



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 19, 2016 – 12:01 PM EST

PDB ID : 5M5Y
EMDB ID: : EMD-3448
Title : RNA Polymerase I elongation complex 2
Authors : Tafur, L.; Sadian, Y.; Hoffmann, N.A.; Jakobi, A.J.; Wetzels, R.; Hagen, W.J.H.; Sachse, C.; Muller, C.W.
Deposited on : 2016-10-23
Resolution : 4.00 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

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

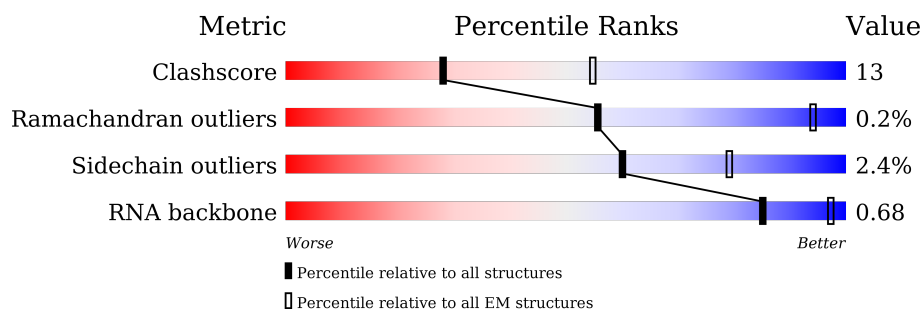
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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.






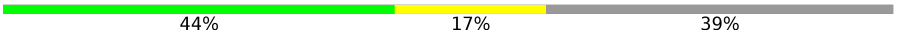

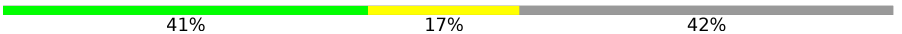



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	1664	62% 25% • 12%
2	B	1203	64% 32% • •
3	C	335	68% 22% • 9%
4	D	137	33% 9% • 57%
5	E	215	74% 25%
6	F	155	49% 15% • 35%
7	G	326	45% 16% 38%
8	H	146	62% 29% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	
10	J	70	
11	K	142	
12	L	70	
13	M	415	
14	N	233	
15	R	10	
16	S	70	
17	T	70	

2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 34104 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 I subunit RPA190.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1466	Total	C	N	O	S	0	0
			11571	7309	2012	2188	62		

- Molecule 2 is a protein called DNA-directed RNA polymerase I subunit RPA135.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1170	Total	C	N	O	S	0	0
			9301	5888	1625	1737	51		

- Molecule 3 is a protein called DNA-directed RNA polymerases I and III subunit RPAC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	304	Total	C	N	O	S	0	0
			2418	1536	414	460	8		

- Molecule 4 is a protein called DNA-directed RNA polymerase I subunit RPA14.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	59	Total	C	N	O	0	0
			467	293	80	94		

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	214	Total	C	N	O	S	0	0
			1751	1111	309	320	11		

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	100	Total	C	N	O	S	0	0
			823	522	144	154	3		

- Molecule 7 is a protein called DNA-directed RNA polymerase I subunit RPA43.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	202	Total	C	N	O	S	0	0
			1600	1026	276	293	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	134	Total	C	N	O	S	0	0
			1075	677	182	212	4		

- Molecule 9 is a protein called DNA-directed RNA polymerase I subunit RPA12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	64	Total	C	N	O	S	0	0
			472	295	78	95	4		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	69	Total	C	N	O	S	0	0
			569	362	101	100	6		

- Molecule 11 is a protein called DNA-directed RNA polymerases I and III subunit RPAC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	100	Total	C	N	O	S	0	0
			784	491	128	160	5		

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	43	Total	C	N	O	S	0	0
			344	211	69	60	4		

- Molecule 13 is a protein called DNA-directed RNA polymerase I subunit RPA49.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	101	Total	C	N	O		0	0
			802	508	132	162			

- Molecule 14 is a protein called DNA-directed RNA polymerase I subunit RPA34.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	135	Total	C	N	O	S	0	0
			1070	685	175	206	4		

