



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 8, 2020 – 06:56 PM BST

PDB ID : 4KQZ
Title : structure of the receptor binding domain (RBD) of MERS-CoV spike
Authors : Lu, G.; Hu, Y.; Wang, Q.; Qi, J.; Gao, F.; Li, Y.; Zhang, Y.; Zhang, W.; Yuan, Y.; Bao, J.; Zhang, B.; Shi, Y.; Yan, J.; Gao, G.F.
Deposited on : 2013-05-15
Resolution : 2.51 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

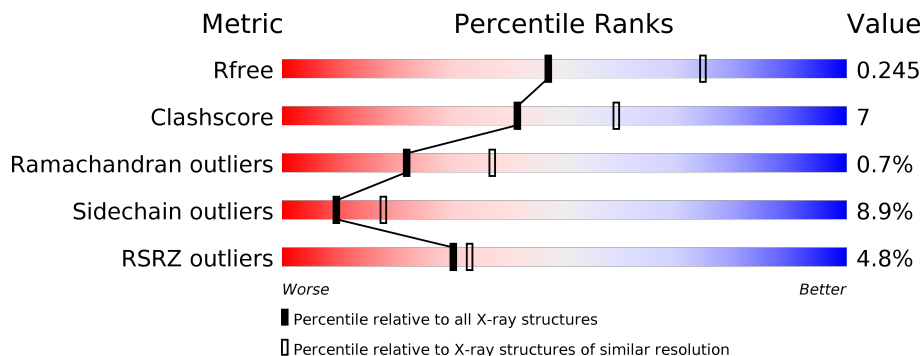
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.51 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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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	251	 8% 67% 16% 17%
1	B	251	 8% 63% 16% 17%

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 3307 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.

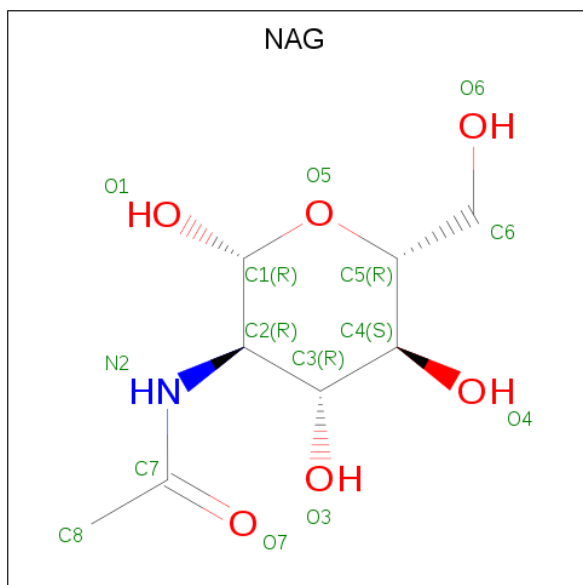
- Molecule 1 is a protein called S protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	208	1608	1026	256	315	11	0	0	0
1	B	208	1608	1026	256	315	11	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	362	ALA	-	expression tag	UNP K0BRG7
A	363	ASP	-	expression tag	UNP K0BRG7
A	364	GLY	-	expression tag	UNP K0BRG7
A	365	ILE	-	expression tag	UNP K0BRG7
A	366	GLN	-	expression tag	UNP K0BRG7
A	607	HIS	-	expression tag	UNP K0BRG7
A	608	HIS	-	expression tag	UNP K0BRG7
A	609	HIS	-	expression tag	UNP K0BRG7
A	610	HIS	-	expression tag	UNP K0BRG7
A	611	HIS	-	expression tag	UNP K0BRG7
A	612	HIS	-	expression tag	UNP K0BRG7
B	362	ALA	-	expression tag	UNP K0BRG7
B	363	ASP	-	expression tag	UNP K0BRG7
B	364	GLY	-	expression tag	UNP K0BRG7
B	365	ILE	-	expression tag	UNP K0BRG7
B	366	GLN	-	expression tag	UNP K0BRG7
B	607	HIS	-	expression tag	UNP K0BRG7
B	608	HIS	-	expression tag	UNP K0BRG7
B	609	HIS	-	expression tag	UNP K0BRG7
B	610	HIS	-	expression tag	UNP K0BRG7
B	611	HIS	-	expression tag	UNP K0BRG7
B	612	HIS	-	expression tag	UNP K0BRG7

