Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2022 – 04:13 pm BST

PDB ID	:	7ZF8
Title	:	SARS-CoV-2 Omicron BA.2 RBD in complex with COVOX-150 Fab
Authors	:	Zhou, D.; Huo, J.; Ren, J.; Stuart, D.I.
Deposited on	:	2022-04-01
Resolution	:	2.95 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
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 2.95 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.
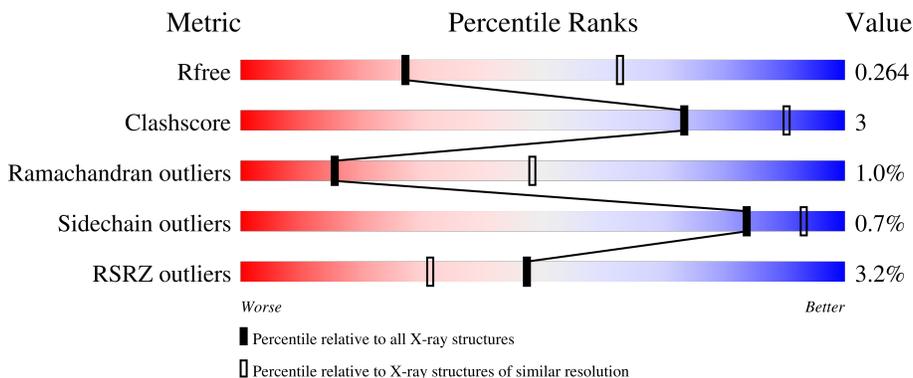
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 <=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	H	222	% 89% 9% .
2	L	216	% 93% 6%
3	E	202	7% 83% 13% ..

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4798 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 COVOX-150 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	216	Total 1597	C 1005	N 266	O 317	S 9	0	0	0

- Molecule 2 is a protein called COVOX-150 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	215	Total 1646	C 1034	N 272	O 335	S 5	0	1	0

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	195	Total 1555	C 1004	N 262	O 281	S 8	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	327	HIS	-	expression tag	UNP P0DTC2
E	328	HIS	-	expression tag	UNP P0DTC2
E	329	HIS	-	expression tag	UNP P0DTC2
E	330	HIS	-	expression tag	UNP P0DTC2
E	331	HIS	-	expression tag	UNP P0DTC2
E	332	HIS	-	expression tag	UNP P0DTC2
E	339	ASP	GLY	variant	UNP P0DTC2
E	371	PHE	SER	variant	UNP P0DTC2
E	373	PRO	SER	variant	UNP P0DTC2
E	375	PHE	SER	variant	UNP P0DTC2
E	376	ALA	THR	variant	UNP P0DTC2
E	405	ASN	ASP	variant	UNP P0DTC2
E	408	SER	ARG	variant	UNP P0DTC2
E	417	ASN	LYS	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	440	LYS	ASN	variant	UNP P0DTC2
E	477	ASN	SER	variant	UNP P0DTC2
E	478	LYS	THR	variant	UNP P0DTC2
E	484	ALA	GLU	variant	UNP P0DTC2
E	493	ARG	GLN	variant	UNP P0DTC2
E	498	ARG	GLN	variant	UNP P0DTC2
E	501	TYR	ASN	variant	UNP P0DTC2
E	505	HIS	TYR	variant	UNP P0DTC2
E	527	LYS	-	expression tag	UNP P0DTC2
E	528	LYS	-	expression tag	UNP P0DTC2

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: COVOX-150 heavy chain
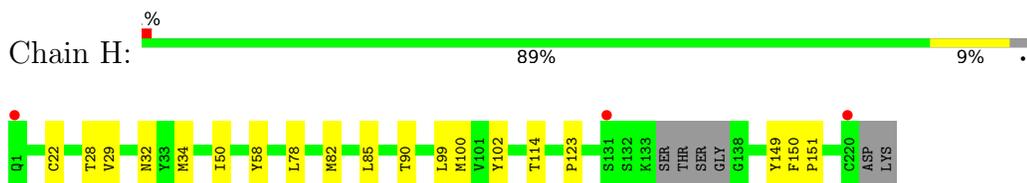
Chain H:

