



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 4, 2017 – 01:32 PM EST

PDB ID : 5WRG
EMDB ID: : EMD-6679
Title : SARS-CoV spike glycoprotein
Authors : Gui, M.; Song, W.; Xiang, Y.; Wang, X.
Deposited on : 2016-12-01
Resolution : 4.30 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

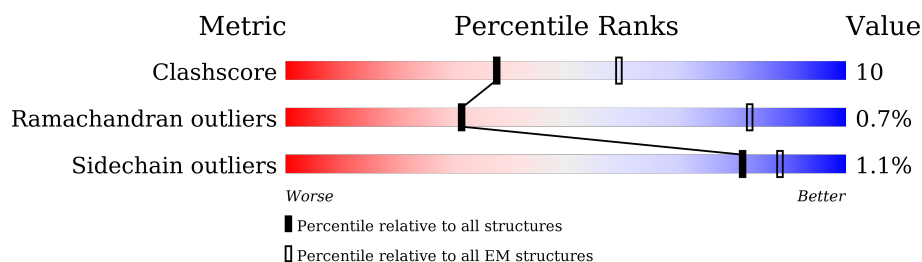
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 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	1203	<div> <div>46%</div> <div>14%</div> <div>.</div> <div>39%</div> </div>
1	B	1203	<div> <div>46%</div> <div>14%</div> <div>.</div> <div>39%</div> </div>
1	C	1203	<div> <div>46%</div> <div>14%</div> <div>.</div> <div>39%</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16317 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 Spike glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	736	Total	C	N	O	S	0	0
			5439	3427	920	1063	29		
1	B	736	Total	C	N	O	S	0	0
			5439	3427	920	1063	29		
1	C	736	Total	C	N	O	S	0	0
			5439	3427	920	1063	29		

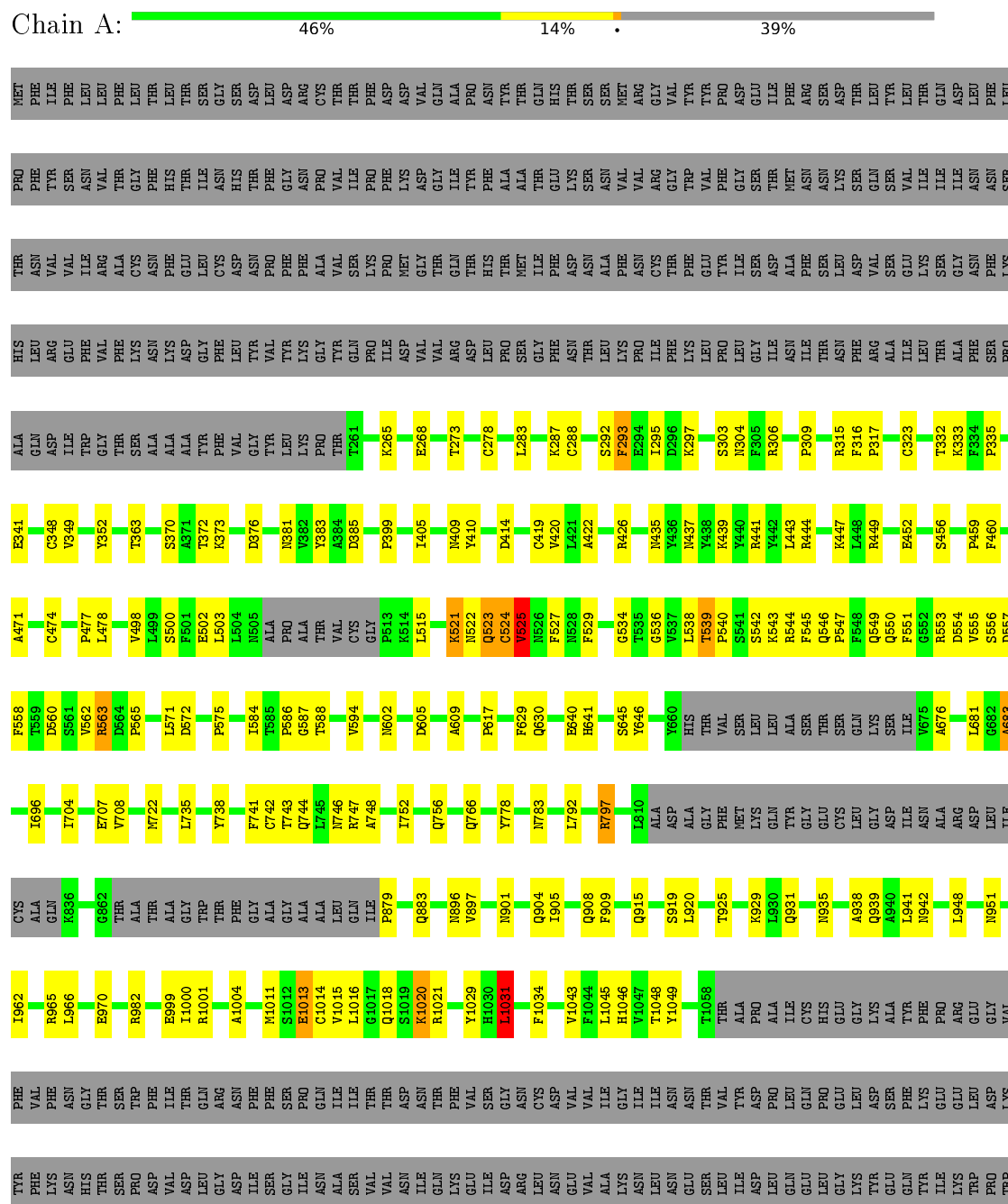
There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	667	ALA	ARG	engineered mutation	UNP P59594
A	1197	SER	-	expression tag	UNP P59594
A	1198	HIS	-	expression tag	UNP P59594
A	1199	PRO	-	expression tag	UNP P59594
A	1200	GLN	-	expression tag	UNP P59594
A	1201	PHE	-	expression tag	UNP P59594
A	1202	GLU	-	expression tag	UNP P59594
A	1203	LYS	-	expression tag	UNP P59594
B	667	ALA	ARG	engineered mutation	UNP P59594
B	1197	SER	-	expression tag	UNP P59594
B	1198	HIS	-	expression tag	UNP P59594
B	1199	PRO	-	expression tag	UNP P59594
B	1200	GLN	-	expression tag	UNP P59594
B	1201	PHE	-	expression tag	UNP P59594
B	1202	GLU	-	expression tag	UNP P59594
B	1203	LYS	-	expression tag	UNP P59594
C	667	ALA	ARG	engineered mutation	UNP P59594
C	1197	SER	-	expression tag	UNP P59594
C	1198	HIS	-	expression tag	UNP P59594
C	1199	PRO	-	expression tag	UNP P59594
C	1200	GLN	-	expression tag	UNP P59594
C	1201	PHE	-	expression tag	UNP P59594
C	1202	GLU	-	expression tag	UNP P59594
C	1203	LYS	-	expression tag	UNP P59594

