



## wwPDB EM Validation Summary Report ⓘ

Nov 20, 2022 – 02:56 am GMT

PDB ID : 6GMH  
EMDB ID : EMD-0031  
Title : Structure of activated transcription complex Pol II-DSIF-PAF-SPT6  
Authors : Vos, S.M.; Farnung, L.; Boehing, M.; Linden, A.; Wigge, C.; Urlaub, H.;  
Cramer, P.  
Deposited on : 2018-05-26  
Resolution : 3.10 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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:

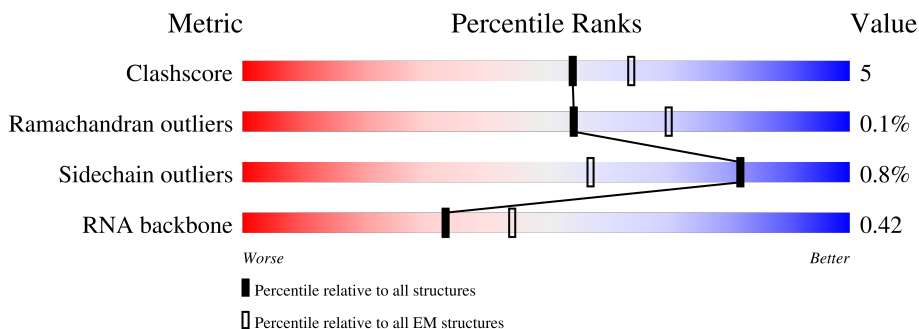
EMDB validation analysis : 0.0.1.dev43  
Mogul : 1.8.4, CSD as541be (2020)  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et 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 3.10 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.








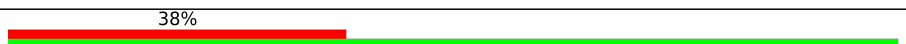

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	1970	
2	B	1174	
3	C	275	
4	D	142	
5	E	210	
6	F	127	
7	G	172	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	150	 89% 10%
9	I	125	 85% 9% 6%
10	J	67	 84% 16%
11	K	117	 83% 15%
12	L	58	 71% 9% 21%
13	M	1726	 44% 56% 42%
14	N	48	 10% 71% 6% 23%
15	P	46	 11% 17% 22% 7% 54%
16	Q	1178	 50% 72% 25%
17	T	48	 12% 75% 23%
18	U	776	 8% 14% 86%
19	V	613	 6% 13% 87%
20	W	305	 56% 88% 10%
21	X	16	 38% 100%
22	Y	121	 76% 87% 9%
23	Z	1087	 29% 39% 6% 55%

## 2 Entry composition [i](#)

There are 25 unique types of molecules in this entry. The entry contains 50239 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 RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	N	O	P			S
1	A	1441	11371	7151	2033	2115	2	70	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1132	9052	5725	1592	1671	64	0	0

- Molecule 3 is a protein called RNA polymerase II subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	263	2115	1324	365	420	6	0	0

- Molecule 4 is a protein called RNA polymerase II subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	126	1004	630	170	200	4	0	0

- Molecule 5 is a protein called RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	209	1720	1089	300	323	8	0	0

- Molecule 6 is a protein called RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	82	657	418	113	121	5	0	0

- Molecule 7 is a protein called RNA polymerase II subunit G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	G	171	1333	866	214	245	8	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	H	148	1186	750	194	237	5	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	I	117	949	587	169	182	11	0	0

- Molecule 10 is a protein called RPB10.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	J	67	533	345	90	92	6	0	0

- Molecule 11 is a protein called RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	K	115	920	593	152	173	2	0	0

- Molecule 12 is a protein called RPB12.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	L	46	388	241	75	66	6	0	0

- Molecule 13 is a protein called Transcription elongation factor SPT6, Transcription elongation factor SPT6, Transcription elongation factor SPT6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	M	1002	4737	2583	1071	1076	7	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-2	SER	-	expression tag	UNP Q7KZ85
M	-1	ASN	-	expression tag	UNP Q7KZ85
M	0	ALA	-	expression tag	UNP Q7KZ85

- Molecule 14 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
14	N	37	773	361	158	217	37	0	0

- Molecule 15 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
15	P	21	452	202	87	142	21	0	0

- Molecule 16 is a protein called CTR9,RNA polymerase-associated protein CTR9 homolog,RNA polymerase-associated protein CTR9 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	Q	884	4116	2199	953	961	3	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	1174	GLU	-	expression tag	UNP Q6PD62
Q	1175	ASN	-	expression tag	UNP Q6PD62
Q	1176	LEU	-	expression tag	UNP Q6PD62
Q	1177	TYR	-	expression tag	UNP Q6PD62
Q	1178	GLN	-	expression tag	UNP Q6PD62

- Molecule 17 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
17	T	48	974	462	168	296	48	0	0

- Molecule 18 is a protein called LEO1,LEO1,RNA polymerase-associated protein LEO1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
18	U	110	440	220	110	110	0	0

- Molecule 19 is a protein called PAF1,RNA polymerase II-associated factor 1 homolog,RNA polymerase II-associated factor 1 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
19	V	81	324	162	81	81	0	0

- Molecule 20 is a protein called WD repeat-containing protein 61.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	W	300	2333	1483	392	454	4	0	0

- Molecule 21 is a protein called CDC73.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
21	X	16	63	32	16	15	0	0

- Molecule 22 is a protein called Transcription elongation factor SPT4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	Y	116	911	570	159	173	9	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Y	-3	GLY	-	expression tag	UNP P63272
Y	-2	PRO	-	expression tag	UNP P63272
Y	-1	GLY	-	expression tag	UNP P63272
Y	0	SER	-	expression tag	UNP P63272

- Molecule 23 is a protein called Transcription elongation factor SPT5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	Z	486	3878	2465	684	712	17	0	0

- Molecule 24 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 25 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

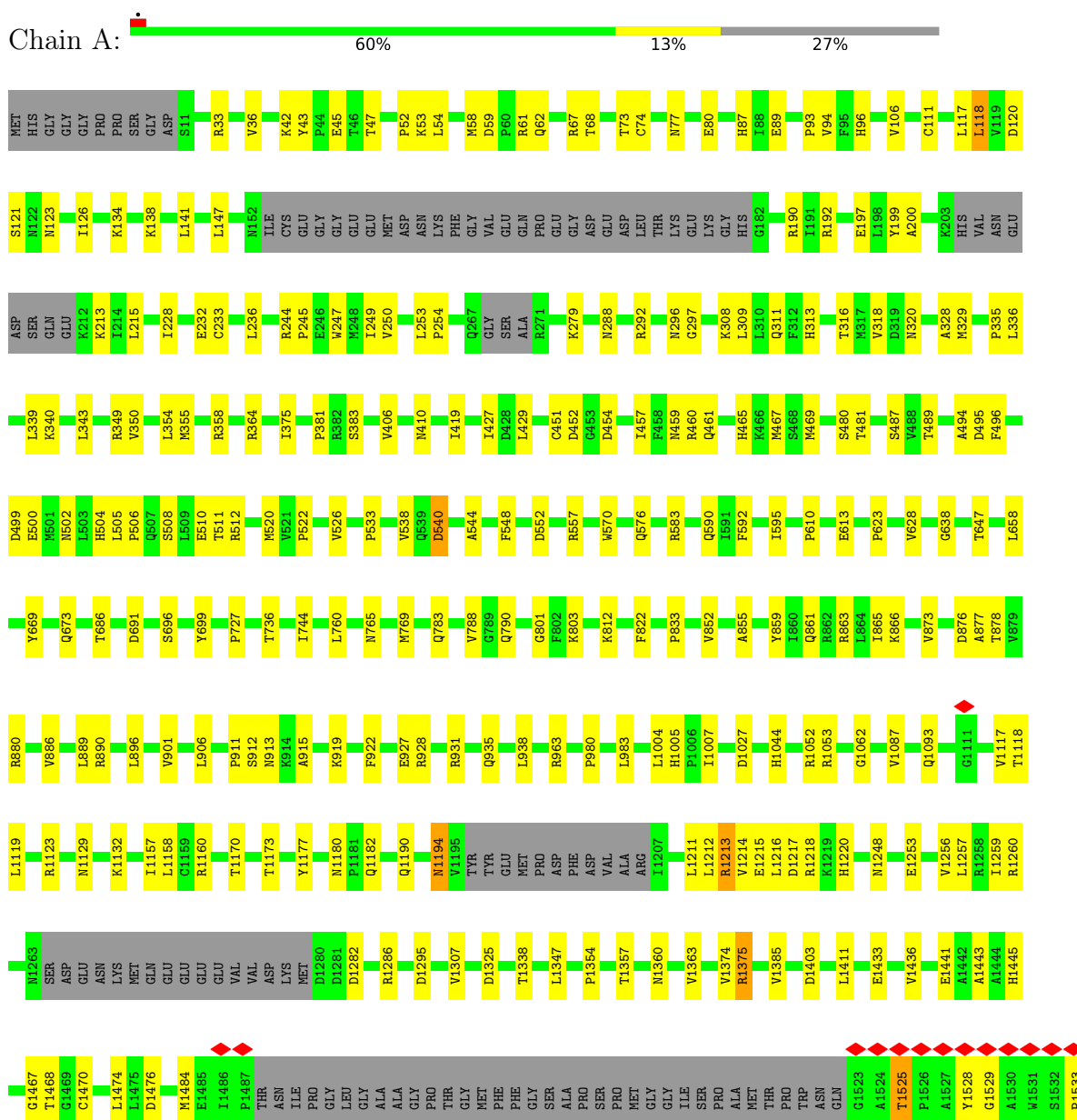
Mol	Chain	Residues	Atoms		AltConf
25	A	1	Total 1	Mg 1	0



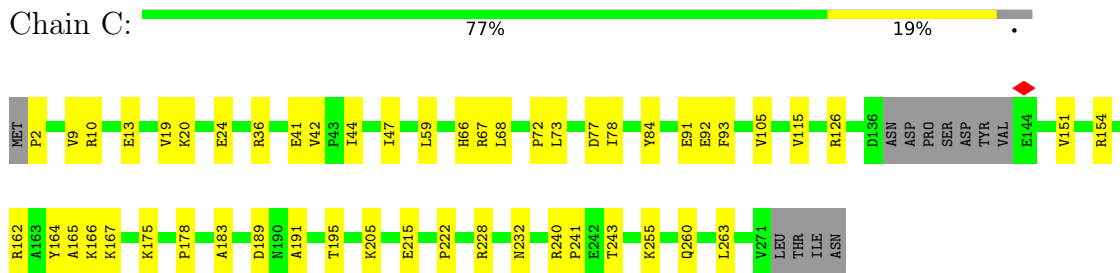
### 3 Residue-property plots

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.

