



wwPDB EM Map/Model Validation Report ⓘ

May 22, 2016 – 11:05 PM EDT

PDB ID : 5IZ7
EMDB ID: : EMD-8139
Title : Cryo-EM structure of thermally stable Zika virus strain H/PF/2013
Authors : Kostyuchenko, V.A.; Zhang, S.; Fibriansah, G.; Lok, S.M.
Deposited on : 2016-03-25
Resolution : 3.70 Å(reported)
Based on PDB ID : 3J27

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

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-20027457

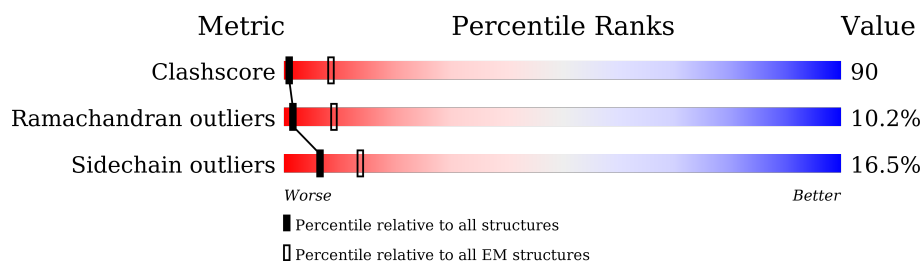
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.70 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	504	
1	B	504	
1	C	504	
2	D	75	
2	E	75	
2	F	75	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13268 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.

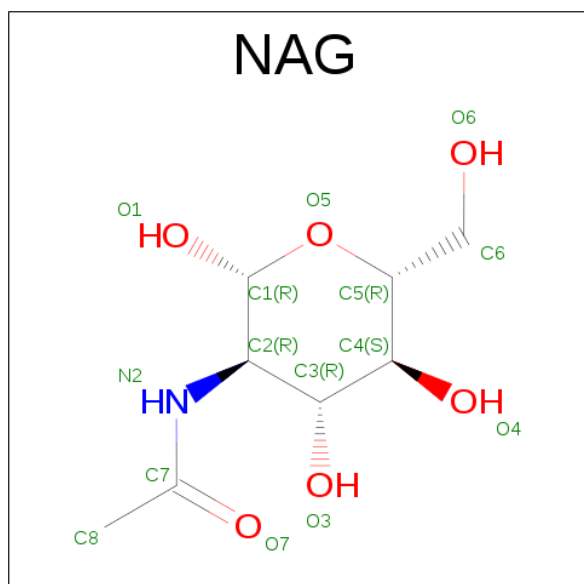
- Molecule 1 is a protein called structural protein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	504	Total	C	N	O	S	0	0
			3809	2395	660	723	31		
1	C	504	Total	C	N	O	S	0	0
			3808	2395	660	722	31		
1	B	504	Total	C	N	O	S	0	0
			3809	2395	660	723	31		

- Molecule 2 is a protein called structural protein M.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	D	75	Total	C	N	O	S	0	0
			600	391	105	103	1		
2	E	75	Total	C	N	O	S	0	0
			600	391	105	103	1		
2	F	75	Total	C	N	O	S	0	0
			600	391	105	103	1		

