



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Feb 13, 2017 – 01:48 PM EST

PDB ID : 5U4W
EMDB ID: : EMD-8508
Title : Cryo-EM Structure of Immature Zika Virus
Authors : Mangala Prasad, V.; Miller, A.S.; Klose, T.; Sirohi, D.; Buda, G.; Jiang, W.;
Kuhn, R.J.; Rossmann, M.G.
Deposited on : 2016-12-06
Resolution : 9.10 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 : 1.7.1 (RC1), CSD as537be (2016)
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

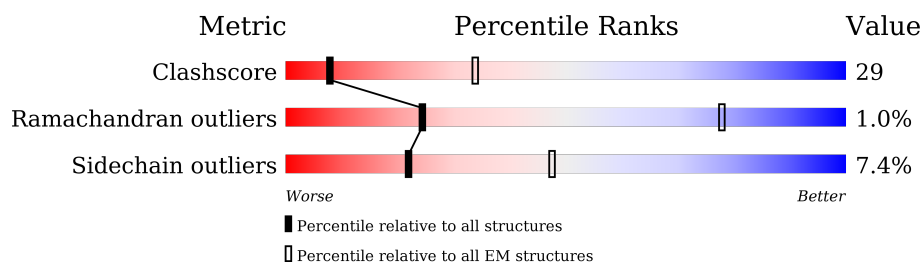
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 9.10 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	402	59% 33% . . .
1	C	402	57% 36% . . .
1	E	402	59% 34% . . .
2	B	81	52% 41% 6% .
2	D	81	53% 40% 6% .
2	F	81	54% 40% 5% .
3	G	66	89% 6% 5%
3	I	66	86% 9% 5%
3	K	66	86% 9% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	53	 85% 15%
4	J	53	 85% 15%
4	L	53	 85% 15%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	A	401	-	-	X	-
5	NAG	B	104	X	-	X	-
5	NAG	D	104	X	-	X	-
5	NAG	F	104	X	-	X	-
6	BMA	B	103	X	-	-	-
6	BMA	D	103	X	-	-	-
6	BMA	F	103	X	-	-	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13368 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 E protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	390	Total	C	N	O	S	0	0
			3046	1926	524	569	27		
1	C	390	Total	C	N	O	S	0	0
			3046	1926	524	569	27		
1	E	390	Total	C	N	O	S	0	0
			3046	1926	524	569	27		

- Molecule 2 is a protein called pr domain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	81	Total	C	N	O	S	0	0
			640	396	106	128	10		
2	D	81	Total	C	N	O	S	0	0
			640	396	106	128	10		
2	F	81	Total	C	N	O	S	0	0
			640	396	106	128	10		

- Molecule 3 is a protein called Protein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	G	66	Total	C	N	O	0	0
			375	233	71	71		
3	I	66	Total	C	N	O	0	0
			375	233	71	71		
3	K	66	Total	C	N	O	0	0
			375	233	71	71		

- Molecule 4 is a protein called M protein.

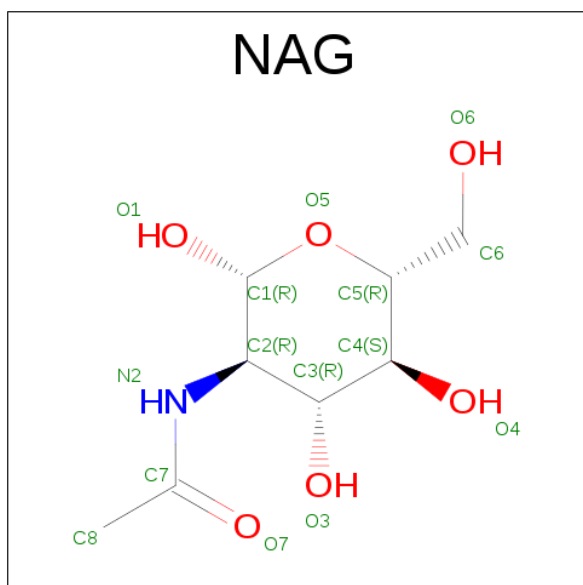
Mol	Chain	Residues	Atoms				AltConf	Trace
4	H	53	Total	C	N	O	0	0
			315	198	57	60		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
4	J	53	Total	C	N	O	0	0
			315	198	57	60		
4	L	53	Total	C	N	O	0	0
			315	198	57	60		