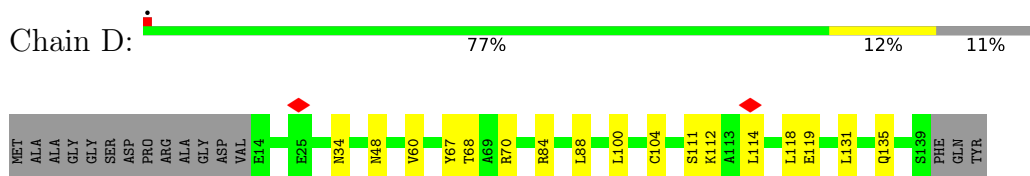
#### • Molecule 1: RPB1



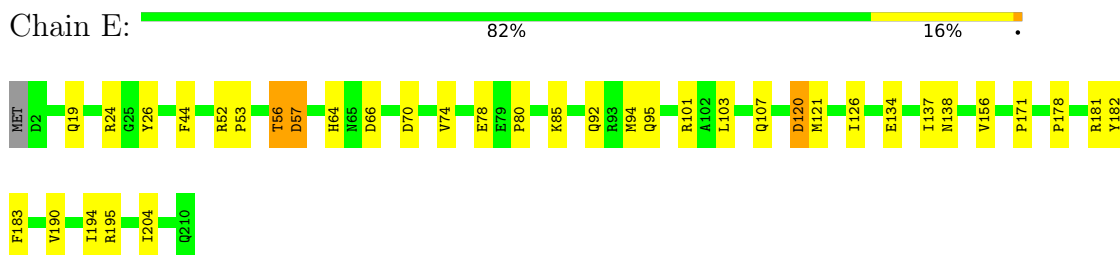




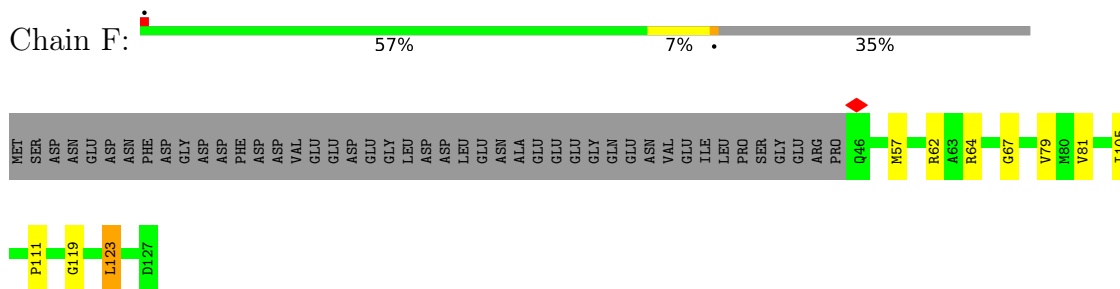
• Molecule 4: RNA polymerase II subunit D



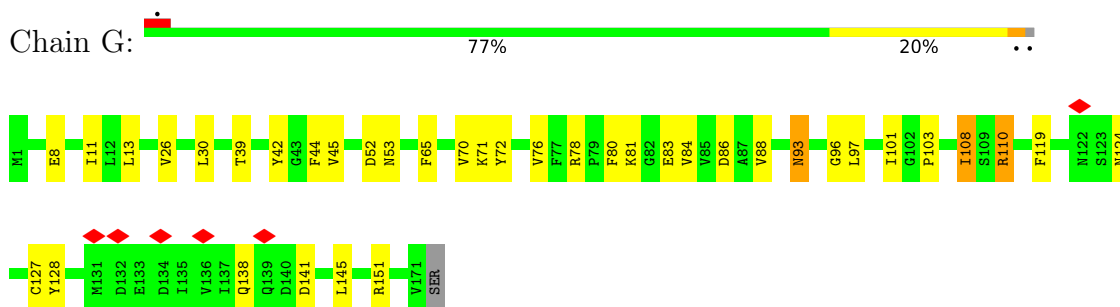
• Molecule 5: RNA polymerase II subunit E



