



## wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Jul 24, 2017 – 01:08 PM EDT

PDB ID : 4V1M  
EMDB ID: : EMD-2784  
Title : Architecture of the RNA polymerase II-Mediator core transcription initiation complex  
Authors : Plaschka, C.; Lariviere, L.; Wenzek, L.; Hemann, M.; Tegunov, D.; Petrotchenko, E.V.; Borchers, C.H.; Baumeister, W.; Herzog, F.; Villa, E.; Cramer, P.  
Deposited on : unknown  
Resolution : 6.60 Å(reported)  
Based on PDB ID : 4A3D

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

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

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

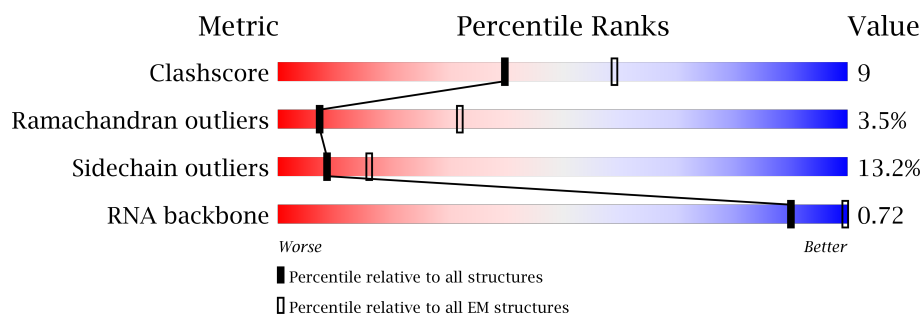
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 6.60 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1733	58% 20% . . 18%
2	B	1224	63% 28% . 6%
3	C	318	61% 19% . 16%
4	E	215	74% 23% .
5	F	155	37% 15% . 46%
6	H	146	55% 30% 5% . 9%
7	I	122	74% 21% . .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
8	J	70	
9	K	120	
10	L	70	
11	N	10	
12	P	6	
13	T	20	

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 29447 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 SUBUNIT RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1422	Total	C	N	O	S	0	0
			11174	7037	1954	2121	62		

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1156	Total	C	N	O	S	0	0
			9143	5784	1606	1697	56		

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 1.

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

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 2.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 3.

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

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

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

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	K	115	Total	C	N	O	S	0	1
			920	590	157	171	2		

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	L	46	Total	C	N	O	S	0	0
			363	224	72	63	4		

- Molecule 11 is a DNA chain called 5'-D(\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	N	10	Total	C	N	O	P	0	0
			207	99	39	59	10		

- Molecule 12 is a RNA chain called 5'-D(\*CP\*CP\*AP\*GP\*GP\*AP)-3'.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	P	6	Total	C	N	O	P	0	0
			130	58	26	40	6		

- Molecule 13 is a DNA chain called 5'-D(\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*TP\*TP\*TP\*CP\*CP\*BRUP\*GP\*GP\*TP\*C)-3'.

Mol	Chain	Residues	Atoms						AltConf	Trace
13	T	20	Total	Br	C	N	O	P	0	0
			404	1	194	63	126	20		