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



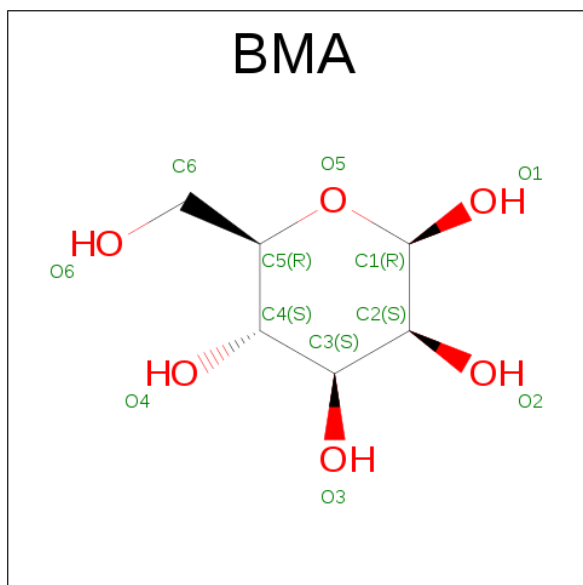
Mol	Chain	Residues	Atoms				AltConf
5	A	1	Total	C	N	O	0
			14	8	1	5	
5	B	1	Total	C	N	O	0
			44	24	3	17	
5	B	1	Total	C	N	O	0
			44	24	3	17	
5	B	1	Total	C	N	O	0
			44	24	3	17	
5	C	1	Total	C	N	O	0
			14	8	1	5	
5	D	1	Total	C	N	O	0
			44	24	3	17	
5	D	1	Total	C	N	O	0
			44	24	3	17	
5	D	1	Total	C	N	O	0
			44	24	3	17	
5	E	1	Total	C	N	O	0
			14	8	1	5	

Continued on next page...

Continued from previous page...

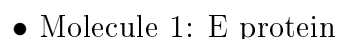
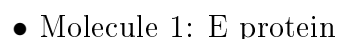
Mol	Chain	Residues	Atoms				AltConf
5	F	1	Total	C	N	O	0
			44	24	3	17	
5	F	1	Total	C	N	O	0
			44	24	3	17	
5	F	1	Total	C	N	O	0
			44	24	3	17	

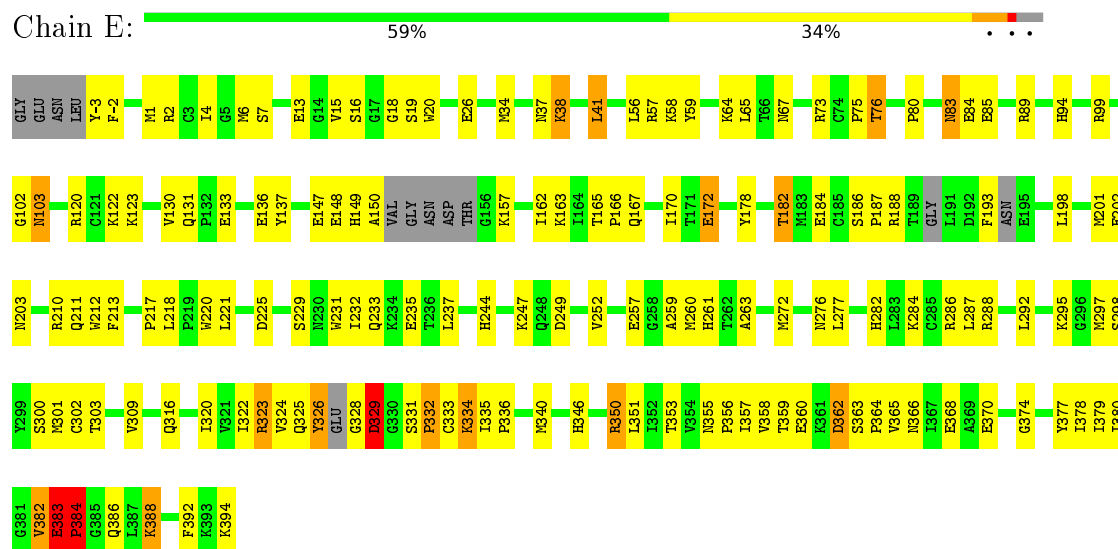
- Molecule 6 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).