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: Spike glycoprotein



[illegible]

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	34152	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

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	0.38	0/5536	0.72	8/7535 (0.1%)
1	B	0.38	0/5536	0.72	7/7535 (0.1%)
1	C	0.38	0/5536	0.72	8/7535 (0.1%)
All	All	0.38	0/16608	0.72	23/22605 (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
1	A	0	14
1	B	0	14
1	C	0	14
All	All	0	42

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1031	LEU	CA-CB-CG	6.54	130.34	115.30
1	A	1031	LEU	CA-CB-CG	6.51	130.27	115.30
1	C	1031	LEU	CA-CB-CG	6.51	130.27	115.30
1	C	1013	GLU	C-N-CA	6.02	136.74	121.70
1	A	1013	GLU	C-N-CA	6.00	136.71	121.70

There are no chirality outliers.

5 of 42 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	268	GLU	Peptide
1	A	521	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	522	ASN	Peptide
1	A	523	GLN	Peptide
1	A	524	CYS	Peptide

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	5439	0	5089	113	0
1	B	5439	0	5089	110	0
1	C	5439	0	5089	109	0
All	All	16317	0	15267	310	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 10.

The worst 5 of 310 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:722:MET:SD	1:C:577:SER:OG	2.47	0.73
1:A:1018:GLN:HB2	1:C:1021:ARG:HH22	1.54	0.70
1:C:297:LYS:HA	1:C:586:PRO:HG3	1.73	0.69
1:A:297:LYS:HA	1:A:586:PRO:HG3	1.73	0.69
1:B:297:LYS:HA	1:B:586:PRO:HG3	1.73	0.69

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	726/1203 (60%)	591 (81%)	130 (18%)	5 (1%)	26	71
1	B	726/1203 (60%)	591 (81%)	130 (18%)	5 (1%)	26	71
1	C	726/1203 (60%)	590 (81%)	131 (18%)	5 (1%)	26	71
All	All	2178/3609 (60%)	1772 (81%)	391 (18%)	15 (1%)	31	71

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	742	CYS
1	A	1031	LEU
1	B	742	CYS
1	B	1031	LEU
1	C	742	CYS

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	562/1047 (54%)	556 (99%)	6 (1%)	80	90
1	B	562/1047 (54%)	556 (99%)	6 (1%)	80	90
1	C	562/1047 (54%)	556 (99%)	6 (1%)	80	90
All	All	1686/3141 (54%)	1668 (99%)	18 (1%)	81	90

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	544	ARG
1	B	563	ARG
1	C	544	ARG
1	B	426	ARG
1	B	525	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	746	ASN
1	B	935	ASN
1	C	721	ASN
1	B	437	ASN
1	B	630	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.