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

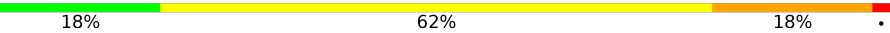


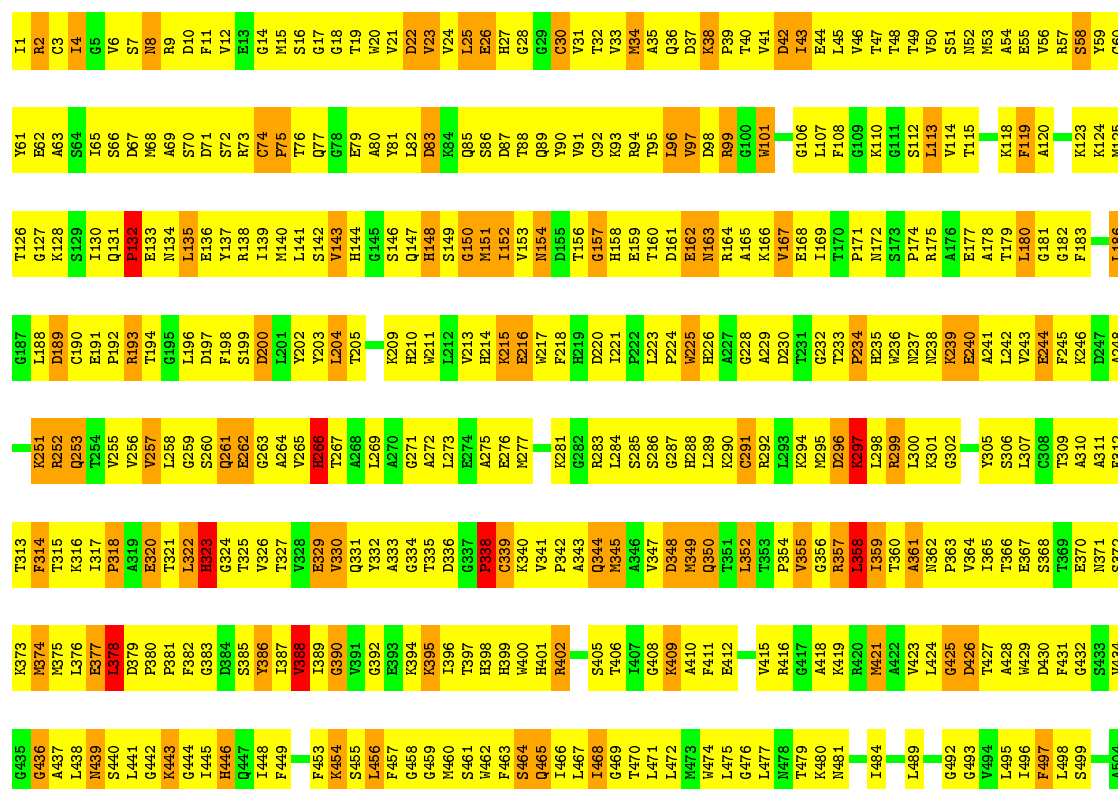
Mol	Chain	Residues	Atoms				AltConf
3	A	1	Total	C	N	O	0
			14	8	1	5	
3	C	1	Total	C	N	O	0
			14	8	1	5	
3	B	1	Total	C	N	O	0
			14	8	1	5	

3 Residue-property plots

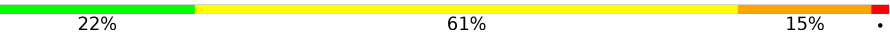
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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.

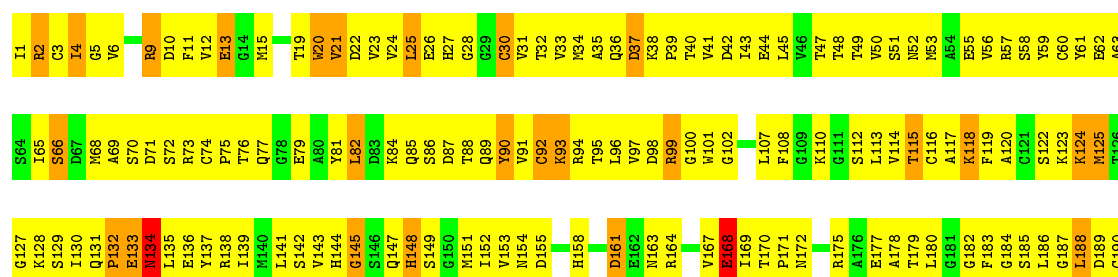
• Molecule 1: structural protein E

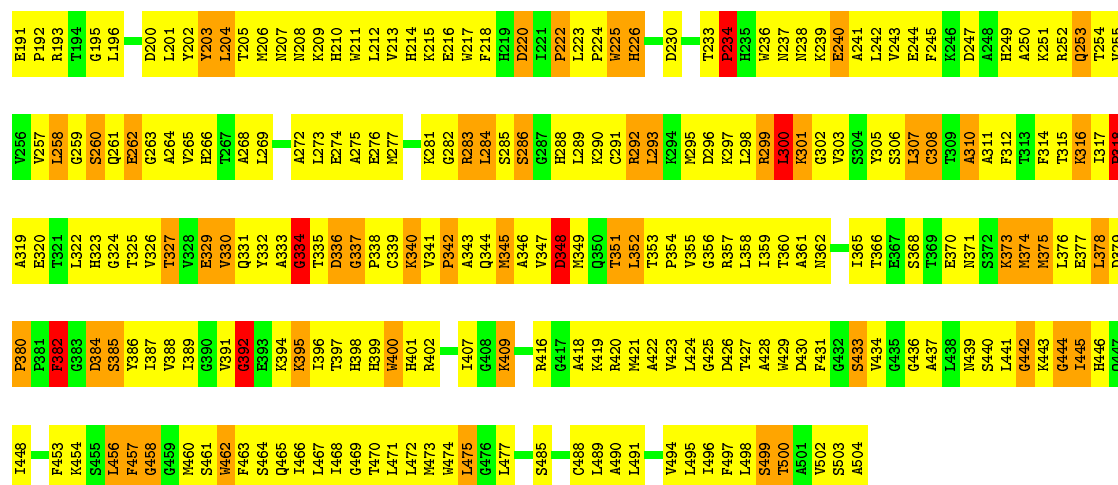
Chain A: 