• Molecule 6: RNA polymerase II subunit F

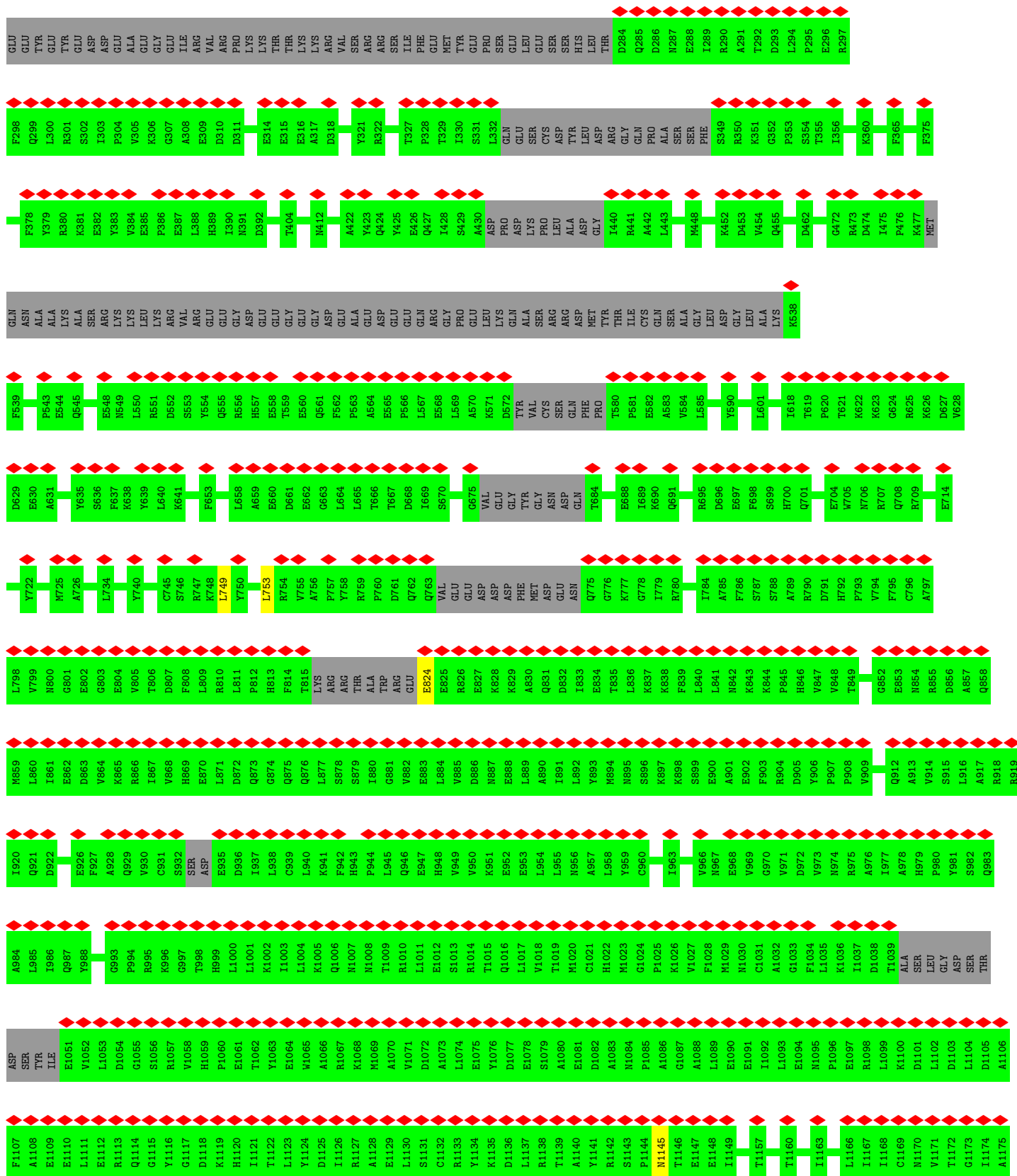


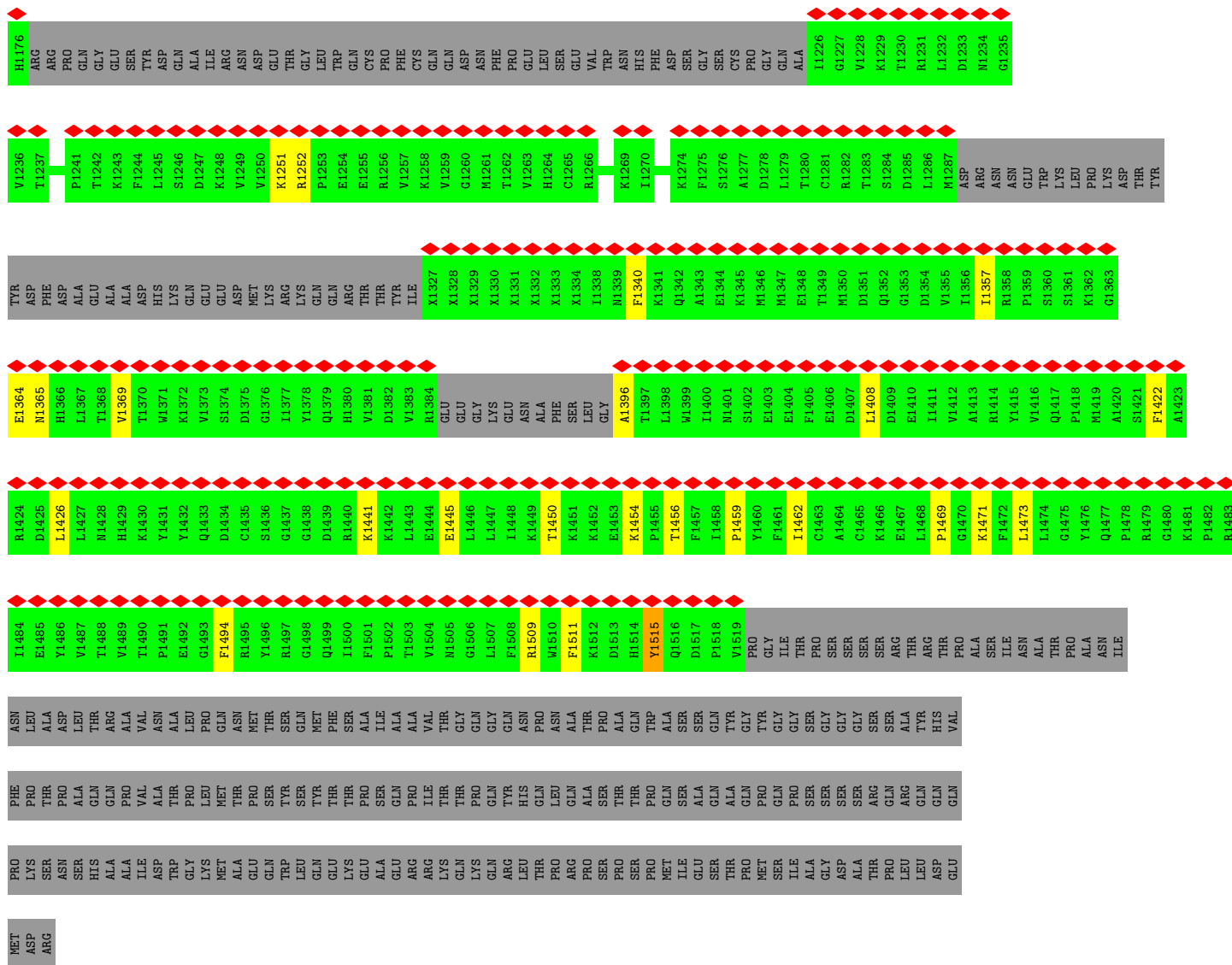
• Molecule 7: RNA polymerase II subunit G



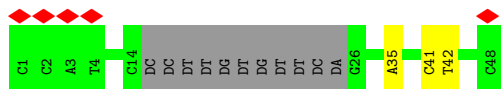
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



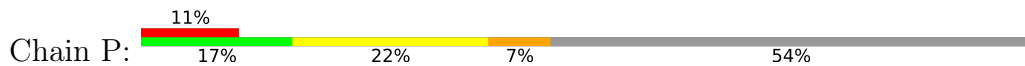




• Molecule 14: Non-template DNA



• Molecule 15: RNA



• Molecule 16: CTR9, RNA polymerase-associated protein CTR9 homolog, RNA polymerase-associated protein CTR9 homolog











VAL	THR	GLU	SER	THR	GLY	LYS	GLU	R416	E417	H418	N419	F420	Q421	P422	G423	D424	N425	V426	E427	V428	C429	E430	G431	E432	L433	I434	N435	L436	Q437	G438	K439	I440	L441	S442	V443	D444	G445	N446	K447	I448	T449	I450	M451	P452	K453	H454	E455	D456	L457	K458	D459	M460	L461	E462	F463	P464	A465	Q466	E467	
L468	R469	K470	Y471	F472	K473	D476	K479	V480	I481	A482	G483	R484	F485	E486	G487	D488	T489	G490	L491	E496	E497	L502	D505	L506	H509	E510	L511	R516	D517	L518	Q519	L520	C521	S522	E523	T524	A525	S526	G527	V528	D529	V530	G531	G532	Q533	H534	E535	W536	G537	E538	L539									
L542	D543	P544	Q545	T546	V549	I550	V551	R552	L553	E554	R555	E556	T557	F558	Q559	Y564	G565	K566	V567	V568	T569	V570	R571	H572	Q573	A574	V575	T576	R577	K578	K579	D580	N581	R582	F583	A584	V585	A586	L587	D588	S589	E590	Q591	N592	N593	I594	H595	V596	K597	D598	I599	V600	K601	V602	I603	G605				
P606	H607	S608	G609	R610	E611	G612	E613	I614	R615	H616	L617	F618	R619	S620	F621	A622	F623	L624	H625	C626	K627	K628	L629	V630	E631	N632	G633	G634	M635	F636	K639	T640	R641	H642	L643	V644	L645	A646	GLY	GLY	SER	LYS	PRO	ARG	ASP	THR	PRO	ASN	THR	VAL	THR	PHE	THR	VAL	GLY	PHE	ALA	PRO	MET	SER
PRO	ARG	ILE	SER	GLY	MET	HIS	PRO	ALA	GLY	GLN	ARG	GLY	PHE	GLY	SER	PRO	GLY	GLY	SER	GLY	MET	GLY	ARG	ARG	GLY	ASP	N703	R711	Q714	K725	R749	G754	SER	ARG	ARG	PRO	GLY	GLY	TRP	LYS	PRO	THR	PRO	ASN	THR	VAL	THR	THR	VAL	GLY	ALA	PRO	MET	SER						
THR	PRO	MET	TYR	GLY	GLN	THR	PRO	TYR	TYR	GLY	GLY	GLY	THR	PRO	TYR	PRO	GLY	GLY	THR	GLN	ASP	GLY	ARG	GLY	ARG	PRO	GLY	THR	PRO	LEU	HIS	ASP	GLY	SER	VAL	GLN	THR	PRO	ALA	GLN	SER	GLY	ALA	THR	TRP	ASP	PRO	ASN	THR	TYR	ASN	PRO	ASN	THR						
THR	PRO	ARG	ALA	GLY	GLU	THR	PRO	TYR	GLY	ASP	ASP	TYR	THR	THR	PRO	SER	PRO	GLY	ALA	TYR	GLY	THR	GLY	THR	GLY	THR	PRO	GLY	PRO	LEU	VAL	ASP	PRO	THR	THR	ASN	PRO	ALA	THR	THR	PRO	GLY	GLY	ALA	THR	TRP	ASP	PRO	ASN	THR	TYR	ASN	PRO	ASN	THR					
ASP	GLN	PHE	SER	PRO	TYR	ALA	PRO	PRO	GLN	SER	TYR	GLN	PRO	PRO	HIS	PRO	PRO	PRO	GLN	VAL	VAL	ASN	PRO	ASP	ALA	ALA	TRP	VAL	VAL	THR	GLN	HIS	SER	PRO	PRO	ALA	THR	THR	TYR	THR	PRO	GLY	GLY	ALA	THR	TRP	ASP	PRO	ASN	THR	TYR	ASN	PRO	ASN	THR					
VAL	GLY	TYR	SER	PRO	CYS	PRO	MET	THR	PRO	GLY	PRO	GLY	TYR	VAL	ASN	PRO	HIS	THR	THR	PRO	PRO	GLY	SER	GLY	LEU	ILE	GLU	GLN	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	
THR	GLY	GLY	MET	GLY	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL		
ASP	LEU	ASP	GLU	LEU	LYS	ARG	PHE	LEU	GLY	LYS	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL	VAL

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	374964	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 ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.155	Depositor
Minimum map value	-0.057	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.0157	Depositor
Map size (Å)	377.64, 377.64, 377.64	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.049, 1.049, 1.049	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: TPO, MG, SEP, 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	0.38	0/11558	0.63	3/15601 (0.0%)
2	B	0.41	0/9232	0.61	2/12462 (0.0%)
3	C	0.42	0/2158	0.58	0/2931
4	D	0.31	0/1017	0.55	0/1368
5	E	0.36	0/1751	0.65	1/2366 (0.0%)
6	F	0.35	0/667	0.55	0/901
7	G	0.34	0/1364	0.64	0/1853
8	H	0.41	0/1207	0.65	0/1628
9	I	0.35	0/972	0.56	0/1316
10	J	0.42	0/542	0.56	0/730
11	K	0.40	0/939	0.60	2/1271 (0.2%)
12	L	0.41	0/394	0.65	0/524
13	M	0.26	0/4724	0.49	0/6031
14	N	0.59	0/870	0.88	0/1341
15	P	0.44	0/506	1.08	5/787 (0.6%)
16	Q	0.25	0/2923	0.44	0/3746
17	T	0.64	0/1087	0.96	1/1674 (0.1%)
20	W	0.27	0/2392	0.55	0/3257
22	Y	0.26	0/927	0.48	0/1250
23	Z	0.28	0/3946	0.53	0/5314
All	All	0.37	0/49176	0.61	14/66351 (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
1	A	0	2
2	B	0	3
3	C	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
5	E	0	2
7	G	0	1
13	M	0	3
All	All	0	12