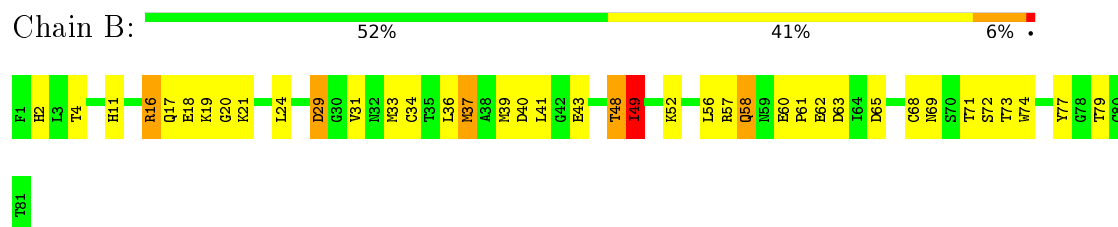
Mol	Chain	Residues	Atoms			AltConf
6	B	1	Total	C	O	0
			22	12	10	
6	B	1	Total	C	O	0
			22	12	10	
6	D	1	Total	C	O	0
			22	12	10	
6	D	1	Total	C	O	0
			22	12	10	
6	F	1	Total	C	O	0
			22	12	10	
6	F	1	Total	C	O	0
			22	12	10	

- Molecule 1: E protein

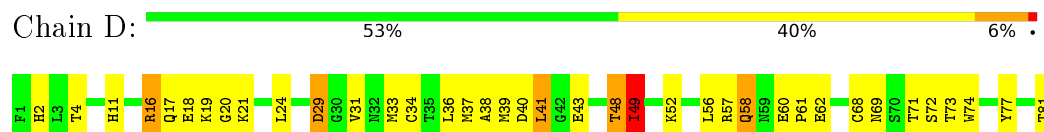




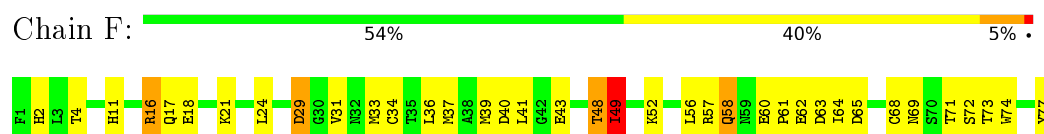
- Molecule 2: pr domain



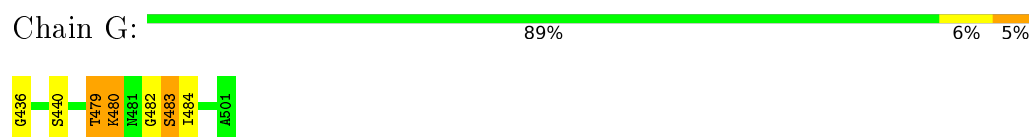
- Molecule 2: pr domain



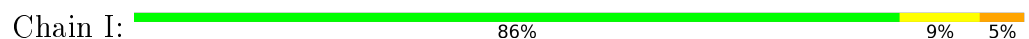
- Molecule 2: pr domain



- Molecule 3: Protein E

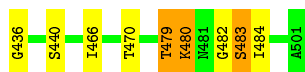
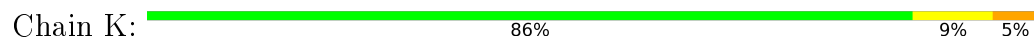


