



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 13, 2025 – 02:47 PM JST

PDB ID : 8XYG
Title : Crystal structure of SARS-CoV-2 BQ.1.1 RBD and human ACE2
Authors : Lan, J.; Wang, C.H.
Deposited on : 2024-01-19
Resolution : 3.64 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

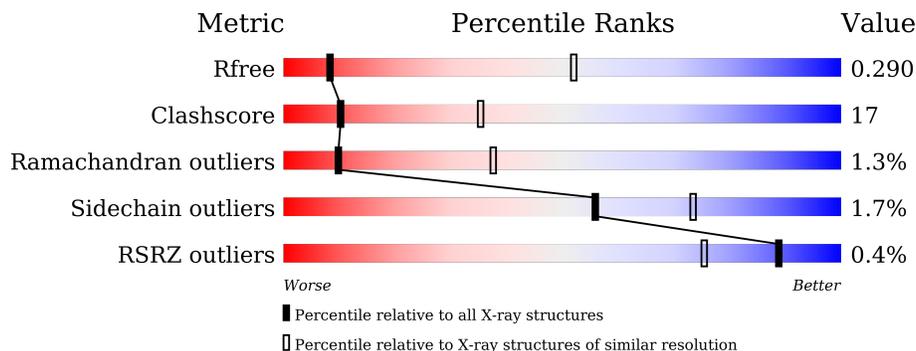
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.64 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}	164625	1000 (3.76-3.52)
Clashscore	180529	1046 (3.76-3.52)
Ramachandran outliers	177936	1031 (3.76-3.52)
Sidechain outliers	177891	1029 (3.76-3.52)
RSRZ outliers	164620	1682 (3.78-3.50)

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 of 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	596	
1	C	596	
2	B	192	
2	D	192	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 12793 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 Processed angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	596	4871	3118	807	917	29	0	1	0
1	C	595	4864	3113	806	916	29	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	PRO	-	expression tag	UNP Q9BYF1
C	18	PRO	-	expression tag	UNP Q9BYF1

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	192	1533	989	256	280	8	0	0	0
2	D	191	1525	985	254	278	8	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	339	ASP	GLY	variant	UNP P0DTC2
B	346	THR	ARG	variant	UNP P0DTC2
B	371	PHE	SER	variant	UNP P0DTC2
B	373	PRO	SER	variant	UNP P0DTC2
B	375	PHE	SER	variant	UNP P0DTC2
B	405	ASN	ASP	variant	UNP P0DTC2
B	408	SER	ARG	variant	UNP P0DTC2
B	417	ASN	LYS	variant	UNP P0DTC2
B	440	LYS	ASN	variant	UNP P0DTC2
B	444	THR	LYS	variant	UNP P0DTC2

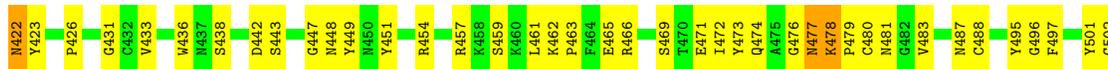
Continued on next page...

Continued from previous page...

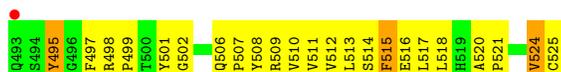
Chain	Residue	Modelled	Actual	Comment	Reference
B	452	ARG	LEU	variant	UNP P0DTC2
B	460	LYS	ASN	variant	UNP P0DTC2
B	477	ASN	SER	variant	UNP P0DTC2
B	478	LYS	THR	variant	UNP P0DTC2
B	484	ALA	GLU	variant	UNP P0DTC2
B	486	VAL	PHE	variant	UNP P0DTC2
B	498	ARG	GLN	variant	UNP P0DTC2
B	501	TYR	ASN	variant	UNP P0DTC2
B	505	HIS	TYR	variant	UNP P0DTC2
D	339	ASP	GLY	variant	UNP P0DTC2
D	346	THR	ARG	variant	UNP P0DTC2
D	371	PHE	SER	variant	UNP P0DTC2
D	373	PRO	SER	variant	UNP P0DTC2
D	375	PHE	SER	variant	UNP P0DTC2
D	405	ASN	ASP	variant	UNP P0DTC2
D	408	SER	ARG	variant	UNP P0DTC2
D	417	ASN	LYS	variant	UNP P0DTC2
D	440	LYS	ASN	variant	UNP P0DTC2
D	444	THR	LYS	variant	UNP P0DTC2
D	452	ARG	LEU	variant	UNP P0DTC2
D	460	LYS	ASN	variant	UNP P0DTC2
D	477	ASN	SER	variant	UNP P0DTC2
D	478	LYS	THR	variant	UNP P0DTC2
D	484	ALA	GLU	variant	UNP P0DTC2
D	486	VAL	PHE	variant	UNP P0DTC2
D	498	ARG	GLN	variant	UNP P0DTC2
D	501	TYR	ASN	variant	UNP P0DTC2
D	505	HIS	TYR	variant	UNP P0DTC2