- Molecule 15 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	R	7	Total	C	N	O	P	0	0
			145	65	22	51	7		

- Molecule 16 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	S	19	Total	C	N	O	P	0	0
			397	188	82	108	19		

- Molecule 17 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	T	25	Total	C	N	O	P	0	0
			509	244	86	154	25		

- Molecule 18 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

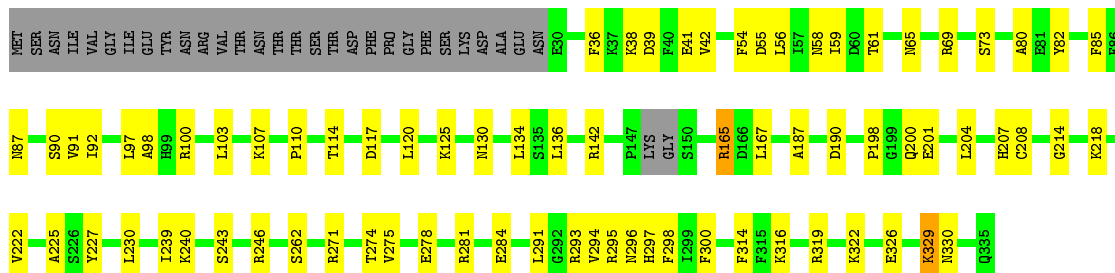
- Chain A:  62% 25% 12%



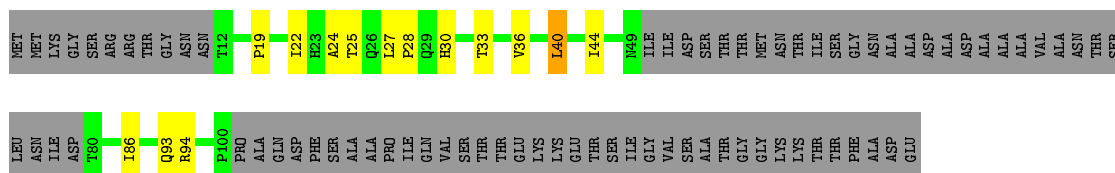




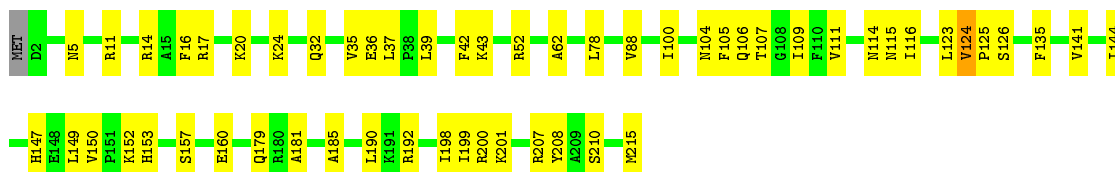
- Molecule 3: DNA-directed RNA polymerases I and III subunit RPA1



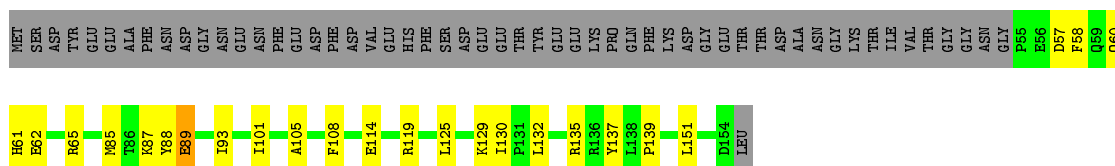
- Molecule 4: DNA-directed RNA polymerase I subunit RPA14



- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

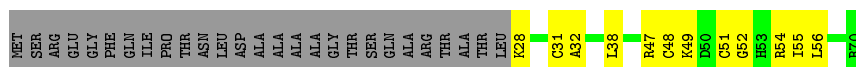


- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



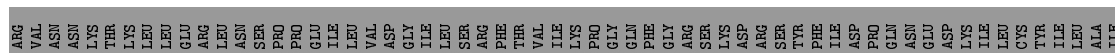
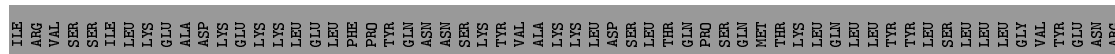
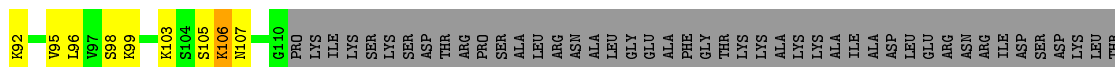
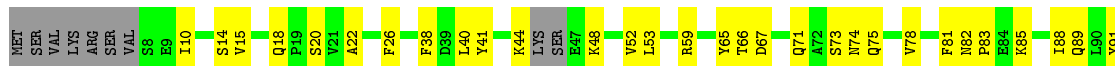
- Molecule 7: DNA-directed RNA polymerase I subunit RPA43





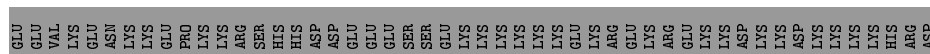
• Molecule 13: DNA-directed RNA polymerase I subunit RPA49

Chain M: 15% 9% 76%



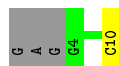
• Molecule 14: DNA-directed RNA polymerase I subunit RPA34

Chain N: 41% 17% 42%



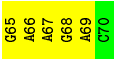
• Molecule 15: RNA

Chain R: 60% 10% 30%

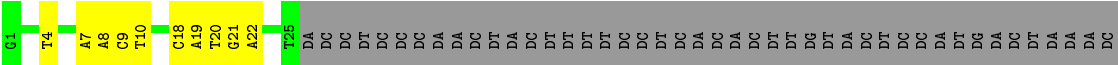


• Molecule 16: Non-template DNA

Chain S: 11% 16% 73%



- Molecule 17: Template DNA



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	50784	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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

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	0.25	1/11782 (0.0%)	0.47	6/15913 (0.0%)
10	J	0.22	0/578	0.44	0/775
11	K	0.48	2/794 (0.3%)	0.81	4/1070 (0.4%)
12	L	0.23	0/346	0.44	0/457
13	M	0.26	0/816	0.48	0/1095
14	N	0.25	0/1090	0.49	0/1466
15	R	0.10	0/160	0.69	0/246
16	S	0.48	0/448	0.83	0/690
17	T	0.49	0/568	0.96	0/874
2	B	0.24	1/9506 (0.0%)	0.46	3/12847 (0.0%)
3	C	0.24	0/2469	0.43	0/3347
4	D	0.23	0/473	0.44	0/641
5	E	0.24	0/1787	0.40	0/2406
6	F	0.23	0/838	0.39	0/1129
7	G	0.23	0/1637	0.44	0/2226
8	H	0.24	0/1093	0.47	0/1480
9	I	0.24	0/478	0.46	0/647
All	All	0.26	4/34863 (0.0%)	0.49	13/47309 (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	1
11	K	0	1
2	B	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	K	57	ASP	N-CA	10.21	1.66	1.46
1	A	763	GLY	N-CA	6.49	1.55	1.46
2	B	1064	LYS	N-CA	5.14	1.56	1.46
11	K	57	ASP	CA-CB	5.07	1.65	1.53

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	445	GLY	N-CA-C	-14.36	77.20	113.10
11	K	57	ASP	N-CA-C	-14.09	72.96	111.00
2	B	1064	LYS	N-CA-C	-12.01	78.58	111.00
11	K	57	ASP	CB-CA-C	-11.69	87.02	110.40
1	A	763	GLY	N-CA-C	-9.02	90.55	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	445	GLY	Mainchain
2	B	877	SER	Peptide
11	K	58	GLY	Mainchain

