



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Dec 10, 2019 – 07:00 PM EST

PDB ID : 6C04
EMDB ID: : EMD-7320
Title : Mtb RNAP Holo/RbpA/double fork DNA -closed clamp
Authors : Darst, S.A.; Campbell, E.A.; Boyaci Selcuk, H.; Chen, J.; Lilic, M.
Deposited on : 2017-12-27
Resolution : 3.27 Å(reported)

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

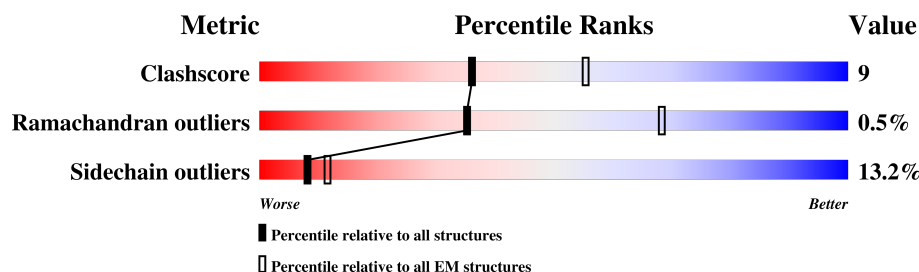
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.27 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.




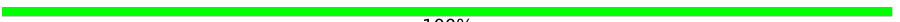
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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 ≥ 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	347	
1	B	347	
2	C	1179	
3	D	1326	
4	E	110	
5	F	531	
6	J	111	
7	H	31	
7	O	31	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	G	26	 62% 8% 31%
8	P	26	 100%

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 27982 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 subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	226	Total	C	N	O	S	0	0
			1724	1085	297	339	3		
1	B	237	Total	C	N	O	S	0	0
			1775	1120	304	348	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1111	Total	C	N	O	S	0	0
			8585	5378	1507	1661	39		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1179	LEU	-	expression tag	UNP V9Z879
C	1180	ALA	-	expression tag	UNP V9Z879
C	1181	ARG	-	expression tag	UNP V9Z879
C	1182	HIS	-	expression tag	UNP V9Z879
C	1183	GLY	-	expression tag	UNP V9Z879
C	1184	GLY	-	expression tag	UNP V9Z879
C	1185	SER	-	expression tag	UNP V9Z879

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1267	Total	C	N	O	S	0	0
			9881	6190	1795	1854	42		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	GLY	-	expression tag	UNP A0A045J9E2
D	0	ALA	-	expression tag	UNP A0A045J9E2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1317	HIS	-	expression tag	UNP A0A045J9E2
D	1318	HIS	-	expression tag	UNP A0A045J9E2
D	1319	HIS	-	expression tag	UNP A0A045J9E2
D	1320	HIS	-	expression tag	UNP A0A045J9E2
D	1321	HIS	-	expression tag	UNP A0A045J9E2
D	1322	HIS	-	expression tag	UNP A0A045J9E2
D	1323	HIS	-	expression tag	UNP A0A045J9E2
D	1324	HIS	-	expression tag	UNP A0A045J9E2

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	E	83	Total	C	N	O	0	0
			649	414	108	127		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	GLY	-	expression tag	UNP A0A0T9N9K3

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	319	Total	C	N	O	S	0	0
			2508	1566	453	480	9		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	expression tag	UNP A0A045HD00
F	-1	PRO	-	expression tag	UNP A0A045HD00
F	0	HIS	-	expression tag	UNP A0A045HD00

- Molecule 6 is a protein called RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	108	Total	C	N	O	S	0	0
			881	543	168	167	3		

- Molecule 7 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	O	31	Total	C	N	O	P	0	0
			634	305	114	185	30		
7	H	22	Total	C	N	O	P	0	0
			454	218	82	132	22		

- Molecule 8 is a DNA chain called DNA (26-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	P	26	Total	C	N	O	P	0	0
			526	254	94	153	25		
8	G	18	Total	C	N	O	P	0	0
			362	176	64	105	17		

