



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:09 PM BST

PDB ID : 5FT2  
EMDB ID: : EMD-3292  
Title : Sub-tomogram averaging of Lassa virus glycoprotein spike from virus- like particles at pH 5  
Authors : Li, S.; Zhaoyang, S.; Pryce, R.; Parsy, M.L.; Fehling, S.K.; Schlie, K.; Siebert, C.A.; Garten, W.; Bowden, T.A.; Strecker, T.; Huiskonen, J.T.  
Deposited on : 2016-01-09  
Resolution : 16.40 Å(reported)  
Based on PDB ID : 4ZJF

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](mailto: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) : trunk27241

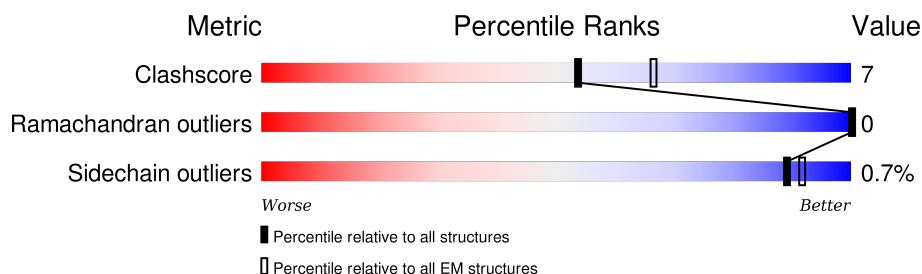
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 16.40 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  $\geq 3$ , 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	B	172	 78% 13% • 9%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1340 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 PRE-GLYCOPROTEIN POLYPROTEIN GP COMPLEX.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	157	Total	C	N	O	S	0	0
			1233	770	213	237	13		

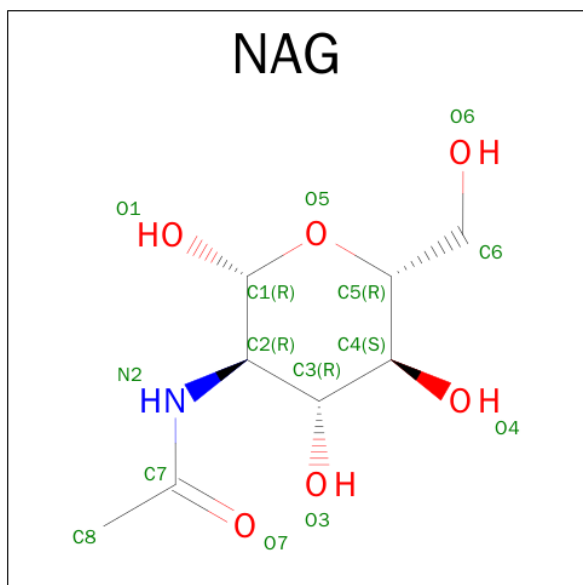
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	66	SER	-	EXPRESSION TAG	UNP P08669
B	67	HIS	-	EXPRESSION TAG	UNP P08669
B	68	HIS	-	EXPRESSION TAG	UNP P08669
B	69	HIS	-	EXPRESSION TAG	UNP P08669
B	70	HIS	-	EXPRESSION TAG	UNP P08669
B	71	HIS	-	EXPRESSION TAG	UNP P08669
B	72	HIS	-	EXPRESSION TAG	UNP P08669
B	73	GLY	-	EXPRESSION TAG	UNP P08669
B	74	GLY	-	EXPRESSION TAG	UNP P08669

- Molecule 2 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				AltConf
2	B	2	Total	C	N	O	0
			56	32	4	20	
2	B	2	Total	C	N	O	0
			56	32	4	20	

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				AltConf
3	B	1	Total	C	N	O	0
			28	16	2	10	
3	B	1	Total	C	N	O	0
			28	16	2	10	

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	B	23	Total	O	0
			23	23	

### 3 Residue-property plots [i](#)

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: PRE-GLYCOPROTEIN POLYPROTEIN GP COMPLEX

Chain B: 



## 4 Experimental information

Property	Value	Source
Reconstruction method	TOMOGRAPHY	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH TILTED IMAGE	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	Depositor
Minimum defocus (nm)	2800	Depositor
Maximum defocus (nm)	6700	Depositor
Magnification	160000	Depositor
Image detector	GATAN K2 SUMMIT (4K X 4K)	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	B	0.27	0/1264	0.46	0/1713

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	B	1233	0	1147	17	0
2	B	56	0	50	0	0
3	B	28	0	26	0	0
4	B	23	0	0	2	0
All	All	1340	0	1223	17	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 7.

The worst 5 of 17 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:B:89:ASN:ND2	4:B:403:HOH:O	2.27	0.67
1:B:151:GLU:HG3	1:B:184:ALA:HB3	1.87	0.56
1:B:206:GLY:N	4:B:407:HOH:O	2.38	0.56
1:B:151:GLU:OE2	1:B:170:HIS:NE2	2.37	0.53
1:B:192:MET:HE1	1:B:213:ILE:HD12	1.94	0.50

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	B	155/172 (90%)	151 (97%)	4 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	B	137/150 (91%)	136 (99%)	1 (1%)	88	94

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	112	ILE



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

4 carbohydrates 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	$\# Z  > 2$
2	NAG	B	301	1,2	14,14,15	0.87	1 (7%)	15,19,21	1.29	2 (13%)
2	NAG	B	302	2	14,14,15	0.38	0	15,19,21	0.47	0
2	NAG	B	305	1,2	14,14,15	0.31	0	15,19,21	0.44	0
2	NAG	B	306	2	14,14,15	0.28	0	15,19,21	0.44	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
2	NAG	B	301	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	302	2	-	0/6/23/26	0/1/1/1
2	NAG	B	305	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	306	2	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	301	NAG	O5-C1	2.75	1.48	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	301	NAG	O4-C4-C5	2.15	114.89	109.23
2	B	301	NAG	C1-O5-C5	3.30	116.99	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

2 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	B	303	1	14,14,15	0.21	0	15,19,21	0.43	0
3	NAG	B	304	1	14,14,15	0.30	0	15,19,21	0.26	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	B	303	1	-	0/6/23/26	0/1/1/1
3	NAG	B	304	1	-	0/6/23/26	0/1/1/1

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.