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	11571	0	11649	340	0
2	B	9301	0	9194	283	0
3	C	2418	0	2401	51	0
4	D	467	0	468	11	0
5	E	1751	0	1776	36	0
6	F	823	0	841	19	0
7	G	1600	0	1600	33	0
8	H	1075	0	1046	28	0
9	I	472	0	473	15	0
10	J	569	0	585	25	0
11	K	784	0	781	20	0
12	L	344	0	363	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	M	802	0	787	34	0
14	N	1070	0	1085	33	0
15	R	145	0	75	1	0
16	S	397	0	213	8	0
17	T	509	0	285	9	0
18	A	2	0	0	0	0
18	B	1	0	0	0	0
18	I	1	0	0	0	0
18	J	1	0	0	0	0
18	L	1	0	0	0	0
All	All	34104	0	33622	846	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 13.

The worst 5 of 846 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:450:LYS:O	1:A:451:VAL:HG22	1.31	1.26
1:A:865:ASP:C	1:A:866:LYS:HD2	1.63	1.18
1:A:865:ASP:HB3	1:A:866:LYS:NZ	1.59	1.16
1:A:450:LYS:O	1:A:451:VAL:CG2	2.04	1.05
1:A:865:ASP:CB	1:A:866:LYS:NZ	2.29	0.96

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	1450/1664 (87%)	1364 (94%)	80 (6%)	6 (0%)	39 79
2	B	1160/1203 (96%)	1094 (94%)	64 (6%)	2 (0%)	52 86

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	300/335 (90%)	289 (96%)	11 (4%)	0	100	100
4	D	55/137 (40%)	54 (98%)	1 (2%)	0	100	100
5	E	212/215 (99%)	208 (98%)	4 (2%)	0	100	100
6	F	98/155 (63%)	94 (96%)	4 (4%)	0	100	100
7	G	196/326 (60%)	184 (94%)	12 (6%)	0	100	100
8	H	130/146 (89%)	125 (96%)	5 (4%)	0	100	100
9	I	62/125 (50%)	59 (95%)	3 (5%)	0	100	100
10	J	67/70 (96%)	62 (92%)	5 (8%)	0	100	100
11	K	98/142 (69%)	92 (94%)	6 (6%)	0	100	100
12	L	41/70 (59%)	40 (98%)	1 (2%)	0	100	100
13	M	97/415 (23%)	90 (93%)	7 (7%)	0	100	100
14	N	127/233 (54%)	105 (83%)	22 (17%)	0	100	100
All	All	4093/5236 (78%)	3860 (94%)	225 (6%)	8 (0%)	56	86

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	451	VAL
2	B	812	ALA
1	A	227	LEU
1	A	248	PHE
1	A	765	LEU

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	1293/1465 (88%)	1264 (98%)	29 (2%)	60	84
2	B	1025/1053 (97%)	999 (98%)	26 (2%)	55	82
3	C	269/296 (91%)	266 (99%)	3 (1%)	80	91
4	D	56/116 (48%)	54 (96%)	2 (4%)	42	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	196/197 (100%)	193 (98%)	3 (2%)	72	89
6	F	90/137 (66%)	88 (98%)	2 (2%)	60	84
7	G	180/291 (62%)	174 (97%)	6 (3%)	45	77
8	H	117/128 (91%)	110 (94%)	7 (6%)	24	63
9	I	56/110 (51%)	54 (96%)	2 (4%)	42	76
10	J	64/65 (98%)	61 (95%)	3 (5%)	32	70
11	K	89/130 (68%)	88 (99%)	1 (1%)	80	91
12	L	38/57 (67%)	37 (97%)	1 (3%)	54	81
13	M	91/371 (24%)	90 (99%)	1 (1%)	80	91
14	N	125/220 (57%)	124 (99%)	1 (1%)	86	93
All	All	3689/4636 (80%)	3602 (98%)	87 (2%)	60	82

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	377	MET
2	B	959	THR
10	J	39	LEU
2	B	495	ARG
2	B	658	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	587	GLN
2	B	724	GLN
11	K	102	ASN
2	B	682	GLN
2	B	710	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	R	6/10 (60%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 6 ligands modelled in this entry, 6 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.