



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 11:17 PM BST

PDB ID : 5YY5
Title : Structural definition of a unique neutralization epitope on the receptor-binding domain of MERS-CoV spike glycoprotein
Authors : Zhang, S.; Wang, P.; Zhou, P.; Wang, X.; Zhang, L.
Deposited on : 2017-12-08
Resolution : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.13.1

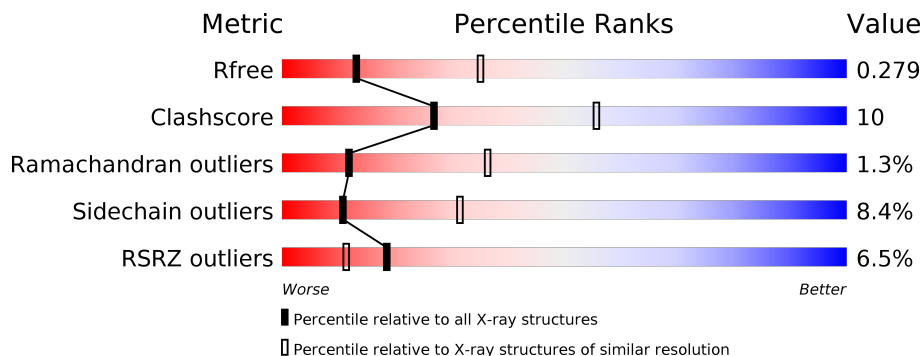
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	209	
1	B	209	
2	H	111	
3	L	112	
4	C	109	
5	D	105	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6592 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 MERS-CoV RBD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	209	1612	1028	257	316	11	0	0	0
1	B	209	1612	1028	257	316	11	0	0	0

- Molecule 2 is a protein called Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	111	878	557	146	171	4	0	0	0

- Molecule 3 is a protein called Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	112	825	512	140	171	2	0	0	0

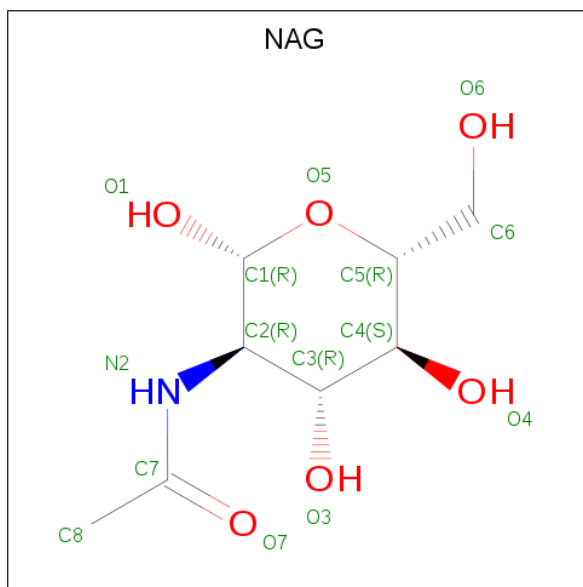
- Molecule 4 is a protein called Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	C	109	860	546	144	166	4	0	0	0

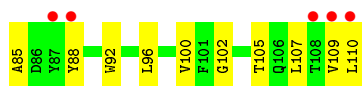
- Molecule 5 is a protein called Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	D	105	777	481	132	162	2	0	0	0

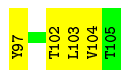
- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



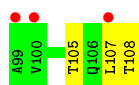
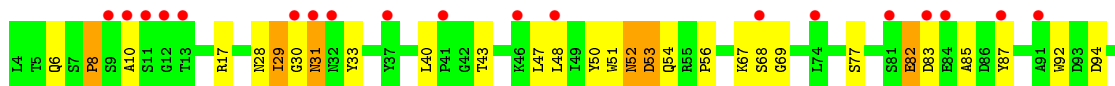
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		



- Molecule 4: Heavy chain



- Molecule 5: Light chain



4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	152.31Å 152.31Å 148.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.51 – 2.80 41.51 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (41.51-2.80) 99.0 (41.51-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 2.61Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.246 , 0.279 0.246 , 0.279	Depositor DCC
R_{free} test set	1927 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	63.6	Xtrriage
Anisotropy	0.830	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 56.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.017 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6592	wwPDB-VP
Average B, all atoms (Å ²)	104.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: 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 > 5$	RMSZ	# $ Z > 5$
1	A	0.58	1/1651 (0.1%)	0.77	2/2254 (0.1%)
1	B	0.48	0/1651	0.71	2/2254 (0.1%)
2	H	0.54	0/898	0.83	1/1221 (0.1%)
3	L	0.47	0/845	0.66	0/1153
4	C	0.37	0/880	0.64	0/1196
5	D	0.36	0/796	0.59	0/1085
All	All	0.49	1/6721 (0.0%)	0.71	5/9163 (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	H	0	1
4	C	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	478	CYS	CB-SG	-5.76	1.72	1.81