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	120	ASP	CB-CG-OD1	8.67	126.11	118.30
11	K	80	ASP	CB-CG-OD1	7.44	124.99	118.30
15	P	36	G	P-O3'-C3'	6.96	128.05	119.70
15	P	32	G	C4-N9-C1'	6.88	135.44	126.50
1	A	118	LEU	CA-CB-CG	6.87	131.09	115.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	460	ARG	Peptide
1	A	538	VAL	Peptide
2	B	629	GLU	Peptide
2	B	71	ALA	Peptide
2	B	98	HIS	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11371	0	11488	183	0
2	B	9052	0	9081	130	0
3	C	2115	0	2057	36	0
4	D	1004	0	980	12	0
5	E	1720	0	1737	25	0
6	F	657	0	684	8	0
7	G	1333	0	1321	26	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	1186	0	1147	11	0
9	I	949	0	879	9	0
10	J	533	0	554	9	0
11	K	920	0	942	12	0
12	L	388	0	393	6	0
13	M	4737	0	2252	21	0
14	N	773	0	412	3	0
15	P	452	0	229	5	0
16	Q	4116	0	1657	20	0
17	T	974	0	541	9	0
18	U	440	0	18	0	0
19	V	324	0	15	1	0
20	W	2333	0	2246	21	0
21	X	63	0	2	0	0
22	Y	911	0	904	8	0
23	Z	3878	0	3941	38	0
24	A	2	0	0	0	0
24	B	1	0	0	0	0
24	C	1	0	0	0	0
24	I	2	0	0	0	0
24	J	1	0	0	0	0
24	L	1	0	0	0	0
24	Y	1	0	0	0	0
25	A	1	0	0	0	0
All	All	50239	0	43480	495	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 495 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)
5:E:85:LYS:NZ	14:N:41:DC:OP2	1.92	1.03
2:B:816:GLU:OE2	2:B:869:LYS:NZ	2.00	0.95
9:I:65:LEU:O	9:I:122:ARG:NH1	2.06	0.88
1:A:803:LYS:O	1:A:812:LYS:NZ	2.12	0.83
1:A:355:MET:SD	2:B:1091:ARG:NH1	2.50	0.82

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	1425/1970 (72%)	1340 (94%)	82 (6%)	3 (0%)	47	79
2	B	1122/1174 (96%)	1062 (95%)	59 (5%)	1 (0%)	51	83
3	C	259/275 (94%)	245 (95%)	14 (5%)	0	100	100
4	D	124/142 (87%)	122 (98%)	2 (2%)	0	100	100
5	E	207/210 (99%)	199 (96%)	7 (3%)	1 (0%)	29	64
6	F	80/127 (63%)	78 (98%)	2 (2%)	0	100	100
7	G	169/172 (98%)	165 (98%)	4 (2%)	0	100	100
8	H	146/150 (97%)	135 (92%)	11 (8%)	0	100	100
9	I	115/125 (92%)	107 (93%)	8 (7%)	0	100	100
10	J	65/67 (97%)	64 (98%)	1 (2%)	0	100	100
11	K	113/117 (97%)	109 (96%)	4 (4%)	0	100	100
12	L	44/58 (76%)	38 (86%)	6 (14%)	0	100	100
13	M	970/1726 (56%)	918 (95%)	52 (5%)	0	100	100
16	Q	581/1178 (49%)	555 (96%)	25 (4%)	1 (0%)	47	79
20	W	298/305 (98%)	278 (93%)	20 (7%)	0	100	100
22	Y	114/121 (94%)	114 (100%)	0	0	100	100
23	Z	476/1087 (44%)	466 (98%)	9 (2%)	1 (0%)	47	79
All	All	6308/9004 (70%)	5995 (95%)	306 (5%)	7 (0%)	54	83

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	57	ASP
1	A	911	PRO
2	B	492	ASP
1	A	540	ASP
23	Z	506	LEU



### 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	1257/1747 (72%)	1242 (99%)	15 (1%)	71	88
2	B	992/1027 (97%)	985 (99%)	7 (1%)	84	93
3	C	240/252 (95%)	239 (100%)	1 (0%)	91	96
4	D	109/126 (86%)	108 (99%)	1 (1%)	78	91
5	E	191/192 (100%)	189 (99%)	2 (1%)	76	90
6	F	71/111 (64%)	69 (97%)	2 (3%)	43	73
7	G	146/153 (95%)	141 (97%)	5 (3%)	37	69
8	H	129/131 (98%)	129 (100%)	0	100	100
9	I	105/112 (94%)	105 (100%)	0	100	100
10	J	56/56 (100%)	56 (100%)	0	100	100
11	K	104/106 (98%)	104 (100%)	0	100	100
12	L	43/55 (78%)	43 (100%)	0	100	100
13	M	154/1514 (10%)	153 (99%)	1 (1%)	86	94
16	Q	121/752 (16%)	120 (99%)	1 (1%)	81	92
20	W	255/260 (98%)	254 (100%)	1 (0%)	91	96
22	Y	102/105 (97%)	102 (100%)	0	100	100
23	Z	427/940 (45%)	426 (100%)	1 (0%)	93	97
All	All	4502/7639 (59%)	4465 (99%)	37 (1%)	82	92

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	G	53	ASN
20	W	237	ASN
7	G	78	ARG
7	G	110	ARG
1	A	1248	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37

such sidechains are listed below:

Mol	Chain	Res	Type
16	Q	756	ASN
23	Z	272	ASN
16	Q	775	ASN
20	W	221	HIS
2	B	23	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	20/46 (43%)	8 (40%)	3 (15%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
15	P	27	A
15	P	29	C
15	P	30	C
15	P	31	G
15	P	33	A

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
15	P	28	A
15	P	36	G
15	P	38	G

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	SEP	A	1547	1,13	8,9,10	1.69	2 (25%)	8,12,14	2.36	2 (25%)
1	TPO	A	1525	1	8,10,11	1.52	1 (12%)	10,14,16	2.02	1 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SEP	A	1547	1,13	-	1/5/8/10	-
1	TPO	A	1525	1	-	0/9/11/13	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1547	SEP	P-O1P	3.57	1.62	1.50
1	A	1525	TPO	P-O1P	3.24	1.61	1.50
1	A	1547	SEP	P-O2P	2.00	1.62	1.54

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1525	TPO	P-OG1-CB	-5.99	105.11	123.21
1	A	1547	SEP	P-OG-CB	-5.94	101.94	118.30
1	A	1547	SEP	OG-CB-CA	2.46	110.53	108.14

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1547	SEP	N-CA-CB-OG

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1547	SEP	2	0
1	A	1525	TPO	1	0

## 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
18	U	3
19	V	3
13	M	1

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	U	492:UNK	C	498:UNK	N	29.01
1	V	293:UNK	C	307:UNK	N	14.62
1	U	429:UNK	C	448:UNK	N	12.83
1	V	227:UNK	C	250:UNK	N	9.56
1	V	260:UNK	C	285:UNK	N	6.58

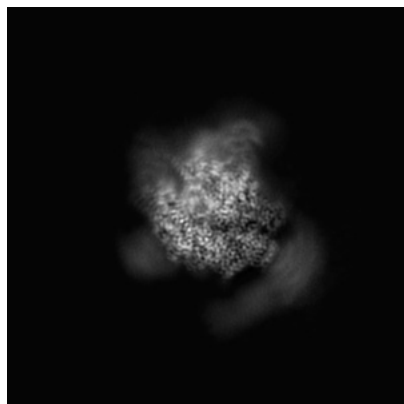
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-0031. These allow visual inspection of the internal detail of the map and identification of artifacts.

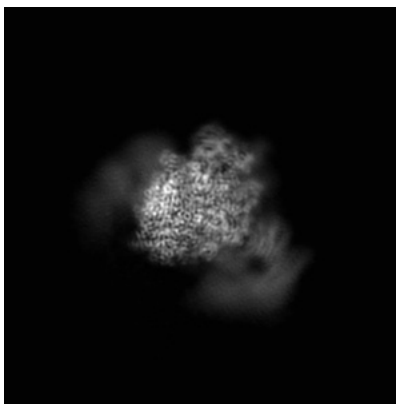
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