• Molecule 1: structural protein E

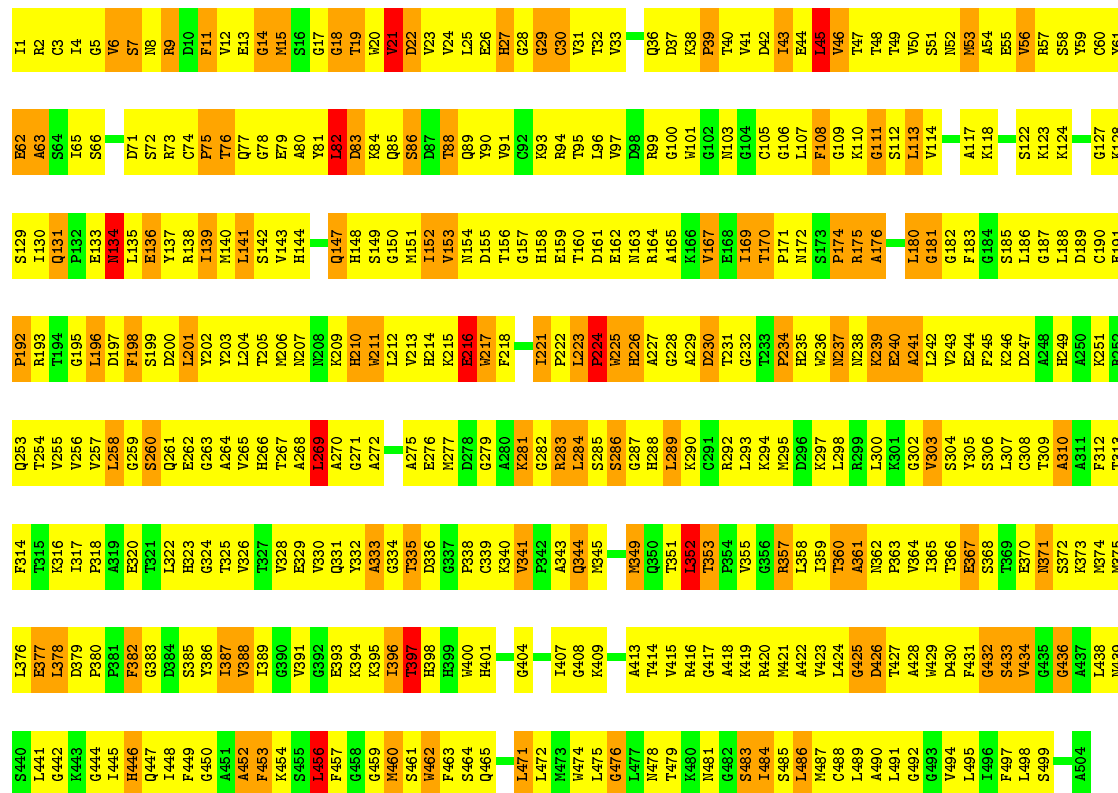
Chain C: 