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
9	D	2	Total	Zn	0
			2	2	

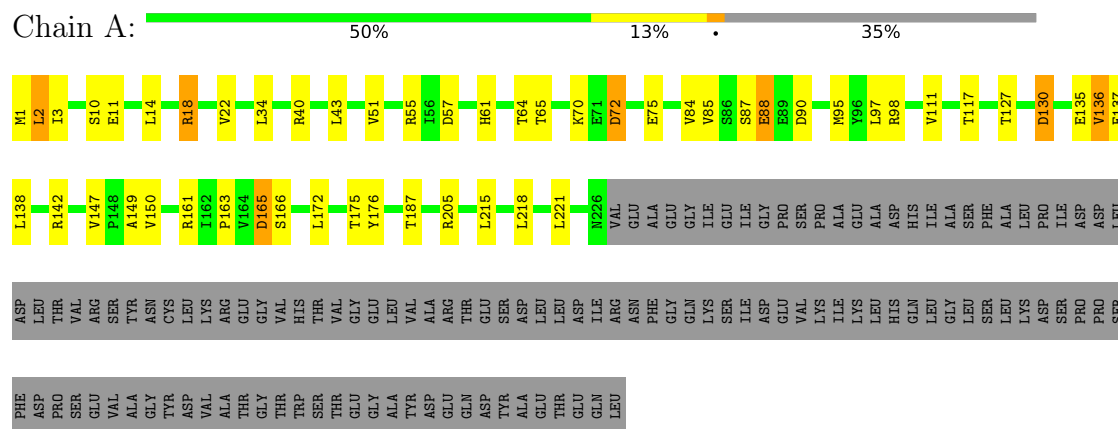
- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
10	D	1	Total	Mg	0
			1	1	

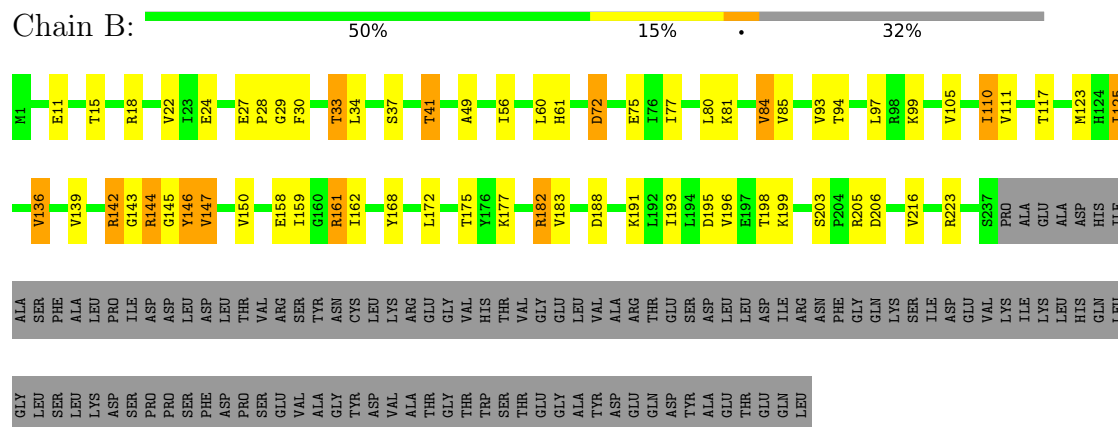
3 Residue-property plots

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.

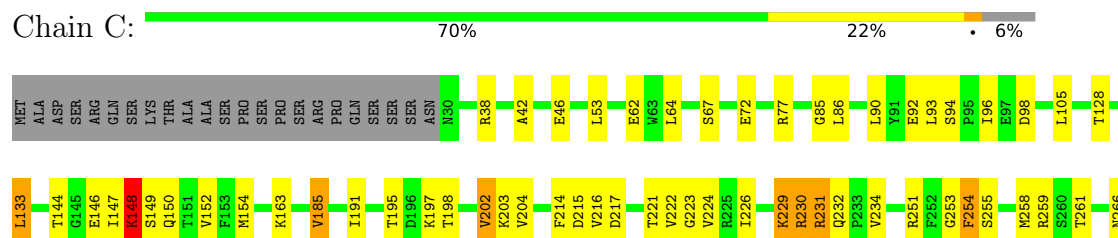
- Molecule 1: DNA-directed RNA polymerase subunit alpha