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

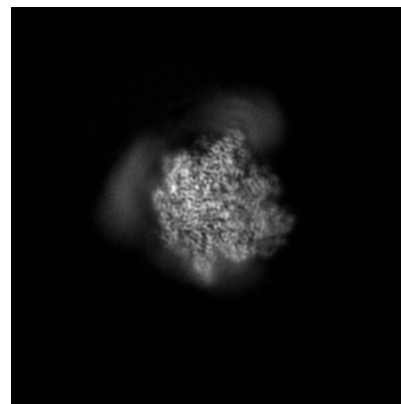
#### 6.1.1 Primary map



X

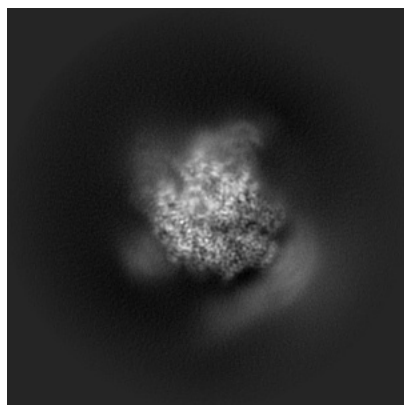


Y

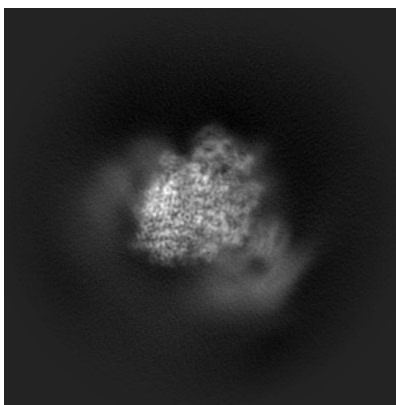


Z

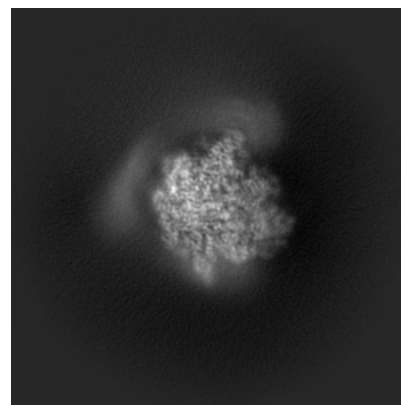
#### 6.1.2 Raw map



X



Y

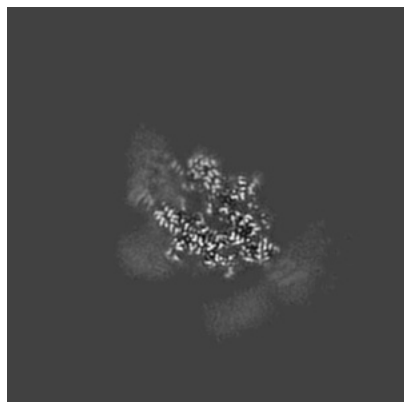


Z

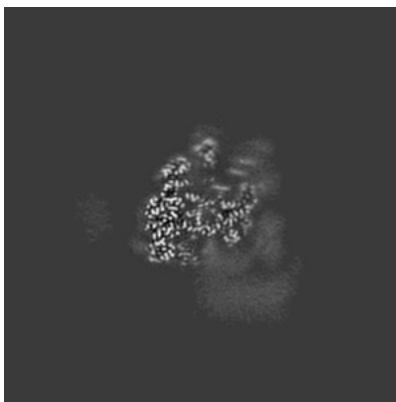
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

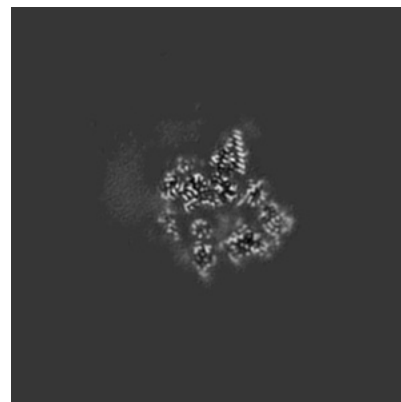
### 6.2.1 Primary map



X Index: 180

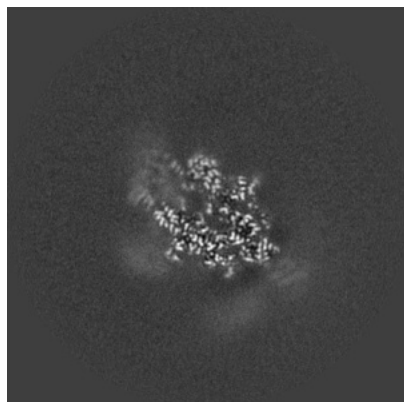


Y Index: 180

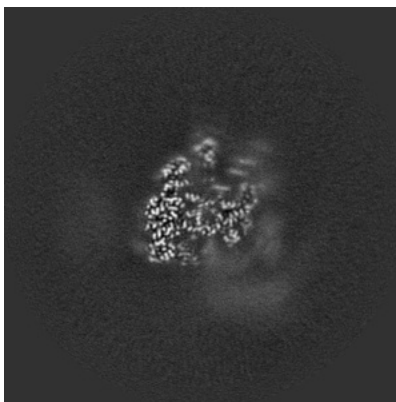


Z Index: 180

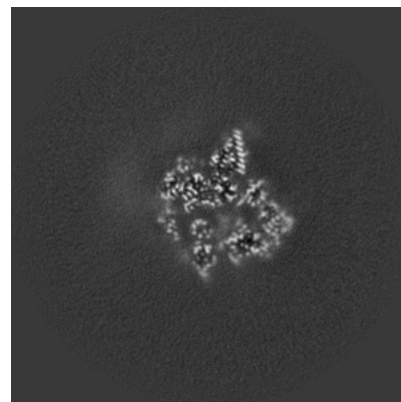
### 6.2.2 Raw map



X Index: 180



Y Index: 180

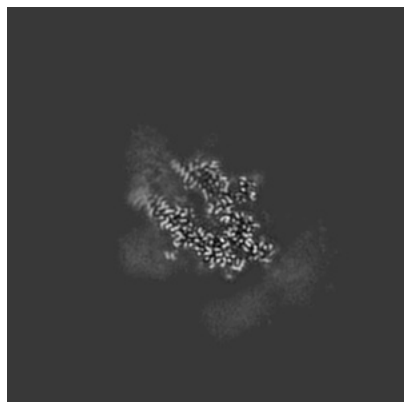


Z Index: 180

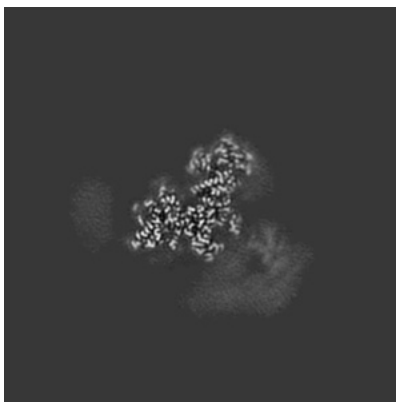
The images above show central slices of the map in three orthogonal directions.

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

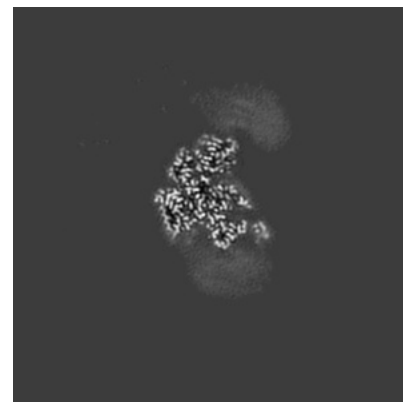
### 6.3.1 Primary map



X Index: 177

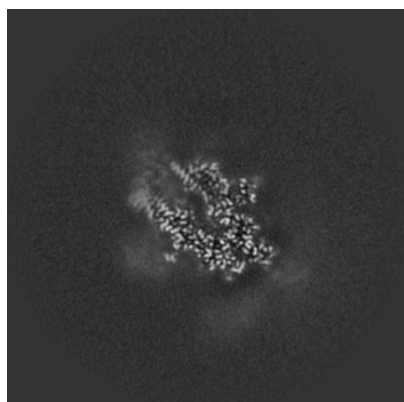


Y Index: 198

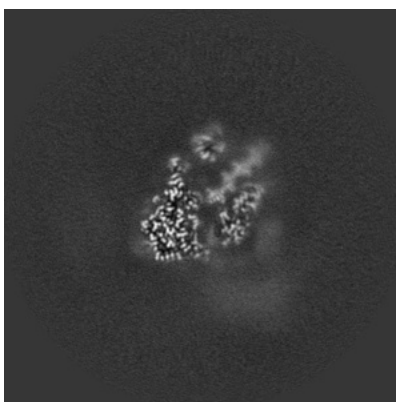


Z Index: 143

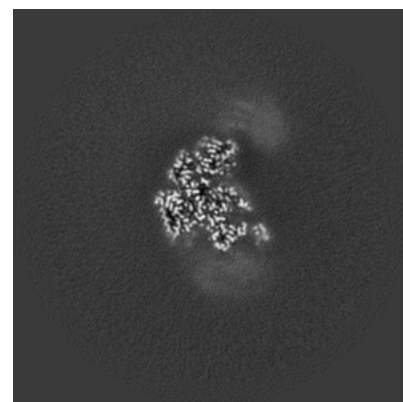
### 6.3.2 Raw map



X Index: 177



Y Index: 173

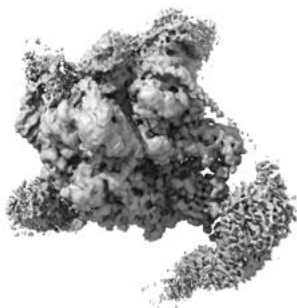


Z Index: 143

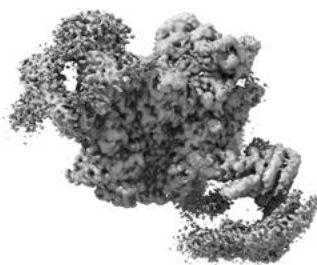
The images above show the largest variance slices of the map in three orthogonal directions.