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	506	LEU	CA-CB-CG	8.05	133.81	115.30
1	B	535	TRP	CA-CB-CG	-7.60	99.26	113.70
1	A	407	CYS	CA-CB-SG	-5.78	103.59	114.00
1	B	389	LEU	CA-CB-CG	-5.39	102.91	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	94	ASN	C-N-CA	-5.21	108.68	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	C	93	THR	Peptide
2	H	93	THR	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	1612	0	1575	33	0
1	B	1612	0	1576	15	0
2	H	878	0	842	27	0
3	L	825	0	778	30	0
4	C	860	0	825	18	0
5	D	777	0	723	18	0
6	A	14	0	13	5	0
6	B	14	0	12	0	0
All	All	6592	0	6344	132	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 132 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)
5:D:30:GLY:H	5:D:67:LYS:HE3	1.41	0.86
2:H:53:ASP:OD1	2:H:55:SER:HB2	1.79	0.82
5:D:40:LEU:HD23	5:D:85:ALA:HB2	1.64	0.79
1:A:413:LYS:HA	6:A:601:NAG:H82	1.64	0.77
4:C:52:SER:HB3	4:C:55:SER:HB3	1.66	0.77

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 X-ray entries followed by that with respect to entries of similar resolution.

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	207/209 (99%)	198 (96%)	9 (4%)	0	100	100
1	B	207/209 (99%)	199 (96%)	8 (4%)	0	100	100
2	H	109/111 (98%)	99 (91%)	8 (7%)	2 (2%)	8	28
3	L	110/112 (98%)	97 (88%)	8 (7%)	5 (4%)	2	8
4	C	107/109 (98%)	94 (88%)	12 (11%)	1 (1%)	17	46
5	D	103/105 (98%)	92 (89%)	8 (8%)	3 (3%)	4	15
All	All	843/855 (99%)	779 (92%)	53 (6%)	11 (1%)	12	36

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	55	SER
3	L	10	ALA
3	L	53	ASP
4	C	55	SER
5	D	10	ALA

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 X-ray entries followed by that with respect to entries of similar resolution.

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	190/190 (100%)	172 (90%)	18 (10%)	8	25
1	B	190/190 (100%)	177 (93%)	13 (7%)	16	42
2	H	94/94 (100%)	85 (90%)	9 (10%)	8	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	L	90/90 (100%)	85 (94%)	5 (6%)	21	51
4	C	91/91 (100%)	82 (90%)	9 (10%)	8	23
5	D	84/84 (100%)	76 (90%)	8 (10%)	8	25
All	All	739/739 (100%)	677 (92%)	62 (8%)	11	31

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	542	ARG
2	H	70	SER
5	D	52	ASN
2	H	12	VAL
2	H	89	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	436	ASN
1	B	486	HIS
1	B	501	ASN
5	D	54	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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
6	NAG	B	601	1	14,14,15	0.46	0	17,19,21	0.90	1 (5%)
6	NAG	A	601	1	14,14,15	0.92	2 (14%)	17,19,21	0.78	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
6	NAG	B	601	1	-	2/6/23/26	0/1/1/1
6	NAG	A	601	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	601	NAG	O5-C1	-2.51	1.39	1.43
6	A	601	NAG	C1-C2	2.18	1.55	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	601	NAG	C1-O5-C5	3.12	116.42	112.19

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	601	NAG	O5-C5-C6-O6
6	A	601	NAG	C4-C5-C6-O6
6	B	601	NAG	O5-C5-C6-O6
6	B	601	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	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.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å ²)	Q < 0.9
1	A	209/209 (100%)	-0.06	0 100 100	66, 74, 94, 102	0
1	B	209/209 (100%)	0.09	6 (2%) 51 41	74, 96, 131, 139	0
2	H	111/111 (100%)	0.14	2 (1%) 68 61	75, 96, 123, 133	0
3	L	112/112 (100%)	0.35	11 (9%) 7 4	72, 97, 119, 128	0
4	C	109/109 (100%)	0.85	15 (13%) 2 1	126, 145, 175, 185	0
5	D	105/105 (100%)	0.98	22 (20%) 1 0	113, 140, 153, 157	0
All	All	855/855 (100%)	0.30	56 (6%) 18 11	66, 97, 151, 185	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	C	68	THR	5.4
4	C	12	VAL	4.7
5	D	10	ALA	4.5
5	D	100	VAL	4.4
5	D	11	SER	4.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
6	NAG	B	601	14/15	0.76	0.25	96,98,102,103	0
6	NAG	A	601	14/15	0.82	0.28	90,92,95,97	0

6.5 Other polymers [i](#)

There are no such residues in this entry.