- Molecule 3: Protein E

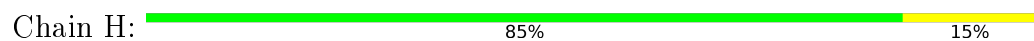




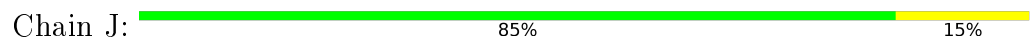
● Molecule 3: Protein E



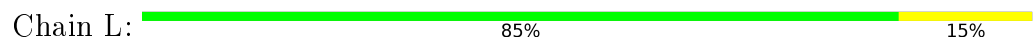
● Molecule 4: M protein



● Molecule 4: M protein



● Molecule 4: M protein



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	9315	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: BMA, NAG

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	2/3106 (0.1%)	0.78	7/4186 (0.2%)
1	C	0.54	2/3106 (0.1%)	0.78	7/4186 (0.2%)
1	E	0.54	2/3106 (0.1%)	0.78	7/4186 (0.2%)
2	B	0.61	2/652 (0.3%)	0.69	1/883 (0.1%)
2	D	0.61	2/652 (0.3%)	0.69	1/883 (0.1%)
2	F	0.61	2/652 (0.3%)	0.69	1/883 (0.1%)
3	G	0.37	0/379	0.48	0/518
3	I	0.37	0/379	0.48	0/518
3	K	0.37	0/379	0.48	0/518
4	H	0.40	0/321	0.46	0/448
4	J	0.40	0/321	0.46	0/448
4	L	0.40	0/321	0.46	0/448
All	All	0.53	12/13374 (0.1%)	0.73	24/18105 (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	B	0	1
2	D	0	1
2	F	0	1
All	All	0	3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	335	ILE	C-N	-15.26	1.05	1.34
1	E	335	ILE	C-N	-15.24	1.05	1.34
1	C	335	ILE	C-N	-15.24	1.05	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	334	LYS	C-N	-13.75	1.02	1.34
1	C	334	LYS	C-N	-13.75	1.02	1.34

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	334	LYS	O-C-N	-9.86	106.93	122.70
1	C	334	LYS	O-C-N	-9.84	106.96	122.70
1	A	334	LYS	O-C-N	-9.83	106.97	122.70
1	A	384	PRO	N-CA-C	-8.40	90.27	112.10
1	C	384	PRO	N-CA-C	-8.39	90.30	112.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	49	ILE	Mainchain
2	D	49	ILE	Mainchain
2	F	49	ILE	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	3046	0	3045	145	0
1	C	3046	0	3044	228	0
1	E	3046	0	3045	167	0
2	B	640	0	607	119	0
2	D	640	0	603	118	0
2	F	640	0	608	86	0
3	G	375	0	264	5	0
3	I	375	0	264	6	0
3	K	375	0	264	6	0
4	H	315	0	225	6	0
4	J	315	0	225	6	0
4	L	315	0	225	6	0
5	A	14	0	11	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	44	0	42	18	0
5	C	14	0	11	6	0
5	D	44	0	42	17	0
5	E	14	0	11	6	0
5	F	44	0	42	17	0
6	B	22	0	17	5	0
6	D	22	0	17	6	0
6	F	22	0	17	6	0
All	All	13368	0	12629	756	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 29.