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

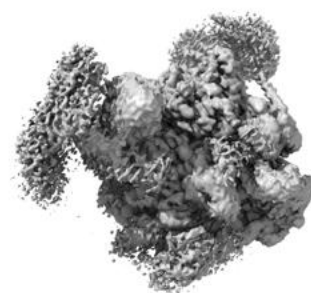
### 6.4.1 Primary map



X



Y



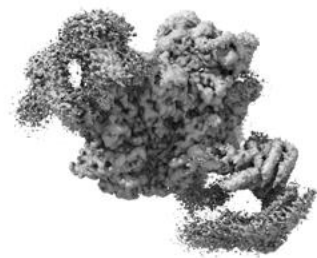
Z

The images above show the 3D surface view of the map at the recommended contour level 0.0157. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.4.2 Raw map



X



Y



Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.



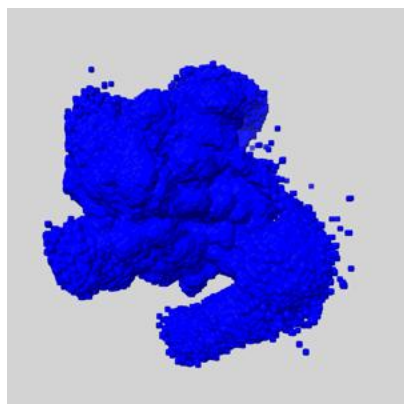
## 6.5 Mask visualisation [i](#)

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

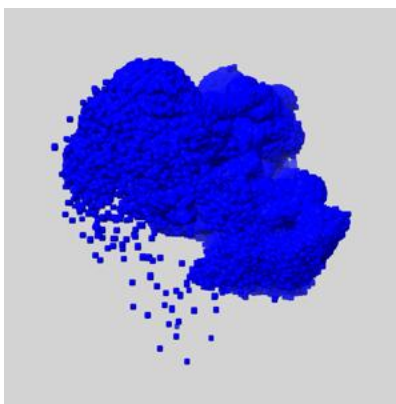
A mask typically either:

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

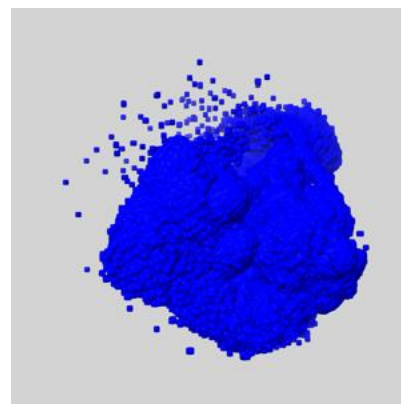
### 6.5.1 emd\_0031\_msk\_1.map [i](#)



X



Y

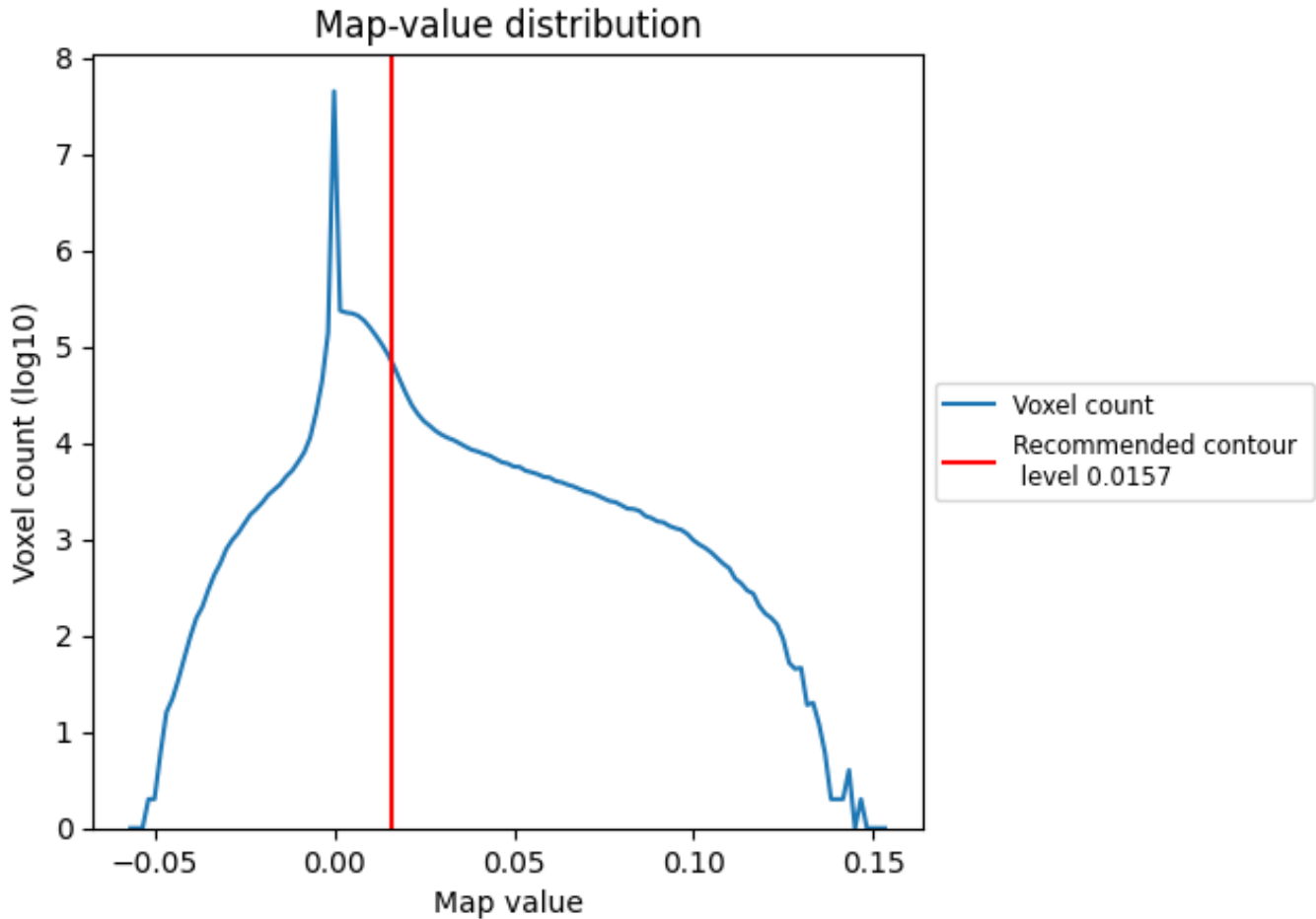


Z

## 7 Map analysis [i](#)

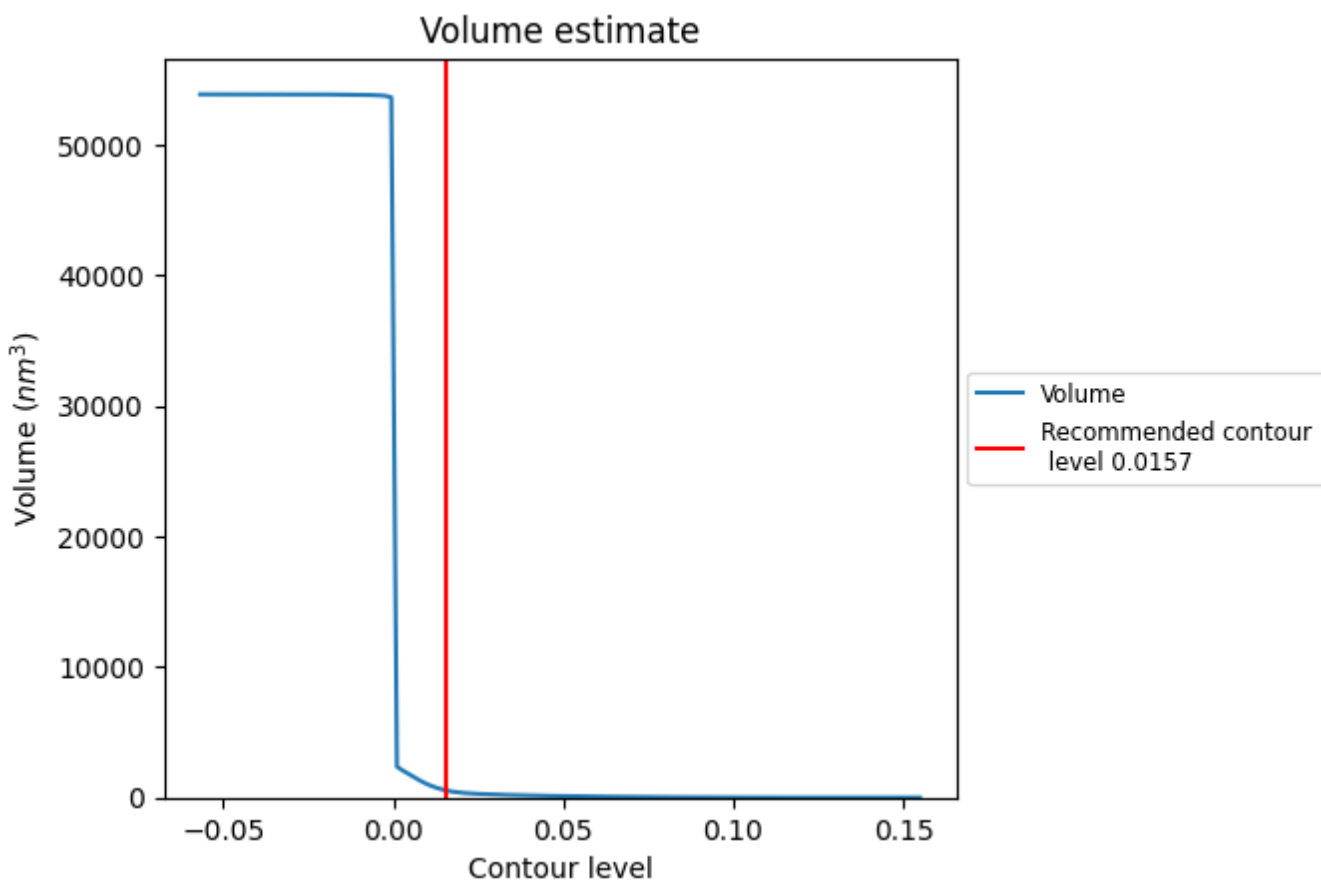
This section contains the results of statistical analysis of the map.