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



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	45	45	45	0	0
3	B	18	18	18	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	46.98Å 108.47Å 125.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.81 – 2.51 49.81 – 2.51	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.81-2.51) 99.2 (49.81-2.51)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.31 (at 2.51Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.208 , 0.251 0.201 , 0.245	Depositor DCC
R_{free} test set	1154 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å ²)	52.0	Xtrriage
Anisotropy	0.651	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 39.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3307	wwPDB-VP
Average B, all atoms (Å ²)	64.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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	0.33	0/1647	0.54	1/2249 (0.0%)
1	B	0.30	0/1647	0.49	0/2249
All	All	0.32	0/3294	0.52	1/4498 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	460	SER	CB-CA-C	9.06	127.32	110.10

There are no chirality outliers.

There are no planarity outliers.

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	1608	0	1573	20	0
1	B	1608	0	1573	27	0
2	A	14	0	13	2	0
2	B	14	0	13	0	0
3	A	45	0	0	2	0
3	B	18	0	0	2	0
All	All	3307	0	3172	45	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.

All (45) 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:456:LEU:HG	1:A:479:LEU:HD21	1.60	0.81
1:B:506:LEU:HB3	1:B:553:TRP:HB2	1.64	0.79
1:B:533:THR:HB	3:B:816:HOH:O	1.90	0.71
1:B:388:LEU:HD21	1:B:571:PHE:HE1	1.58	0.68
1:B:579:THR:OG1	1:B:580:ASP:N	2.28	0.67
1:B:487:ASN:OD1	1:B:487:ASN:N	2.25	0.66
1:A:437:CYS:SG	3:A:840:HOH:O	2.53	0.66
1:B:540:TYR:HD2	1:B:542:ARG:HH22	1.44	0.65
1:A:413:LYS:HA	2:A:701:NAG:H61	1.86	0.57
1:A:394:PRO:HG3	1:A:400:LYS:HG3	1.88	0.55
1:B:406:ASN:HA	1:B:583:SER:HB3	1.90	0.54
1:B:396:VAL:N	3:B:801:HOH:O	2.32	0.51
1:B:518:VAL:HG22	1:B:524:SER:HA	1.93	0.51
1:A:388:LEU:HD11	1:A:571:PHE:CE1	2.46	0.50
1:A:401:ARG:HH12	1:A:521:ASN:HB3	1.77	0.50
1:B:431:ALA:O	1:B:435:SER:OG	2.30	0.50
1:B:395:GLN:HG3	1:B:498:SER:HB2	1.94	0.48
1:B:400:LYS:HB3	1:B:445:TYR:CE2	2.49	0.48
1:B:509:ASP:OD1	1:B:509:ASP:N	2.46	0.48
1:A:393:PRO:HG2	1:A:567:LEU:HD21	1.96	0.47
1:B:524:SER:HB3	1:B:527:VAL:HG13	1.96	0.47
1:B:514:VAL:O	1:B:516:GLN:NE2	2.42	0.46
1:B:540:TYR:HB3	1:B:542:ARG:HH12	1.81	0.46
1:A:533:THR:HG23	3:A:835:HOH:O	2.15	0.46
1:A:535:TRP:CE2	1:B:517:LEU:HB2	2.51	0.45
1:B:580:ASP:OD1	1:B:580:ASP:N	2.49	0.45
1:A:542:ARG:O	1:B:458:VAL:HG22	2.17	0.44
1:A:484:VAL:HA	1:A:485:PRO:HD3	1.85	0.43
1:A:456:LEU:HD13	1:A:464:ILE:HG21	2.00	0.43
1:B:422:ASP:HB3	1:B:481:LEU:HB2	2.00	0.43
1:A:433:ILE:HA	1:A:438:TYR:OH	2.19	0.42
1:A:449:PRO:HG2	1:A:452:MET:HG3	2.01	0.42
1:B:498:SER:HB3	1:B:534:VAL:HG23	2.02	0.42
1:B:413:LYS:O	1:B:416:SER:OG	2.37	0.42
1:B:388:LEU:HD21	1:B:571:PHE:CE1	2.47	0.41
1:A:587:LYS:HB2	1:A:587:LYS:HE3	1.71	0.41
1:B:458:VAL:HG13	1:B:461:ALA:HB3	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:542:ARG:HA	1:B:554:LEU:O	2.21	0.41
1:A:408:ASN:HA	1:A:585:CYS:O	2.21	0.41
1:A:471:GLN:HB3	1:A:477:THR:HG21	2.02	0.40
1:B:392:THR:HA	1:B:393:PRO:HD3	1.90	0.40
1:A:450:LEU:HA	1:A:450:LEU:HD12	1.92	0.40
1:A:587:LYS:HE2	2:A:701:NAG:H81	2.03	0.40
1:A:392:THR:HA	1:A:393:PRO:HD3	1.90	0.40
1:B:482:ALA:HB3	1:B:571:PHE:HE2	1.86	0.40

