



wwPDB EM Model Validation Summary Report ⓘ

May 18, 2020 – 11:02 AM EDT

PDB ID : 6N4C
EMDB ID : EMD-0340
Title : EM structure of the DNA wrapping in bacterial open transcription initiation complex
Authors : Florez-Ariza, A.; Cassago, A.; de Oliveira, P.S.L.; Guerra, D.G.; Portugal, R.V.
Deposited on : 2018-11-19
Resolution : 17.00 Å (reported)
Based on initial model : 4YLN

This is a wwPDB EM 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.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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) : 2.10.1

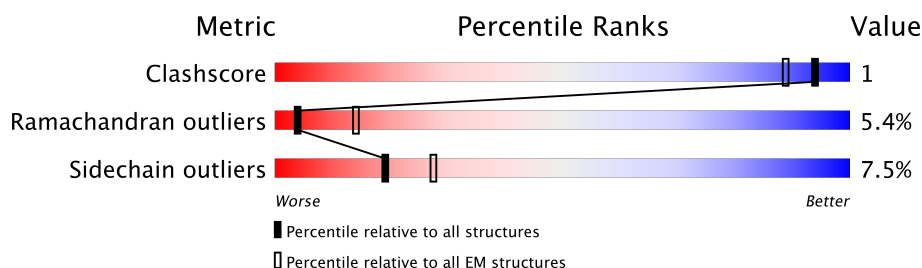
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 17.00 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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 on the bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	F	558	80% 15% . .
2	C	1341	84% 14% .
3	D	1358	85% 12% .
4	A	316	78% 19% . .
5	B	310	76% 19% . .
6	E	90	70% 26% .
7	a	94	12% 57% 31%
8	b	94	15% 51% 34%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 68774 atoms, of which 33722 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 RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	F	558	Total	C	H	N	O	S	0	0
			9018	2799	4523	794	875	27		

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	?	-	GLU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	ALA	deletion	UNP P00579
F	?	-	PRO	deletion	UNP P00579
F	?	-	THR	deletion	UNP P00579
F	?	-	ALA	deletion	UNP P00579
F	?	-	THR	deletion	UNP P00579
F	?	-	HIS	deletion	UNP P00579
F	?	-	VAL	deletion	UNP P00579
F	?	-	GLY	deletion	UNP P00579
F	?	-	SER	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	SER	deletion	UNP P00579
F	?	-	GLN	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	?	-	ASP	deletion	UNP P00579
F	?	-	GLY	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	SER	deletion	UNP P00579
F	?	-	ALA	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579
F	?	-	ILE	deletion	UNP P00579
F	?	-	ALA	deletion	UNP P00579
F	?	-	LYS	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	PRO	deletion	UNP P00579
F	521	ILE	-	insertion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	GLU	deletion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	PRO	deletion	UNP P00579
F	?	-	LEU	deletion	UNP P00579
F	?	-	ASP	deletion	UNP P00579

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

Mol	Chain	Residues	Atoms						AltConf	Trace
2	C	1341	Total	C	H	N	O	S	0	0
			21172	6636	10595	1842	2056	43		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
3	D	1358	Total	C	H	N	O	S	0	0
			21296	6615	10758	1882	1991	50		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	GLU	deletion	UNP A0A369F490
D	?	-	GLY	deletion	UNP A0A369F490
D	?	-	LYS	deletion	UNP A0A369F490
D	?	-	GLU	deletion	UNP A0A369F490

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

Mol	Chain	Residues	Atoms						AltConf	Trace
4	A	316	Total	C	H	N	O	S	0	0
			5001	1545	2528	436	484	8		

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms						AltConf	Trace
5	B	310	Total	C	H	N	O	S	0	0
			4883	1508	2471	422	475	7		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	255	LEU	ARG	conflict	UNP P0A7Z4

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms						AltConf	Trace
6	E	90	Total	C	H	N	O	S	0	0
			1431	430	722	136	142	1		

- Molecule 7 is a DNA chain called DNA (94-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
7	a	94	Total	C	H	N	O	P	0	0
			2995	923	1064	349	566	93		