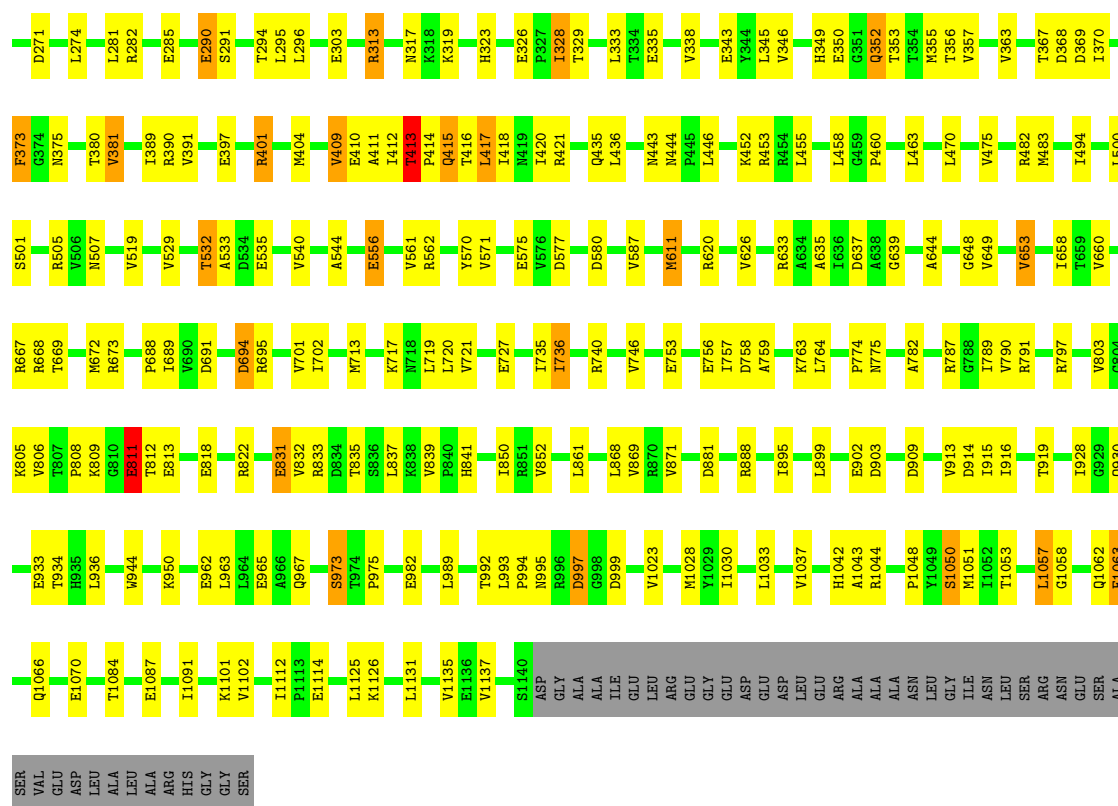


- Molecule 1: DNA-directed RNA polymerase subunit alpha



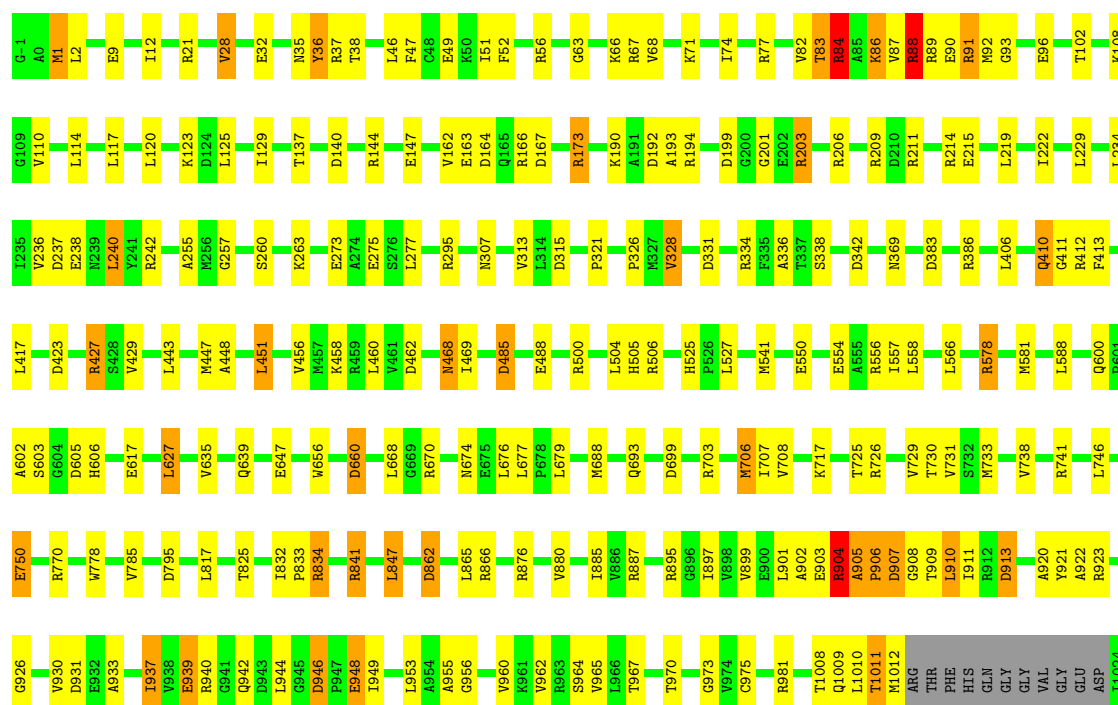
- Molecule 2: DNA-directed RNA polymerase subunit beta

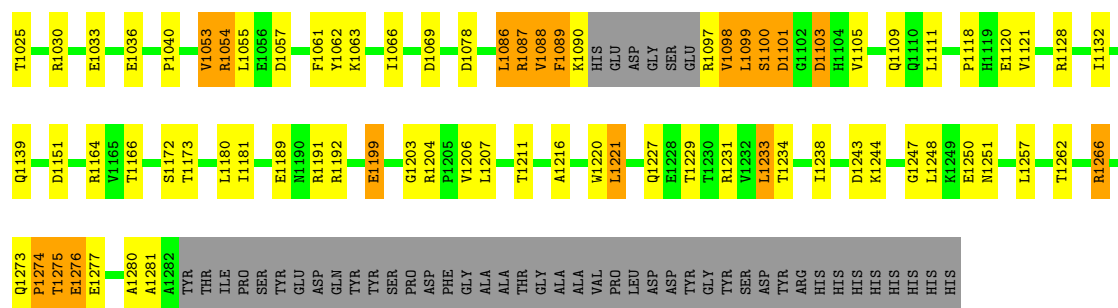




• Molecule 3: DNA-directed RNA polymerase subunit beta'