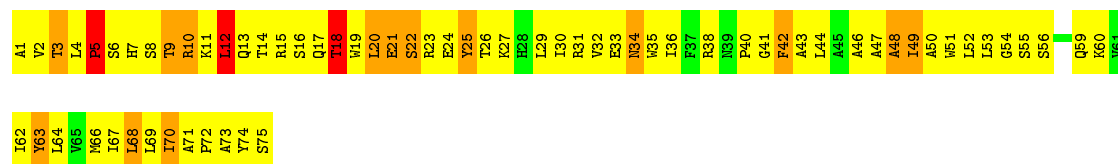
• Molecule 1: structural protein E

Chain B: 19% 59% 20%



• Molecule 2: structural protein M

Chain D: 11% 67% 19%



- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| V62 | Y63 | M66 | L67 | L68 | L69 | I70 | A71 | F72 | A73 | Y74 | S75 | A1 | V2 | T3 | L4 | P5 | S6 | H7 | S8 | T9 | R10 | K11 | L12 | Q13 | T14 | R15 | S16 | S17 | Q18 | T19 | R20 | S21 | S22 | R23 | E24 | S25 | T26 | K27 | H28 | L29 | I30 | R31 | V32 | S33 | N34 | S35 | L36 | F37 | R38 | G41 | F42 | A43 | L44 | A45 | A46 | A47 | A48 | I49 | A50 | W51 | L52 | L53 | G54 | S55 | S56 | T57 | S58 | Q59 | M60 | V61 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	7180	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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.60	0/3888	0.87	5/5261 (0.1%)
1	B	0.65	0/3888	0.93	9/5261 (0.2%)
1	C	0.66	1/3887 (0.0%)	0.91	8/5261 (0.2%)
2	D	0.61	0/615	0.82	1/838 (0.1%)
2	E	0.77	0/615	0.86	0/838
2	F	0.66	0/615	0.85	1/838 (0.1%)
All	All	0.64	1/13508 (0.0%)	0.90	24/18297 (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	F	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	225	TRP	CB-CG	-8.14	1.35	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	334	GLY	N-CA-C	6.91	130.37	113.10
1	B	224	PRO	N-CA-C	-6.73	94.61	112.10
1	B	82	LEU	N-CA-C	6.66	128.98	111.00
1	C	348	ASP	N-CA-C	6.48	128.49	111.00
1	B	456	LEU	CA-CB-CG	6.47	130.18	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	F	74	TYR	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	A	3809	0	3761	684	0
1	B	3809	0	3760	770	0
1	C	3808	0	3760	682	0
2	D	600	0	624	165	0
2	E	600	0	624	119	0
2	F	600	0	624	124	0
3	A	14	0	13	4	0
3	B	14	0	13	1	0
3	C	14	0	13	5	0
All	All	13268	0	13192	2386	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 90.

The worst 5 of 2386 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:C:68:MET:SD	1:C:255:VAL:HB	1.79	1.21
1:B:345:MET:HE3	1:B:380:PRO:HA	1.23	1.16
1:C:178:ALA:HB3	1:C:186:LEU:H	1.08	1.15
1:A:158:HIS:HB3	1:A:166:LYS:HE3	1.20	1.15
1:B:60:CYS:HB3	1:B:224:PRO:HG2	1.21	1.14

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	502/504 (100%)	332 (66%)	123 (24%)	47 (9%)	1	14
1	B	502/504 (100%)	331 (66%)	104 (21%)	67 (13%)	0	6
1	C	502/504 (100%)	365 (73%)	96 (19%)	41 (8%)	1	17
2	D	73/75 (97%)	37 (51%)	26 (36%)	10 (14%)	0	6
2	E	73/75 (97%)	48 (66%)	20 (27%)	5 (7%)	1	23
2	F	73/75 (97%)	43 (59%)	24 (33%)	6 (8%)	1	17
All	All	1725/1737 (99%)	1156 (67%)	393 (23%)	176 (10%)	1	12

5 of 176 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	36	GLN
1	A	132	PRO
1	A	322	LEU
1	A	323	HIS
1	A	357	ARG

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	A	410/410 (100%)	335 (82%)	75 (18%)	2	14
1	B	410/410 (100%)	343 (84%)	67 (16%)	3	21
1	C	410/410 (100%)	349 (85%)	61 (15%)	4	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	64/64 (100%)	53 (83%)	11 (17%)	2	18
2	E	64/64 (100%)	52 (81%)	12 (19%)	2	13
2	F	64/64 (100%)	55 (86%)	9 (14%)	4	29
All	All	1422/1422 (100%)	1187 (84%)	235 (16%)	6	20

5 of 235 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	292	ARG
1	C	409	LYS
2	E	10	ARG
1	C	299	ARG
1	C	340	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	237	ASN
1	B	27	HIS
2	E	7	HIS
1	C	261	GLN
1	A	207	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 ⓘ

3 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	$\# Z > 2$
3	NAG	A	601	1	14,14,15	1.00	0	15,19,21	0.95	0
3	NAG	B	601	1	14,14,15	0.64	0	15,19,21	0.69	0
3	NAG	C	601	1	14,14,15	1.11	2 (14%)	15,19,21	0.86	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
3	NAG	A	601	1	-	0/6/23/26	0/1/1/1
3	NAG	B	601	1	-	0/6/23/26	0/1/1/1
3	NAG	C	601	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	601	NAG	C3-C2	2.17	1.57	1.52
3	C	601	NAG	C1-C2	3.01	1.56	1.52

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	NAG	4	0
3	B	601	NAG	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	601	NAG	5	0

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.