The worst 5 of 756 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)
2:B:40:ASP:CB	2:D:39:MET:HB3	1.29	1.56
1:C:108:PHE:CD1	2:F:63:ASP:HA	1.43	1.52
2:B:40:ASP:HA	2:D:39:MET:CB	1.46	1.46
5:F:101:NAG:O4	6:F:102:BMA:C1	1.66	1.42
1:C:108:PHE:HD1	2:F:63:ASP:CA	1.32	1.42

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	380/402 (94%)	344 (90%)	33 (9%)	3 (1%)	24	69
1	C	380/402 (94%)	345 (91%)	32 (8%)	3 (1%)	24	69
1	E	380/402 (94%)	345 (91%)	32 (8%)	3 (1%)	24	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	79/81 (98%)	70 (89%)	9 (11%)	0	100	100
2	D	79/81 (98%)	70 (89%)	9 (11%)	0	100	100
2	F	79/81 (98%)	70 (89%)	9 (11%)	0	100	100
3	G	64/66 (97%)	50 (78%)	11 (17%)	3 (5%)	3	32
3	I	64/66 (97%)	50 (78%)	11 (17%)	3 (5%)	3	32
3	K	64/66 (97%)	50 (78%)	11 (17%)	3 (5%)	3	32
4	H	51/53 (96%)	47 (92%)	4 (8%)	0	100	100
4	J	51/53 (96%)	47 (92%)	4 (8%)	0	100	100
4	L	51/53 (96%)	47 (92%)	4 (8%)	0	100	100
All	All	1722/1806 (95%)	1535 (89%)	169 (10%)	18 (1%)	24	65

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	ASP
1	A	383	GLU
1	A	384	PRO
1	C	329	ASP
1	C	383	GLU

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	339/348 (97%)	316 (93%)	23 (7%)	20	57
1	C	339/348 (97%)	316 (93%)	23 (7%)	20	57
1	E	339/348 (97%)	316 (93%)	23 (7%)	20	57
2	B	75/75 (100%)	65 (87%)	10 (13%)	5	28
2	D	75/75 (100%)	65 (87%)	10 (13%)	5	28
2	F	75/75 (100%)	65 (87%)	10 (13%)	5	28
3	G	16/50 (32%)	16 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	I	16/50 (32%)	16 (100%)	0	100	100
3	K	16/50 (32%)	16 (100%)	0	100	100
4	H	16/43 (37%)	16 (100%)	0	100	100
4	J	16/43 (37%)	16 (100%)	0	100	100
4	L	16/43 (37%)	16 (100%)	0	100	100
All	All	1338/1548 (86%)	1239 (93%)	99 (7%)	22	54

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	300	SER
1	C	388	LYS
2	F	29	ASP
1	C	326	TYR
1	C	346	HIS

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

Mol	Chain	Res	Type
1	C	200	GLN
1	C	316	GLN
1	E	366	ASN
1	C	248	GLN
2	D	7	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 ⓘ

18 ligands 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$
5	NAG	A	401	-	14,14,15	0.51	0	15,19,21	4.83	2 (13%)
5	NAG	B	101	-	14,14,15	1.79	1 (7%)	15,19,21	1.60	2 (13%)
6	BMA	B	102	-	11,11,12	0.88	1 (9%)	15,15,17	0.89	1 (6%)
6	BMA	B	103	-	11,11,12	1.06	1 (9%)	15,15,17	2.64	4 (26%)
5	NAG	B	104	-	15,15,15	0.39	0	17,21,21	0.64	0
5	NAG	B	105	-	15,15,15	0.47	0	17,21,21	0.52	0
5	NAG	C	401	-	14,14,15	0.50	0	15,19,21	4.82	2 (13%)
5	NAG	D	101	-	14,14,15	1.77	1 (7%)	15,19,21	1.60	2 (13%)
6	BMA	D	102	-	11,11,12	0.87	1 (9%)	15,15,17	0.89	1 (6%)
6	BMA	D	103	-	11,11,12	1.05	1 (9%)	15,15,17	2.64	4 (26%)
5	NAG	D	104	-	15,15,15	0.40	0	17,21,21	0.64	0
5	NAG	D	105	-	15,15,15	0.46	0	17,21,21	0.51	0
5	NAG	E	401	-	14,14,15	0.51	0	15,19,21	4.83	2 (13%)
5	NAG	F	101	-	14,14,15	1.77	1 (7%)	15,19,21	1.60	2 (13%)
6	BMA	F	102	-	11,11,12	0.87	1 (9%)	15,15,17	0.90	1 (6%)
6	BMA	F	103	-	11,11,12	1.06	1 (9%)	15,15,17	2.64	4 (26%)
5	NAG	F	104	-	15,15,15	0.40	0	17,21,21	0.64	0
5	NAG	F	105	-	15,15,15	0.46	0	17,21,21	0.52	0