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

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

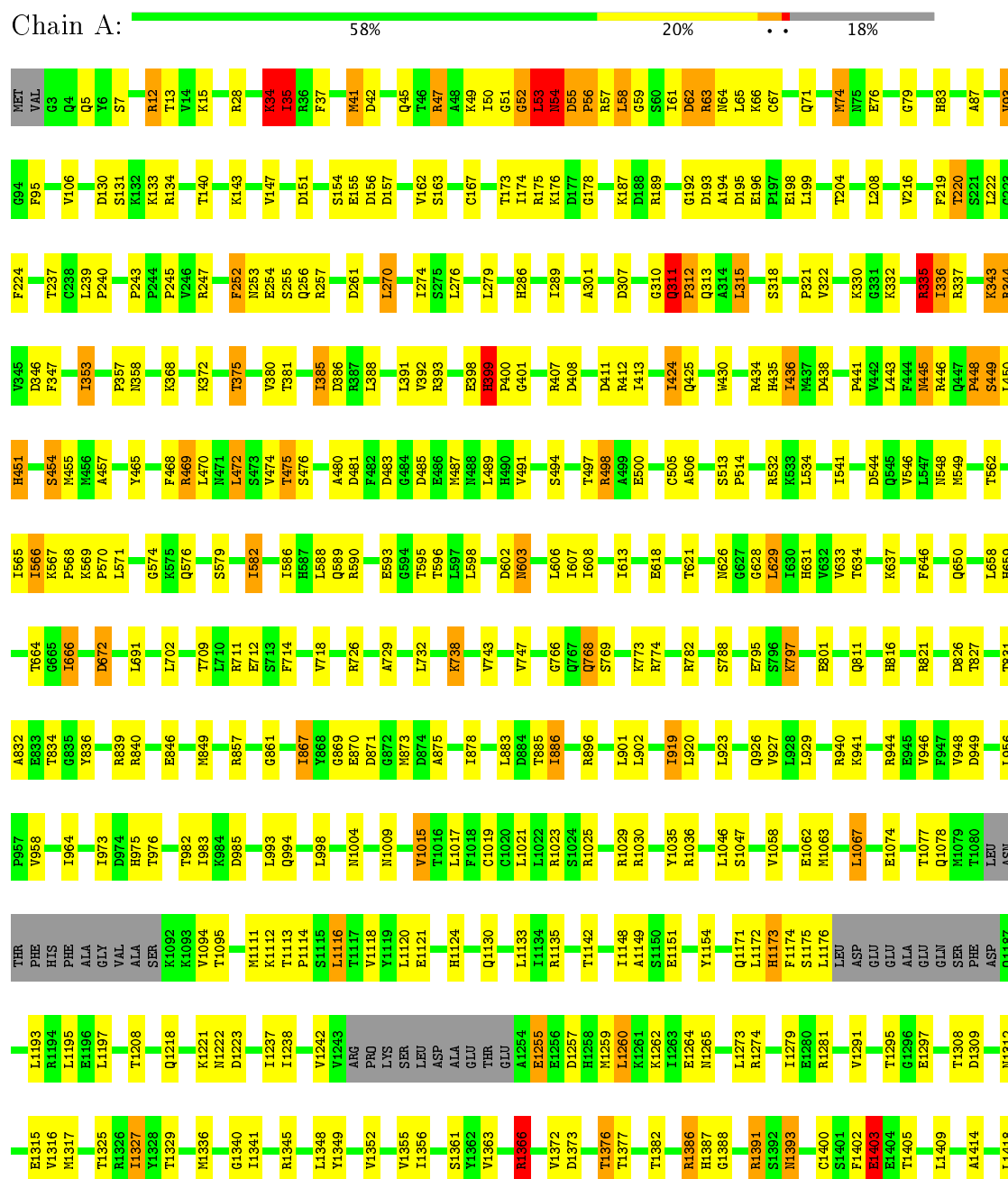
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