● Molecule 2: Spike protein S1



● Molecule 2: Spike protein S1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	73.39Å 152.13Å 117.44Å 90.00° 102.87° 90.00°	Depositor
Resolution (Å)	33.92 – 3.64 33.92 – 3.64	Depositor EDS
% Data completeness (in resolution range)	98.3 (33.92-3.64) 98.2 (33.92-3.64)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.52 (at 3.65Å)	Xtrriage
Refinement program	PHENIX (1.19.1_4122: ???)	Depositor
R, R_{free}	0.245 , 0.291 0.247 , 0.290	Depositor DCC
R_{free} test set	1385 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å ²)	117.0	Xtrriage
Anisotropy	0.511	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 90.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12793	wwPDB-VP
Average B, all atoms (Å ²)	150.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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](#)

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	0.29	0/5013	0.49	0/6811
1	C	0.28	0/5005	0.47	0/6800
2	B	0.39	1/1579 (0.1%)	0.67	1/2151 (0.0%)
2	D	0.41	1/1571 (0.1%)	0.67	0/2140
All	All	0.32	2/13168 (0.0%)	0.53	1/17902 (0.0%)

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	D	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	336	CYS	CB-SG	-6.00	1.72	1.82
2	D	347	PHE	CE1-CZ	5.55	1.47	1.37

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	390	LEU	CA-CB-CG	5.57	128.12	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	392	PHE	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	4871	0	4649	135	0
1	C	4864	0	4641	101	0
2	B	1533	0	1457	95	0
2	D	1525	0	1451	96	0
All	All	12793	0	12198	422	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 17.

The worst 5 of 422 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)
2:B:522:ALA:HB1	2:B:525:CYS:HB2	1.53	0.87
2:D:341:VAL:O	2:D:509:ARG:NH2	2.08	0.85
1:A:456:LEU:HD22	1:A:512:PHE:HD2	1.39	0.84
2:D:431:GLY:HA3	2:D:514:SER:HA	1.59	0.84
2:D:409:GLN:HA	2:D:414:GLN:HG2	1.59	0.82

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	595/596 (100%)	556 (93%)	35 (6%)	4 (1%)	19 50
1	C	594/596 (100%)	563 (95%)	29 (5%)	2 (0%)	37 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	190/192 (99%)	147 (77%)	37 (20%)	6 (3%)	3	22
2	D	189/192 (98%)	139 (74%)	42 (22%)	8 (4%)	2	17
All	All	1568/1576 (100%)	1405 (90%)	143 (9%)	20 (1%)	10	37

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	362	VAL
2	B	477	ASN
2	D	417	ASN
2	B	372	ALA
2	B	524	VAL

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	528/527 (100%)	525 (99%)	3 (1%)	84	91
1	C	527/527 (100%)	521 (99%)	6 (1%)	70	81
2	B	166/166 (100%)	158 (95%)	8 (5%)	21	47
2	D	165/166 (99%)	159 (96%)	6 (4%)	30	55
All	All	1386/1386 (100%)	1363 (98%)	23 (2%)	56	73

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	504	PHE
2	D	338	PHE
1	C	609	ASP
2	D	355	ARG
2	B	422	ASN

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

Mol	Chain	Res	Type
1	A	401	HIS
1	A	535	HIS
1	C	552	GLN
1	C	556	ASN

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 oligosaccharides 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.

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	596/596 (100%)	-0.27	0 100 100	75, 138, 190, 266	1 (0%)
1	C	595/596 (99%)	-0.26	1 (0%) 92 84	73, 139, 203, 278	1 (0%)
2	B	192/192 (100%)	-0.07	2 (1%) 79 59	109, 151, 246, 307	0
2	D	191/192 (99%)	0.07	4 (2%) 63 42	115, 169, 269, 305	0
All	All	1574/1576 (99%)	-0.20	7 (0%) 89 75	73, 143, 222, 307	2 (0%)

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	410	ILE	3.1
2	B	338	PHE	2.8
2	D	390	LEU	2.2
2	D	493	GLN	2.2
1	C	35	GLU	2.1

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.