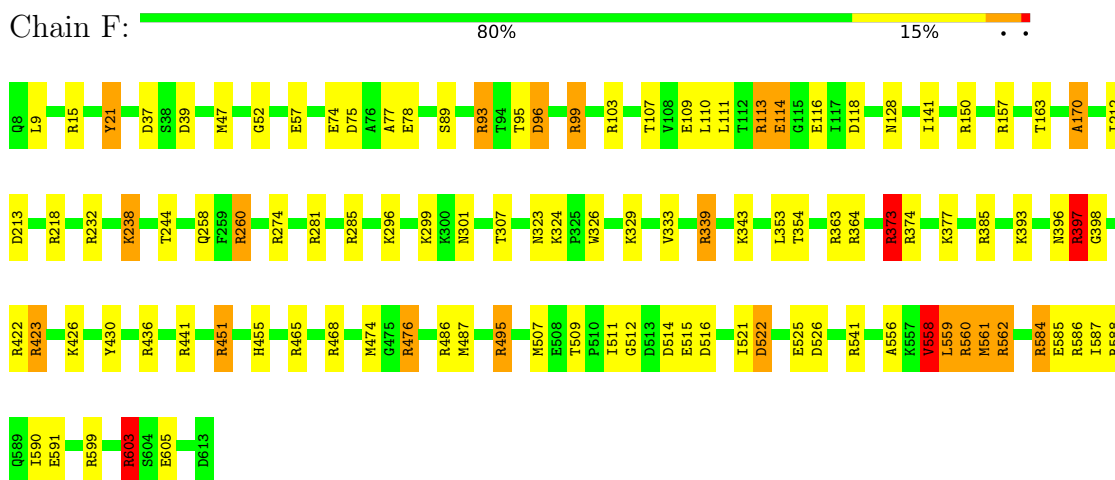
- Molecule 8 is a DNA chain called DNA (94-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
8	b	94	Total	C	H	N	O	P	0	0
			2978	917	1061	349	558	93		