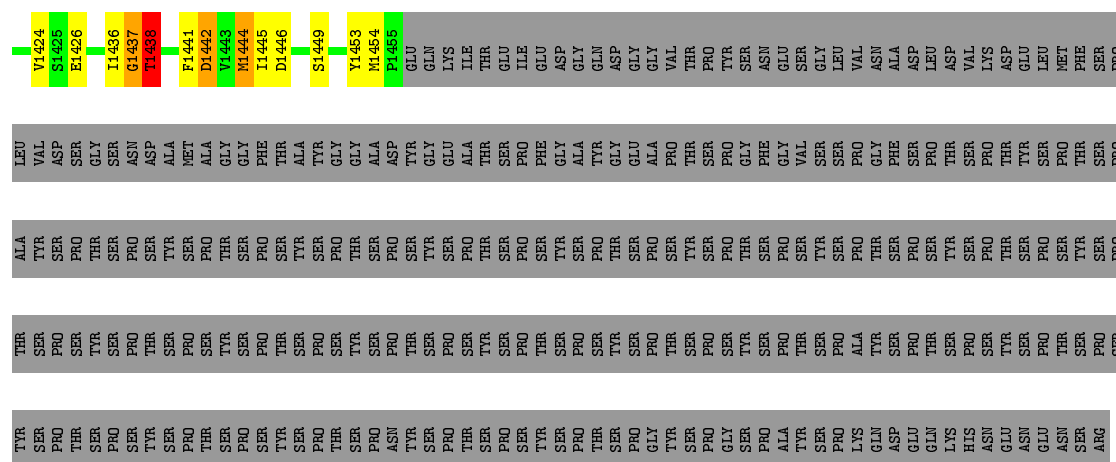
Mol	Chain	Residues	Atoms		AltConf
15	A	1	Total	Mg	0
			1	1	

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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. 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. 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.

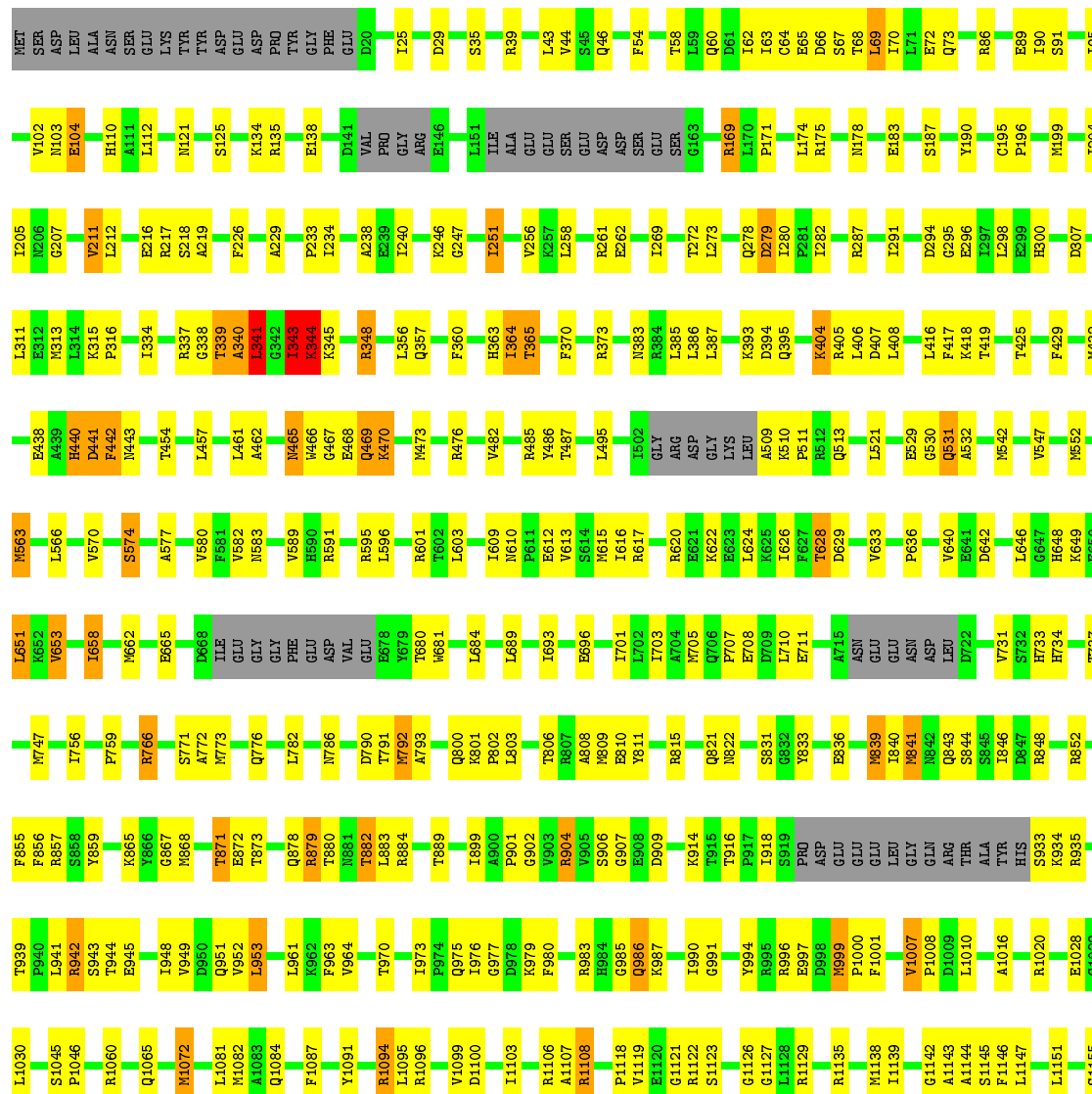
- Molecule 1: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1