### 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.

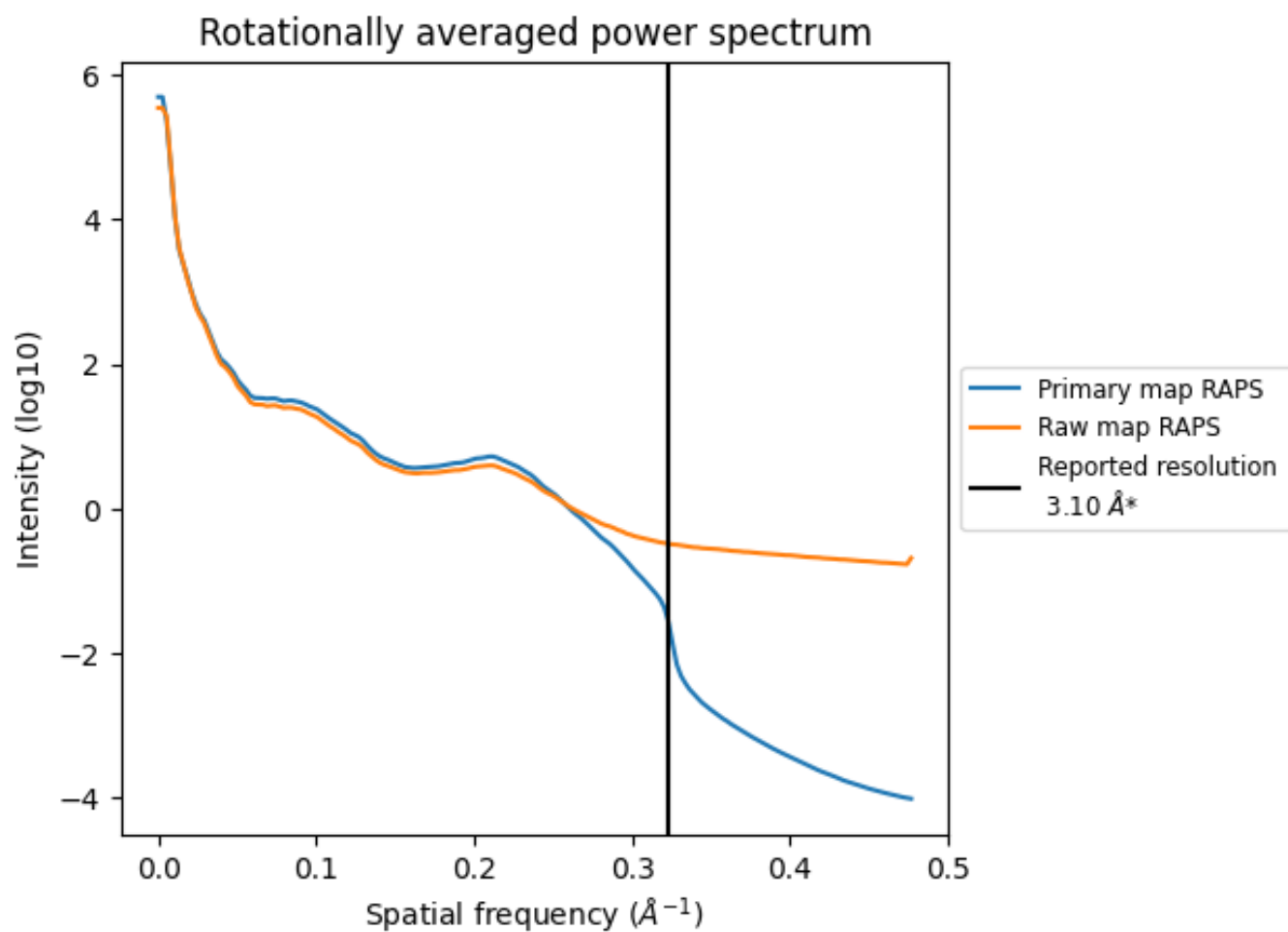
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 541 nm<sup>3</sup>; this corresponds to an approximate mass of 489 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

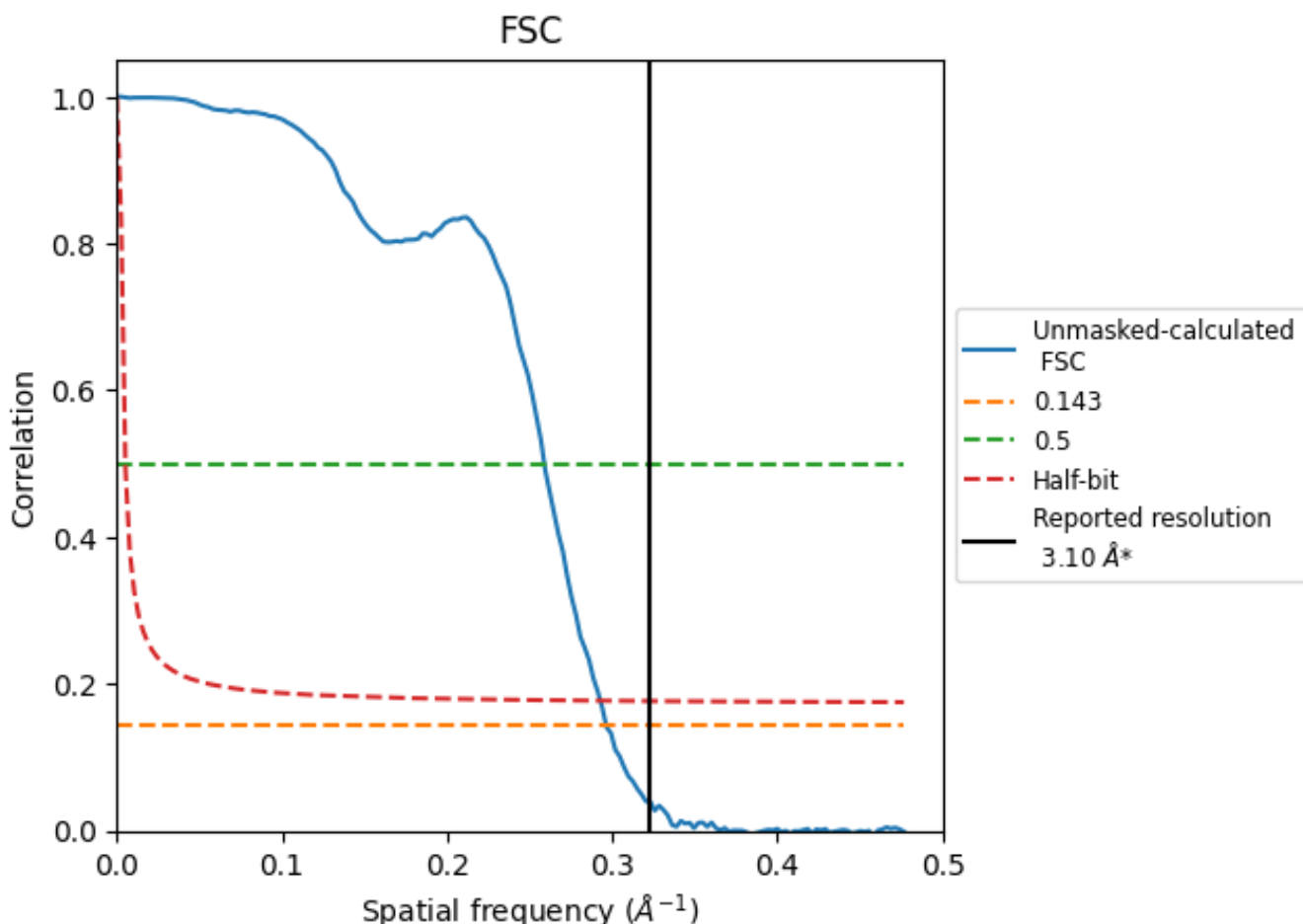


\*Reported resolution corresponds to spatial frequency of 0.323 Å<sup>-1</sup>

## 8 Fourier-Shell correlation [i](#)

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

### 8.1 FSC [i](#)



\*Reported resolution corresponds to spatial frequency of 0.323  $\text{\AA}^{-1}$