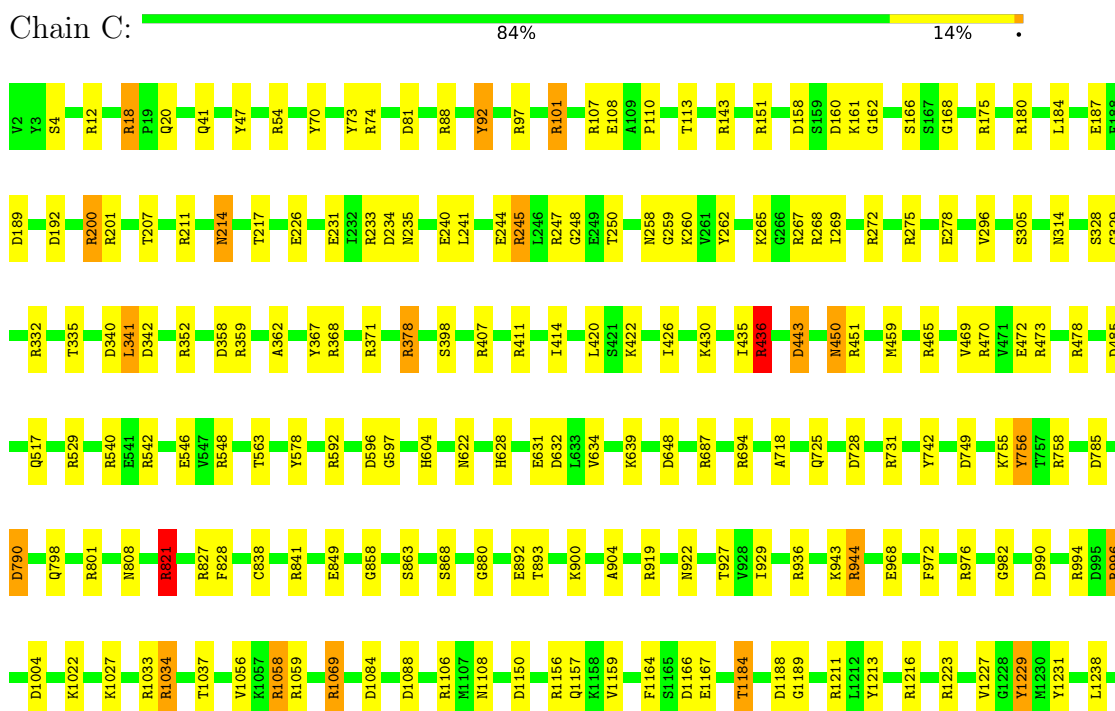
3 Residue-property plots

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: RNA polymerase sigma factor RpoD



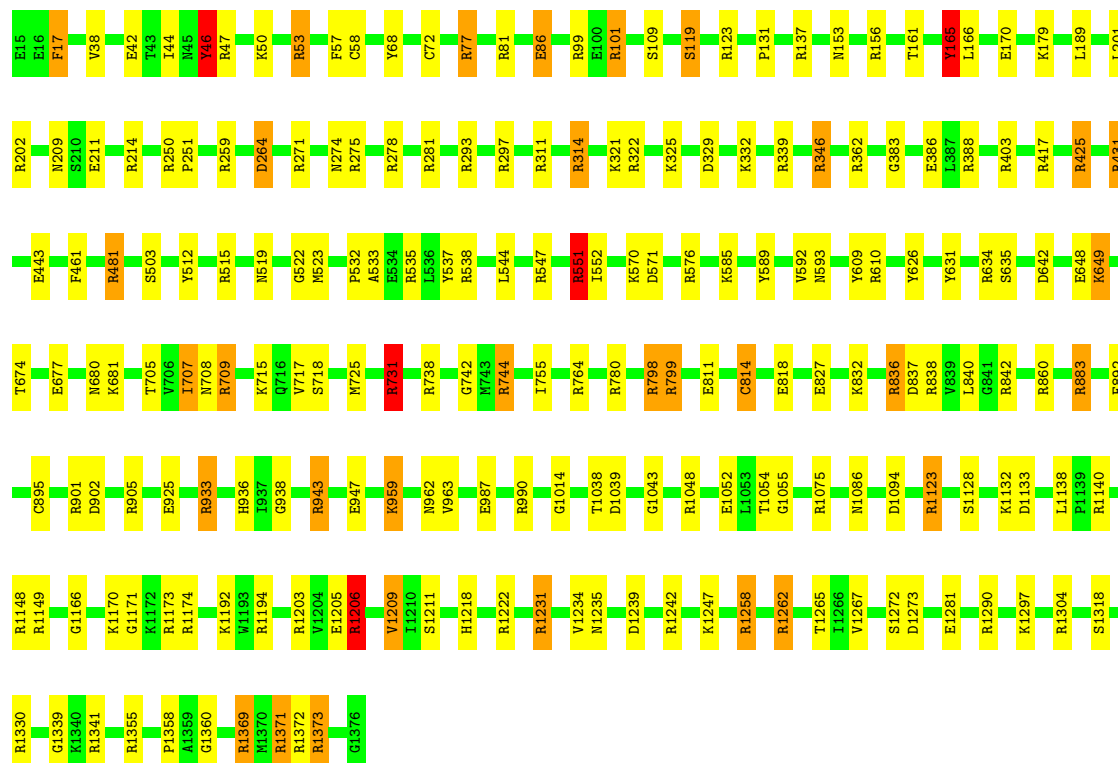
- Molecule 2: DNA-directed RNA polymerase subunit beta





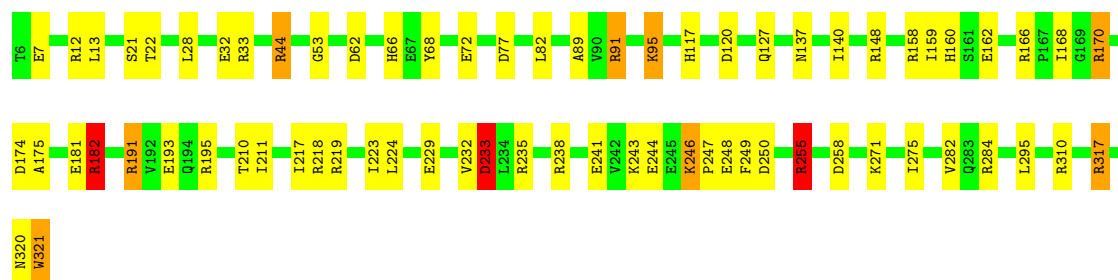
• Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain D: 85% 12% .