## • Molecule 2: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2

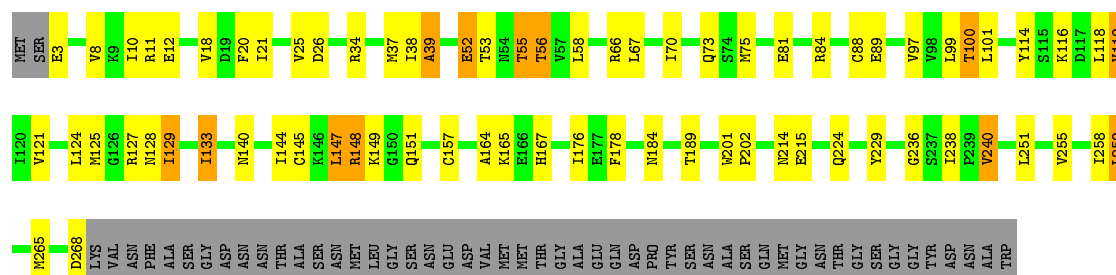
Chain B: 63% 28% 6%



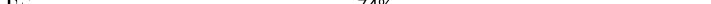


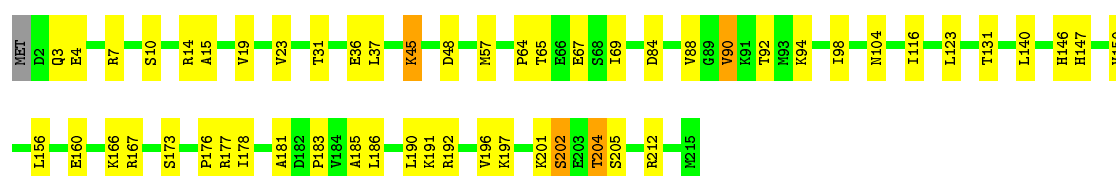
● Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

Chain C:  61% 19% 1% 16%



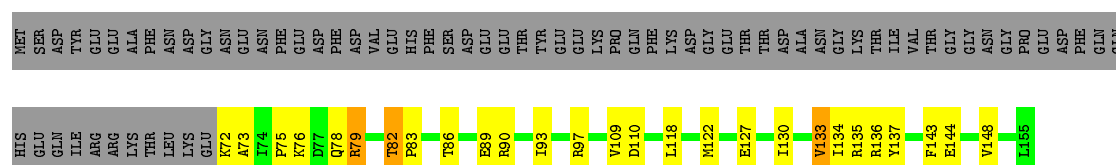
● Molecule 4: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