## 8.2 Resolution estimates [i](#)

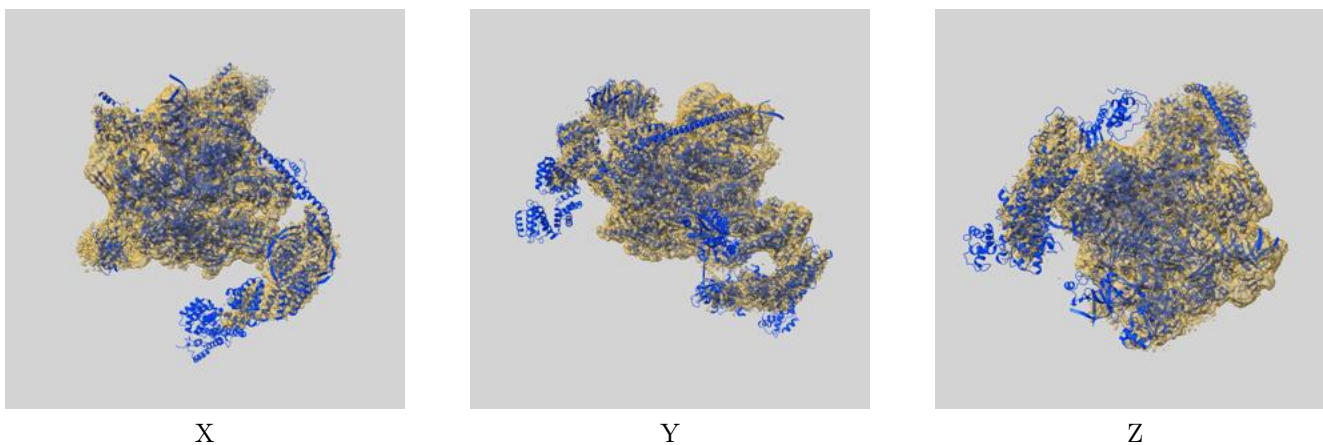
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.37	3.86	3.42

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

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

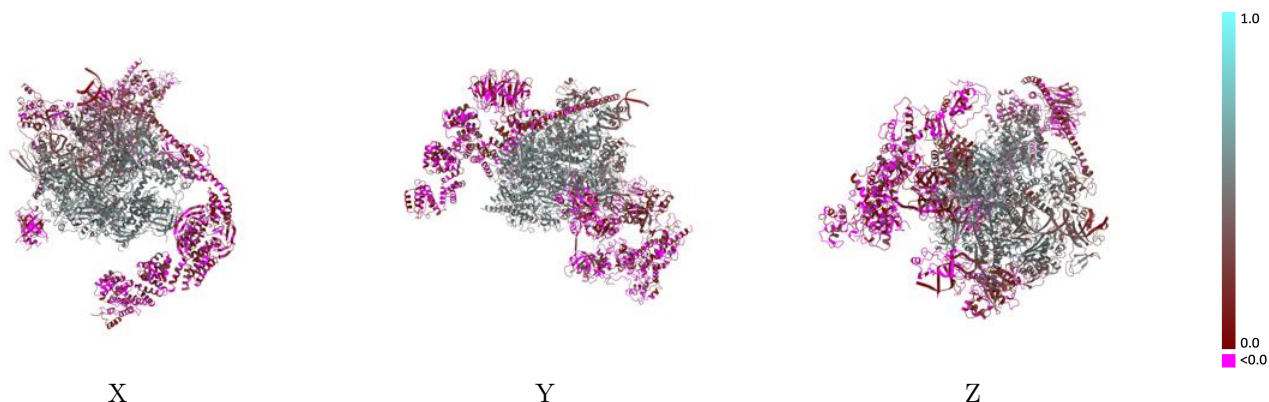
This section contains information regarding the fit between EMDB map EMD-0031 and PDB model 6GMH. Per-residue inclusion information can be found in section [3](#) on page [9](#).

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



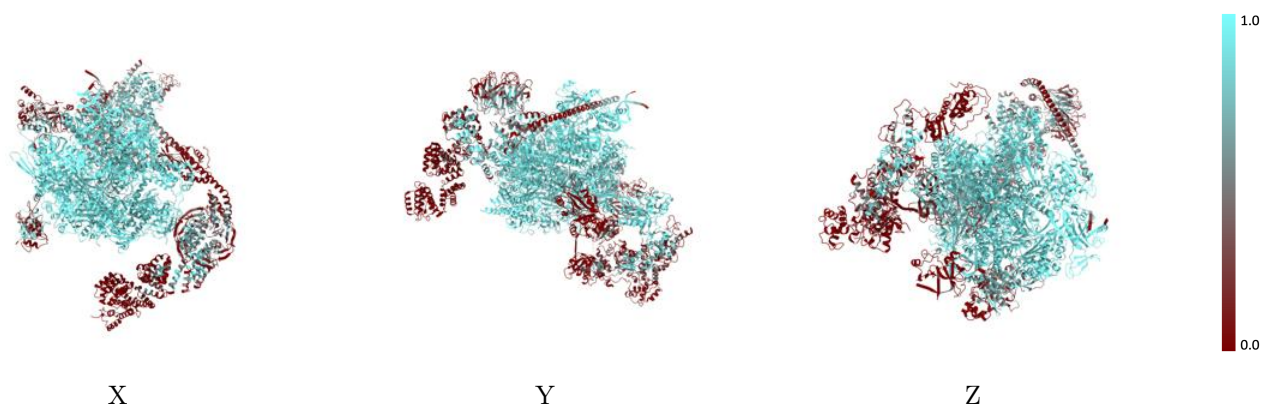
The images above show the 3D surface view of the map at the recommended contour level 0.0157 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 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.

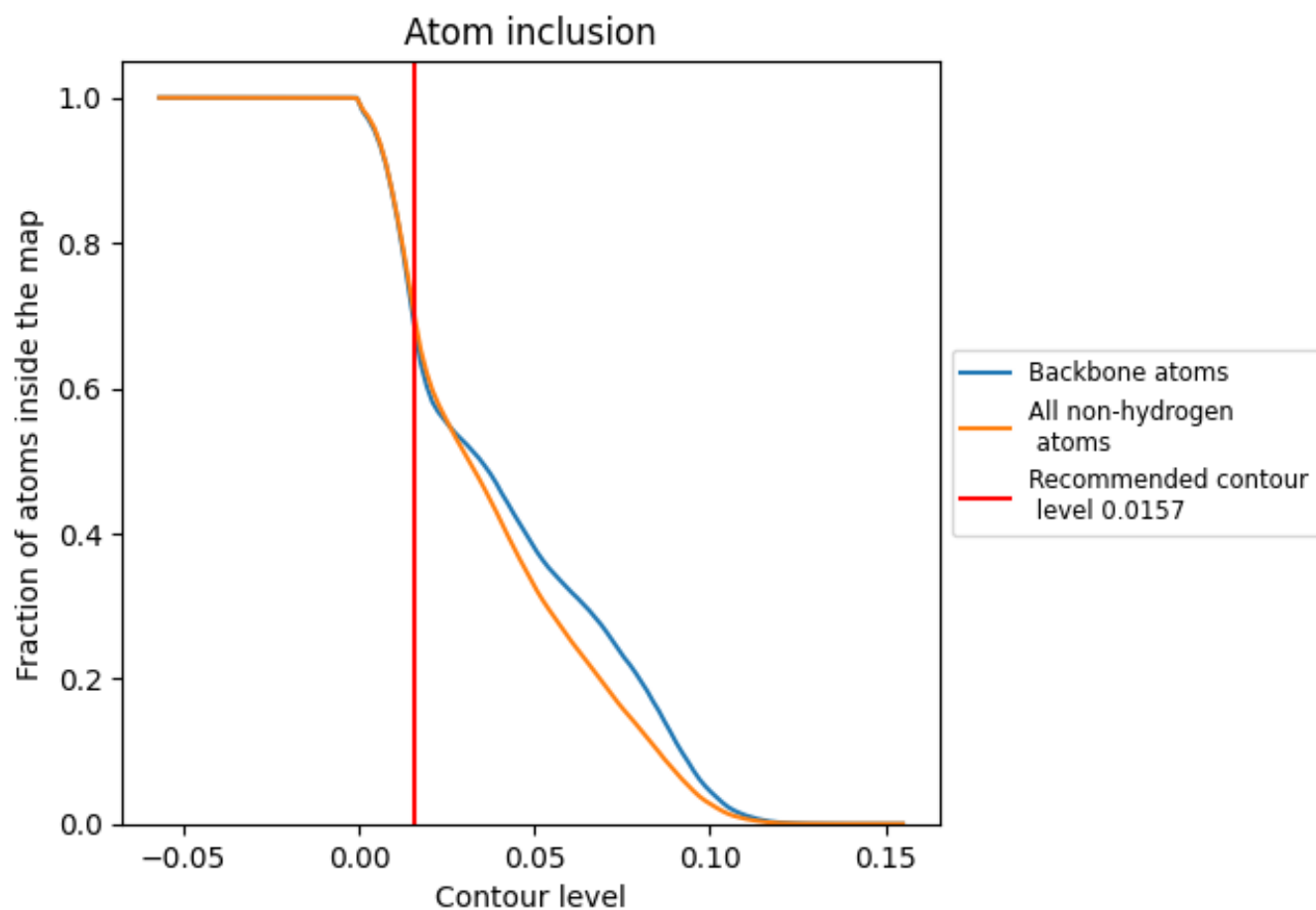
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (0.0157).



















































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 69% of all backbone atoms, 70% 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 (0.0157) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7031	 0.3240
A	 0.9288	 0.4710
B	 0.9537	 0.4950
C	 0.9459	 0.5090
D	 0.8234	 0.2100
E	 0.9541	 0.4400
F	 0.9326	 0.5000
G	 0.8432	 0.2540
H	 0.9284	 0.4960
I	 0.9398	 0.4270
J	 0.9597	 0.5360
K	 0.9616	 0.5230
L	 0.9409	 0.4820
M	 0.1683	 0.0280
N	 0.8370	 0.2560
P	 0.7013	 0.2670
Q	 0.2773	 0.0810
T	 0.8398	 0.3330
U	 0.3477	 0.0230
V	 0.4383	 0.0630
W	 0.3519	 0.0410
X	 0.5714	 0.1690
Y	 0.2036	 0.1250
Z	 0.3032	 0.1100

