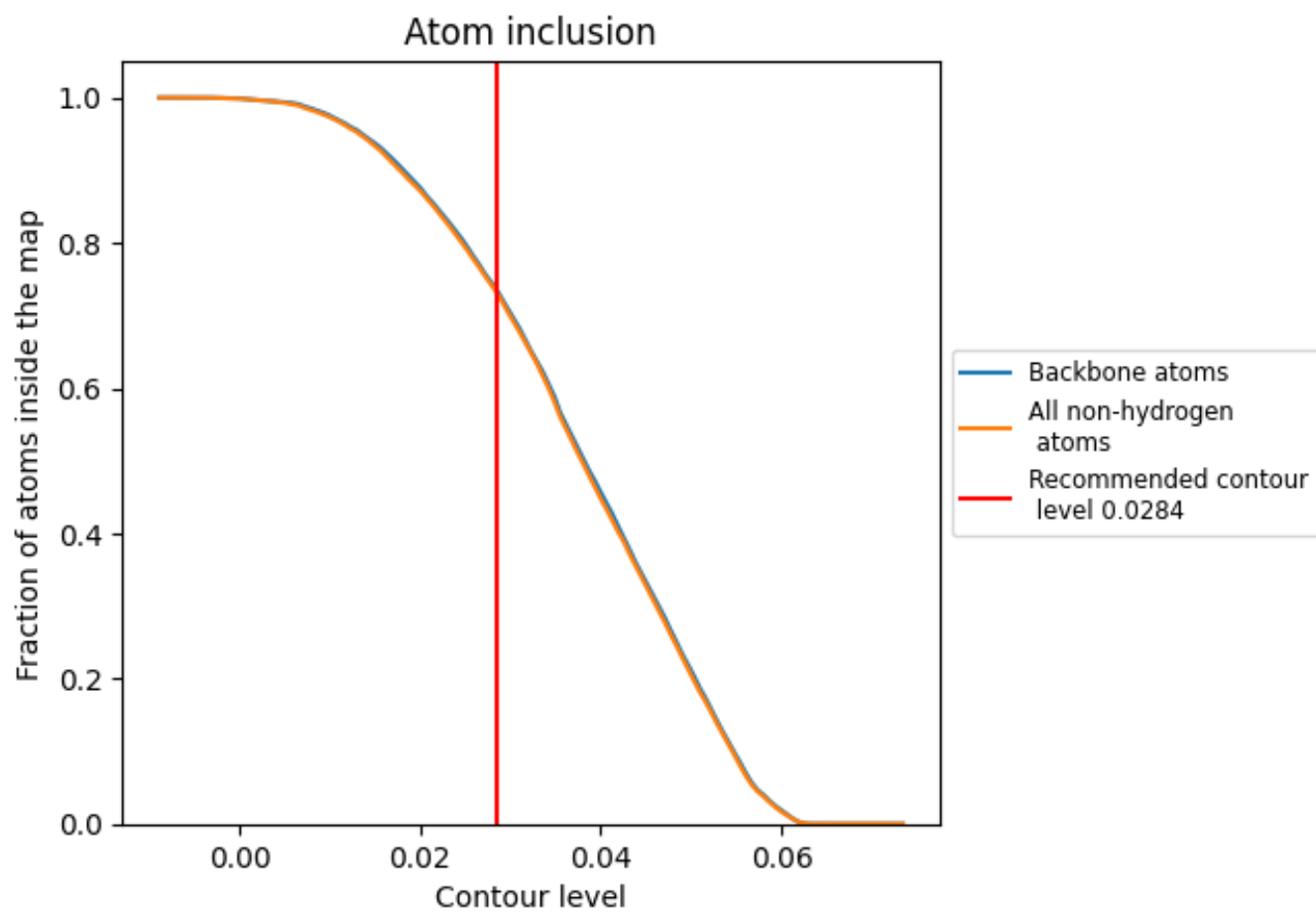
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0284).



























## 9.4 Atom inclusion [i](#)



At the recommended contour level, 74% of all backbone atoms, 73% 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 (0.0284) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7332	 0.0260
A	 0.8979	 0.0290
B	 0.7606	 0.0340
C	 0.3733	 0.0210
D	 0.7522	 0.0200
E	 0.4558	 -0.0130
F	 0.7489	 0.0060
G	 0.7882	 0.0470
H	 0.4846	 0.0300
I	 0.6205	 -0.0000
J	 0.7079	 0.0350
K	 0.3024	 0.0000
L	 0.2443	 0.0090