Chain E:  74% 23%



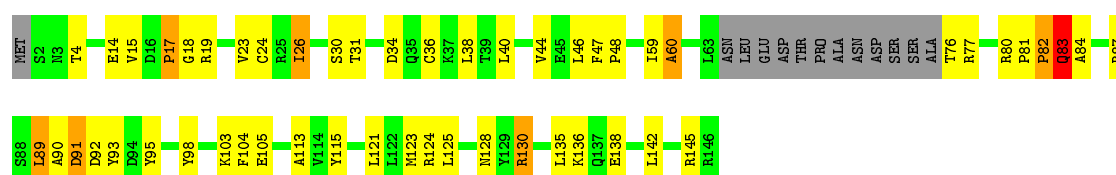
● Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F:  37% 15% 46%



● Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H:  55% 30% 5% 9%

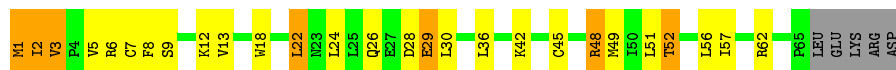


- Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I:  74% 21% 5%



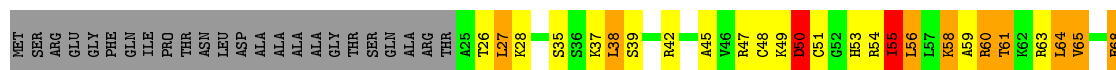
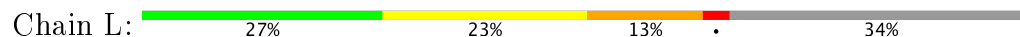
- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5



- Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11



- Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4



- Molecule 11: 5'-D(\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP)-3'

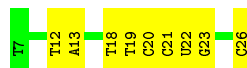


There are no outlier residues recorded for this chain.

- Molecule 12: 5'-D(\*CP\*CP\*AP\*GP\*GP\*AP)-3'



- Molecule 13: 5'-D(\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*TP\*TP\*TP\*CP \*CP\*BRUP\*GP\*GP\*TP\*C)-3'



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	14777	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	37169	Depositor
Image detector	GATAN K2 (4k x 4k)	Depositor

## 5 Model quality

### 5.1 Standard geometry

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

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  >2	RMSZ	# Z  >2
1	A	0.54	0/11374	0.85	10/15383 (0.1%)
10	L	0.54	0/365	0.95	0/485
11	N	1.07	0/232	1.06	0/356
12	P	1.21	0/145	0.79	0/224
13	T	1.26	1/426 (0.2%)	1.08	0/652
2	B	0.49	0/9318	0.79	10/12565 (0.1%)
3	C	0.49	0/2133	0.78	2/2891 (0.1%)
4	E	0.48	0/1788	0.71	0/2406
5	F	0.62	0/691	0.81	0/933
6	H	0.51	0/1086	0.80	0/1470
7	I	0.47	0/989	0.78	0/1331
8	J	0.54	0/541	0.88	1/727 (0.1%)
9	K	0.47	0/938	0.71	0/1267
All	All	0.55	1/30026 (0.0%)	0.82	23/40690 (0.1%)

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	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	T	26	DC	C1'-N1	5.58	1.56	1.49

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	465	ASN	O-C-N	15.46	147.44	122.70
2	B	465	ASN	CA-C-N	-12.77	89.11	117.20
2	B	465	ASN	C-N-CA	-9.63	97.63	121.70
2	B	218	SER	O-C-N	7.31	134.39	122.70
1	A	399	HIS	N-CA-CB	7.26	123.67	110.60

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	34	LYS	Mainchain,Peptide
2	B	404	LYS	Mainchain
2	B	43	LEU	Mainchain

## 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	11174	0	11233	204	0
2	B	9143	0	9122	204	0
3	C	2095	0	2051	42	0
4	E	1752	0	1776	26	0
5	F	679	0	701	19	0
6	H	1068	0	1040	24	0
7	I	971	0	927	15	0
8	J	532	0	542	15	0
9	K	920	0	929	20	0
10	L	363	0	386	20	0
11	N	207	0	114	0	0
12	P	130	0	66	1	0
13	T	404	0	227	6	0
14	A	2	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
15	A	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	29447	0	29114	514	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 9.