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
5	NAG	A	401	-	-	0/6/23/26	0/1/1/1
5	NAG	B	101	-	-	0/6/23/26	0/1/1/1
6	BMA	B	102	-	-	0/2/19/22	0/1/1/1
6	BMA	B	103	-	1/1/4/5	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	B	104	-	1/1/6/7	0/6/26/26	0/1/1/1
5	NAG	B	105	-	-	0/6/26/26	0/1/1/1
5	NAG	C	401	-	-	0/6/23/26	0/1/1/1
5	NAG	D	101	-	-	0/6/23/26	0/1/1/1
6	BMA	D	102	-	-	0/2/19/22	0/1/1/1
6	BMA	D	103	-	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	D	104	-	1/1/6/7	0/6/26/26	0/1/1/1
5	NAG	D	105	-	-	0/6/26/26	0/1/1/1
5	NAG	E	401	-	-	0/6/23/26	0/1/1/1
5	NAG	F	101	-	-	0/6/23/26	0/1/1/1
6	BMA	F	102	-	-	0/2/19/22	0/1/1/1
6	BMA	F	103	-	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	F	104	-	1/1/6/7	0/6/26/26	0/1/1/1
5	NAG	F	105	-	-	0/6/26/26	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	101	NAG	C2-N2	-6.26	1.35	1.46
5	F	101	NAG	C2-N2	-6.20	1.35	1.46
5	D	101	NAG	C2-N2	-6.19	1.35	1.46
6	B	103	BMA	C6-C5	-2.39	1.43	1.51
6	F	103	BMA	C6-C5	-2.38	1.43	1.51

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	401	NAG	O4-C4-C3	-8.54	91.11	110.36
5	A	401	NAG	O4-C4-C3	-8.53	91.12	110.36
5	C	401	NAG	O4-C4-C3	-8.53	91.12	110.36
5	D	101	NAG	O7-C7-N2	2.48	126.91	121.84
5	B	101	NAG	O7-C7-N2	2.49	126.91	121.84

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	B	103	BMA	C5
6	F	103	BMA	C5
5	D	104	NAG	C1
6	D	103	BMA	C5
5	B	104	NAG	C1

There are no torsion outliers.

There are no ring outliers.

18 monomers are involved in 76 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	401	NAG	7	0
5	B	101	NAG	4	0
6	B	102	BMA	1	0
6	B	103	BMA	4	0
5	B	104	NAG	12	0
5	B	105	NAG	6	0
5	C	401	NAG	6	0
5	D	101	NAG	4	0
6	D	102	BMA	2	0
6	D	103	BMA	4	0
5	D	104	NAG	11	0
5	D	105	NAG	6	0
5	E	401	NAG	6	0
5	F	101	NAG	4	0
6	F	102	BMA	2	0
6	F	103	BMA	4	0
5	F	104	NAG	11	0
5	F	105	NAG	6	0

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	D	2
1	E	2
2	B	2
1	C	2
1	A	2
2	F	2

The worst 5 of 12 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	49:ILE	C	50:THR	N	1.17
1	D	49:ILE	C	50:THR	N	1.17
1	F	49:ILE	C	50:THR	N	1.17
1	B	48:THR	C	49:ILE	N	1.12
1	D	48:THR	C	49:ILE	N	1.12