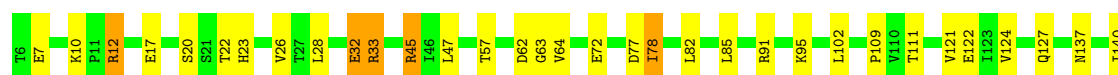
• Molecule 4: DNA-directed RNA polymerase subunit alpha

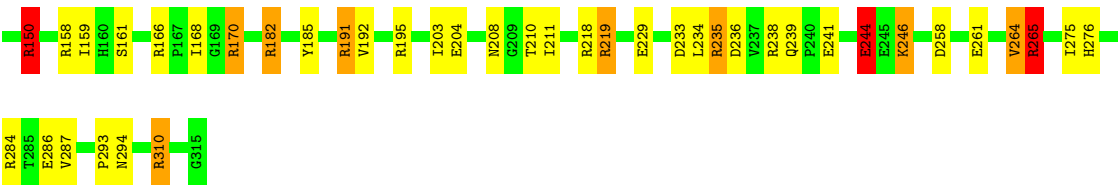
Chain A: 78% 19% . .



• Molecule 5: DNA-directed RNA polymerase subunit alpha

Chain B: 76% 19% . .

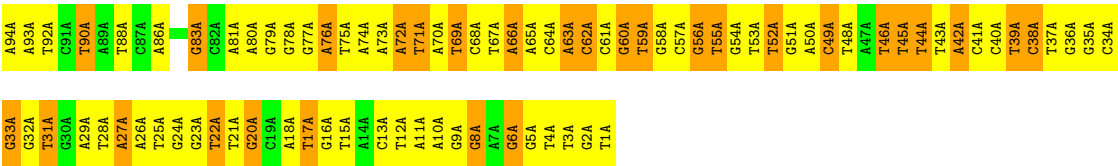
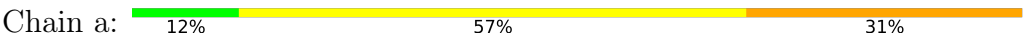




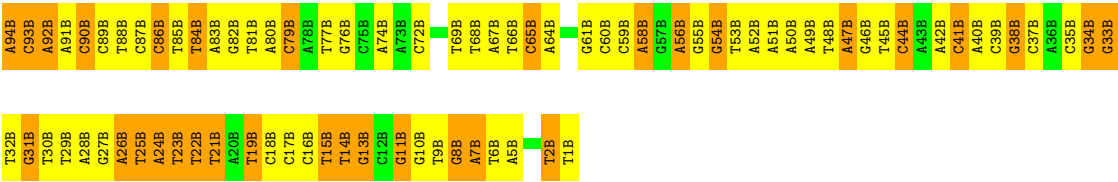
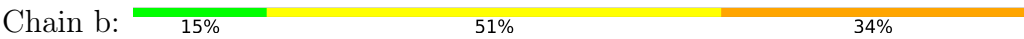
• Molecule 6: DNA-directed RNA polymerase subunit omega



• Molecule 7: DNA (94-MER)



• Molecule 8: DNA (94-MER)