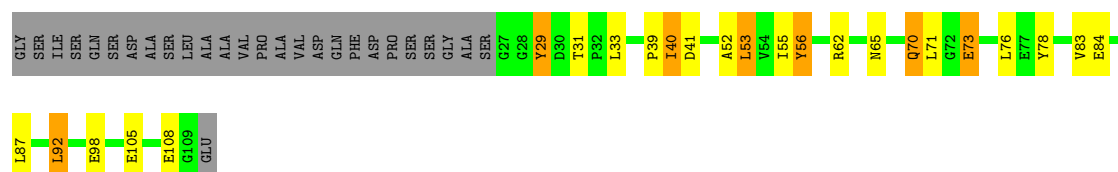
Chain D: 72% 19%





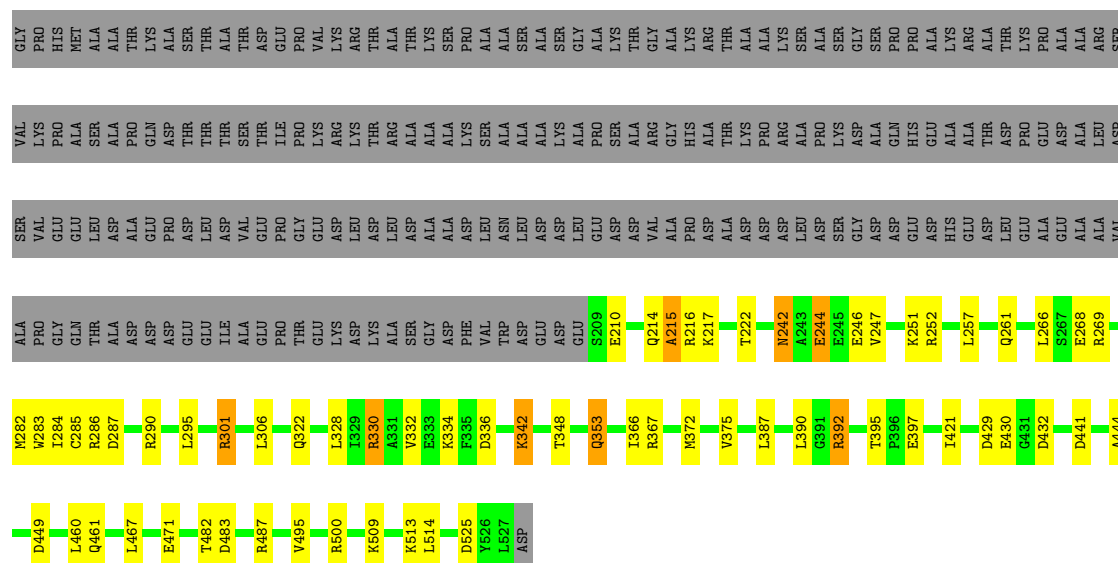
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 54% 15% 6% 25%



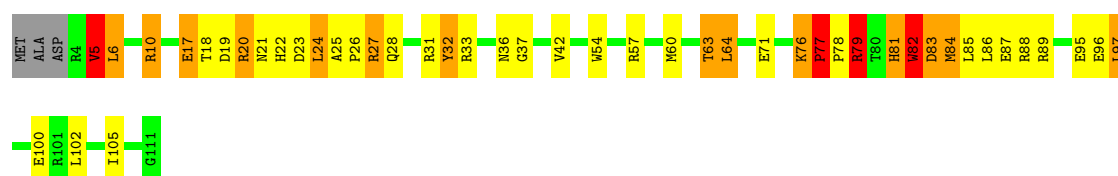
- Molecule 5: RNA polymerase sigma factor SigA

Chain F: 48% 11% 40%




- Molecule 6: RNA polymerase-binding protein RbpA

Chain J: 56% 25% 13%



- Molecule 7: DNA (31-MER)

Chain O:  87% 13%



- Molecule 7: DNA (31-MER)

Chain H:  52% 16% 29%



- Molecule 8: DNA (26-MER)

Chain P:  100%

There are no outlier residues recorded for this chain.

- Molecule 8: DNA (26-MER)

Chain G:  62% 8% 31%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	171547	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 ($e^-/\text{\AA}^2$)	6.7	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	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: ZN, 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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.39	0/1750	0.56	0/2380
1	B	0.73	9/1802 (0.5%)	0.61	1/2454 (0.0%)
2	C	0.44	0/8743	0.59	6/11859 (0.1%)
3	D	0.67	32/10045 (0.3%)	0.63	12/13581 (0.1%)
4	E	0.51	0/662	0.66	0/901
5	F	0.33	0/2539	0.47	0/3426
6	J	0.93	3/897 (0.3%)	1.13	5/1210 (0.4%)
7	H	0.75	1/509 (0.2%)	0.96	0/784
7	O	0.73	0/710	1.01	1/1095 (0.1%)
8	G	0.67	0/405	1.00	0/622
8	P	0.72	0/589	0.96	0/906
All	All	0.58	45/28651 (0.2%)	0.66	25/39218 (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
2	C	0	3
3	D	0	4
All	All	0	7