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	206/251 (82%)	192 (93%)	12 (6%)	2 (1%)	15	27
1	B	206/251 (82%)	194 (94%)	11 (5%)	1 (0%)	29	47
All	All	412/502 (82%)	386 (94%)	23 (6%)	3 (1%)	22	37

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	462	GLY
1	B	382	GLU
1	A	428	ILE

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	190/224 (85%)	179 (94%)	11 (6%)	20	36
1	B	190/224 (85%)	167 (88%)	23 (12%)	5	8
All	All	380/448 (85%)	346 (91%)	34 (9%)	9	18

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	406	ASN
1	A	411	LEU
1	A	412	THR
1	A	423	PHE
1	A	455	ASP
1	A	456	LEU
1	A	484	VAL
1	A	488	LEU
1	A	490	THR
1	A	495	LEU
1	A	581	THR
1	B	400	LYS
1	B	423	PHE
1	B	435	SER
1	B	456	LEU
1	B	458	VAL
1	B	473	PHE
1	B	481	LEU
1	B	484	VAL
1	B	487	ASN
1	B	488	LEU
1	B	490	THR
1	B	492	THR
1	B	509	ASP
1	B	511	ARG
1	B	512	THR
1	B	514	VAL
1	B	533	THR
1	B	540	TYR
1	B	542	ARG
1	B	548	LEU
1	B	549	GLU
1	B	580	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	581	THR

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

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	$\# Z > 2$
2	NAG	A	701	1	14,14,15	0.49	0	17,19,21	0.90	1 (5%)
2	NAG	B	701	1	14,14,15	0.43	0	17,19,21	1.23	2 (11%)

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	A	701	1	-	1/6/23/26	0/1/1/1
2	NAG	B	701	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	701	NAG	C1-O5-C5	3.42	116.83	112.19
2	B	701	NAG	C2-N2-C7	-2.54	119.28	122.90
2	A	701	NAG	C2-N2-C7	-2.18	119.79	122.90

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	701	NAG	C8-C7-N2-C2
2	B	701	NAG	O5-C5-C6-O6
2	A	701	NAG	O5-C5-C6-O6
2	B	701	NAG	O7-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	NAG	2	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

6.1 Protein, DNA and RNA chains

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	208/251 (82%)	-0.05	1 (0%) 91 92	31, 48, 91, 122	0
1	B	208/251 (82%)	0.27	19 (9%) 9 9	40, 66, 149, 179	0
All	All	416/502 (82%)	0.11	20 (4%) 30 33	31, 58, 121, 179	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	547	PRO	4.7
1	B	506	LEU	4.4
1	A	460	SER	4.0
1	B	511	ARG	3.9
1	B	512	THR	3.8
1	B	588	LEU	3.8
1	B	553	TRP	3.5
1	B	507	LEU	3.3
1	B	579	THR	3.2
1	B	551	GLY	2.9
1	B	508	SER	2.8
1	B	554	LEU	2.6
1	B	548	LEU	2.5
1	B	544	GLN	2.4
1	B	587	LYS	2.4
1	B	540	TYR	2.3
1	B	510	ASP	2.2
1	B	552	GLY	2.1
1	B	473	PHE	2.1
1	B	545	LEU	2.0

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
2	NAG	B	701	14/15	0.81	0.29	85,93,96,97	0
2	NAG	A	701	14/15	0.93	0.17	79,83,87,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.