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	16015	Depositor
Resolution determination method	FSC 1/2 BIT CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	JEOL 2100	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	20	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	TVIPS TEMCAM-F416 (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	F	0.80	4/4553 (0.1%)	1.26	41/6127 (0.7%)
2	C	0.73	0/10746	1.18	76/14499 (0.5%)
3	D	0.72	0/10699	1.20	91/14447 (0.6%)
4	A	0.74	0/2505	1.26	20/3392 (0.6%)
5	B	0.74	0/2442	1.30	18/3308 (0.5%)
6	E	0.73	0/711	1.76	10/956 (1.0%)
7	a	1.26	0/2166	1.97	118/3344 (3.5%)
8	b	1.78	16/2150 (0.7%)	2.41	148/3314 (4.5%)
All	All	0.88	20/35972 (0.1%)	1.40	522/49387 (1.1%)

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	F	0	19
2	C	0	25
3	D	0	29
4	A	0	10
5	B	0	10
6	E	0	4
7	a	0	45
8	b	0	44
All	All	0	186

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	b	5(B)	DA	N7-C5	42.03	1.64	1.39
8	b	49(B)	DA	C2-N3	-23.01	1.12	1.33
8	b	50(B)	DA	P-O5'	14.36	1.74	1.59
8	b	92(B)	DA	O3'-P	10.89	1.74	1.61

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	b	51(B)	DA	O3'-P	10.04	1.73	1.61

The worst 5 of 522 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	E	78	ALA	O-C-N	30.92	172.18	122.70
8	b	47(B)	DA	O4'-C1'-N9	30.91	129.63	108.00
8	b	91(B)	DA	O5'-P-OP2	-30.64	73.94	110.70
8	b	47(B)	DA	N1-C6-N6	-18.52	107.49	118.60
8	b	50(B)	DA	P-O3'-C3'	17.60	140.82	119.70

There are no chirality outliers.

5 of 186 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	157	ARG	Sidechain
1	F	170	ALA	Peptide
1	F	232	ARG	Sidechain
1	F	93	ARG	Sidechain
1	F	99	ARG	Sidechain

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	F	4495	4523	4521	2	0
2	C	10577	10595	10591	6	0
3	D	10538	10758	10755	6	0
4	A	2473	2528	2525	1	0
5	B	2412	2471	2468	2	0
6	E	709	722	719	5	0
7	a	1931	1064	1065	0	0
8	b	1917	1061	1061	0	0
All	All	35052	33722	33705	22	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 1.

The worst 5 of 22 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:B:264:VAL:HG12	5:B:265:ARG:H	1.56	0.71
2:C:546:GLU:H	2:C:546:GLU:CD	2.11	0.53
2:C:1184:THR:HG23	2:C:1189:GLY:HA2	1.91	0.53
6:E:38:LEU:H	6:E:59:ILE:HD13	1.76	0.51
4:A:232:VAL:HG22	4:A:233:ASP:H	1.78	0.49

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	F	556/558 (100%)	421 (76%)	103 (18%)	32 (6%)	2	23
2	C	1339/1341 (100%)	1085 (81%)	193 (14%)	61 (5%)	2	27
3	D	1356/1358 (100%)	1089 (80%)	203 (15%)	64 (5%)	2	27
4	A	314/316 (99%)	223 (71%)	71 (23%)	20 (6%)	1	21
5	B	308/310 (99%)	202 (66%)	73 (24%)	33 (11%)	0	9
6	E	88/90 (98%)	71 (81%)	12 (14%)	5 (6%)	2	23
All	All	3961/3973 (100%)	3091 (78%)	655 (16%)	215 (5%)	4	24

5 of 215 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	9	LEU
1	F	37	ASP
1	F	89	SER
1	F	96	ASP
1	F	110	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	F	491/491 (100%)	448 (91%)	43 (9%)	11	37
2	C	1156/1156 (100%)	1083 (94%)	73 (6%)	20	50
3	D	1132/1132 (100%)	1067 (94%)	65 (6%)	23	53
4	A	276/276 (100%)	243 (88%)	33 (12%)	5	24
5	B	270/270 (100%)	239 (88%)	31 (12%)	6	25
6	E	74/74 (100%)	65 (88%)	9 (12%)	5	24
All	All	3399/3399 (100%)	3145 (92%)	254 (8%)	19	45

5 of 254 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1338	GLU
3	D	649	LYS
5	B	211	ILE
3	D	44	ILE
3	D	264	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	446	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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.