The worst 5 of 45 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	146	TYR	CE1-CZ	-12.56	1.22	1.38
3	D	36	TYR	CE1-CZ	-9.97	1.25	1.38
3	D	52	PHE	CG-CD1	-9.75	1.24	1.38
3	D	88	ARG	C-O	-8.92	1.06	1.23
3	D	92	MET	C-O	-8.56	1.07	1.23

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	J	77	PRO	C-N-CD	-20.77	74.89	120.60
6	J	77	PRO	C-N-CA	13.80	179.95	122.00
2	C	148	LYS	N-CA-C	-7.88	89.72	111.00
3	D	51	ILE	CG1-CB-CG2	-7.34	95.25	111.40
2	C	413	THR	C-N-CD	6.17	141.36	128.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	229	LYS	Peptide
2	C	254	PHE	Peptide
2	C	774	PRO	Peptide
3	D	578	ARG	Peptide
3	D	600	GLN	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	1724	0	1768	28	0
1	B	1775	0	1809	27	0
2	C	8585	0	8508	156	0
3	D	9881	0	9949	214	0
4	E	649	0	645	18	0
5	F	2508	0	2525	44	0
6	J	881	0	861	54	0
7	H	454	0	251	12	0
7	O	634	0	350	4	0
8	G	362	0	206	1	0
8	P	526	0	296	0	0
9	D	2	0	0	0	0
10	D	1	0	0	0	0
All	All	27982	0	27168	509	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 509 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:D:1089:PHE:CE1	3:D:1109:GLN:NE2	1.91	1.38
6:J:76:LYS:HB2	6:J:77:PRO:CD	1.67	1.23
3:D:1089:PHE:CD1	3:D:1109:GLN:NE2	2.09	1.20
5:F:261:GLN:HG2	6:J:82:TRP:CZ2	1.78	1.17
6:J:76:LYS:HB2	6:J:77:PRO:HD3	1.16	1.15

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	224/347 (65%)	206 (92%)	18 (8%)	0	100	100
1	B	235/347 (68%)	205 (87%)	30 (13%)	0	100	100
2	C	1109/1179 (94%)	1001 (90%)	106 (10%)	2 (0%)	49	79
3	D	1261/1326 (95%)	1171 (93%)	83 (7%)	7 (1%)	27	62
4	E	81/110 (74%)	71 (88%)	9 (11%)	1 (1%)	14	48
5	F	317/531 (60%)	307 (97%)	9 (3%)	1 (0%)	43	74
6	J	106/111 (96%)	91 (86%)	11 (10%)	4 (4%)	3	22
All	All	3333/3951 (84%)	3052 (92%)	266 (8%)	15 (0%)	35	65

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	904	ARG
3	D	905	ALA
3	D	909	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	J	78	PRO
3	D	1100	SER

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	195/297 (66%)	167 (86%)	28 (14%)	3	17
1	B	197/297 (66%)	164 (83%)	33 (17%)	2	11
2	C	932/997 (94%)	816 (88%)	116 (12%)	5	22
3	D	1043/1103 (95%)	919 (88%)	124 (12%)	6	23
4	E	69/89 (78%)	56 (81%)	13 (19%)	1	8
5	F	262/429 (61%)	233 (89%)	29 (11%)	7	26
6	J	93/97 (96%)	67 (72%)	26 (28%)	0	2
All	All	2791/3309 (84%)	2422 (87%)	369 (13%)	8	19

5 of 369 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1050	SER
3	D	275	GLU
5	F	495	VAL
2	C	1070	GLU
3	D	86	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	920	HIS
3	D	329	GLN
5	F	516	HIS
2	C	1062	GLN
3	D	368	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 3 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.