The worst 5 of 514 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:A:867:ILE:CD1	1:A:867:ILE:CG1	1.83	1.51
2:B:104:GLU:OE2	10:L:54:ARG:HD3	1.50	1.10
1:A:53:LEU:HD23	1:A:54:ASN:H	1.03	1.10
2:B:121:ASN:ND2	2:B:963:PHE:CZ	2.25	1.04
2:B:104:GLU:OE2	10:L:54:ARG:CD	2.10	1.00

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	1414/1733 (82%)	1236 (87%)	126 (9%)	52 (4%)	4	33
2	B	1138/1224 (93%)	1008 (89%)	86 (8%)	44 (4%)	3	31
3	C	264/318 (83%)	242 (92%)	20 (8%)	2 (1%)	22	66
4	E	212/215 (99%)	195 (92%)	13 (6%)	4 (2%)	9	47
5	F	82/155 (53%)	76 (93%)	6 (7%)	0	100	100
6	H	129/146 (88%)	106 (82%)	14 (11%)	9 (7%)	1	19
7	I	117/122 (96%)	98 (84%)	16 (14%)	3 (3%)	6	40
8	J	63/70 (90%)	51 (81%)	9 (14%)	3 (5%)	2	27
9	K	113/120 (94%)	109 (96%)	4 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	L	44/70 (63%)	27 (61%)	9 (20%)	8 (18%)	0	3
All	All	3576/4173 (86%)	3148 (88%)	303 (8%)	125 (4%)	7	34

5 of 125 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ARG
1	A	58	LEU
1	A	76	GLU
1	A	189	ARG
1	A	195	ASP

### 5.3.2 Protein sidechains ⓘ

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	1240/1520 (82%)	1062 (86%)	178 (14%)	4	22
2	B	986/1061 (93%)	861 (87%)	125 (13%)	5	25
3	C	234/274 (85%)	206 (88%)	28 (12%)	6	27
4	E	196/197 (100%)	175 (89%)	21 (11%)	8	31
5	F	74/137 (54%)	67 (90%)	7 (10%)	10	36
6	H	117/128 (91%)	103 (88%)	14 (12%)	6	27
7	I	113/116 (97%)	106 (94%)	7 (6%)	21	54
8	J	60/65 (92%)	49 (82%)	11 (18%)	2	12
9	K	99/102 (97%)	87 (88%)	12 (12%)	6	27
10	L	40/57 (70%)	27 (68%)	13 (32%)	0	2
All	All	3159/3657 (86%)	2743 (87%)	416 (13%)	8	24

5 of 416 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	175	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	616	ILE
8	J	12	LYS
2	B	251	ILE
2	B	440	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1393	ASN
2	B	357	GLN
7	I	83	ASN
2	B	300	HIS
1	A	545	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
12	P	5/6 (83%)	1 (20%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
12	P	8	G

There are no RNA pucker outliers to report.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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
13	BRU	T	22	13,12	13,21,22	1.26	2 (15%)	16,30,33	3.57	4 (25%)

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
13	BRU	T	22	13,12	-	0/3/21/22	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	T	22	BRU	O5'-C5'	-2.39	1.41	1.44
13	T	22	BRU	C4-C5	2.96	1.42	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	T	22	BRU	C5-C4-N3	-7.36	114.83	123.64
13	T	22	BRU	C2'-C1'-N1	-2.53	108.25	114.23
13	T	22	BRU	O4'-C1'-N1	4.47	115.31	107.78
13	T	22	BRU	C4-N3-C2	11.03	124.81	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	T	22	BRU	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 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
2	B	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	934:LYS	C	935:ARG	N	5.39
1	B	351:TYR	C	352:ALA	N	3.28