- Molecule 2: COVOX-150 light chain

Chain L:
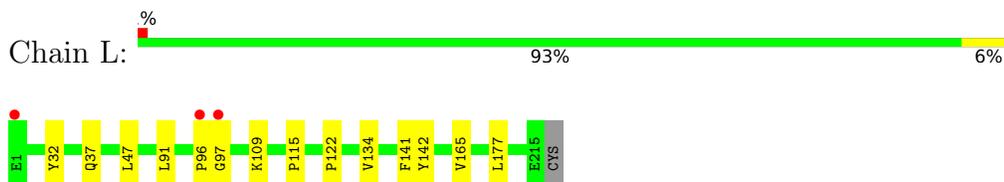
- Molecule 3: Spike protein S1

Chain E:
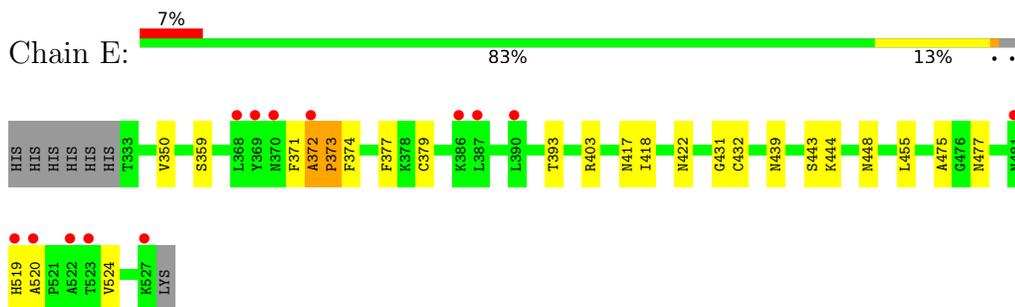
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	194.22Å 84.93Å 58.42Å 90.00° 101.02° 90.00°	Depositor
Resolution (Å)	57.34 – 2.95 95.32 – 2.95	Depositor EDS
% Data completeness (in resolution range)	97.9 (57.34-2.95) 97.9 (95.32-2.95)	Depositor EDS
R_{merge}	0.49	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.00 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.220 , 0.258 0.227 , 0.264	Depositor DCC
R_{free} test set	977 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	52.5	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.40, < L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4798	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	H	0.25	0/1631	0.49	0/2223
2	L	0.25	0/1685	0.49	0/2290
3	E	0.25	0/1602	0.48	0/2180
All	All	0.25	0/4918	0.49	0/6693

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 (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	H	1597	0	1573	12	0
2	L	1646	0	1607	7	0
3	E	1555	0	1475	15	0
All	All	4798	0	4655	31	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 3.

All (31) 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:H:29:VAL:HG13	1:H:34:MET:HG3	1.74	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:372:ALA:HB3	3:E:373:PRO:HD3	1.75	0.67
2:L:122:PRO:HD3	2:L:134:VAL:HG22	1.81	0.60
3:E:518:LEU:O	3:E:520:ALA:N	2.36	0.58
1:H:22:CYS:HB3	1:H:78:LEU:HB3	1.85	0.58
3:E:439:ASN:O	3:E:443:SER:OG	2.20	0.58
2:L:32:TYR:OH	3:E:403:ARG:NH2	2.37	0.57
1:H:90:THR:HG23	1:H:114:THR:HA	1.89	0.54
2:L:37:GLN:HB2	2:L:47:LEU:HD11	1.89	0.54
3:E:444:LYS:HG3	3:E:448:ASN:HB2	1.90	0.52
1:H:123:PRO:HB3	1:H:149:TYR:HB3	1.93	0.51
1:H:82:MET:HB3	1:H:85:LEU:HD21	1.93	0.50
1:H:100:MET:HG3	3:E:455:LEU:HD11	1.94	0.50
2:L:165[A]:VAL:HG22	2:L:177:LEU:HD12	1.94	0.49
3:E:417:ASN:OD1	3:E:418:ILE:N	2.45	0.49
1:H:99:LEU:HB3	1:H:102:TYR:HB2	1.94	0.48
2:L:165[B]:VAL:HG22	2:L:177:LEU:HD12	1.96	0.48
3:E:393:THR:HG23	3:E:517:LEU:HA	1.95	0.47
2:L:115:PRO:HB3	2:L:141:PHE:HB3	1.96	0.47
1:H:50:ILE:HG22	1:H:58:TYR:HB2	1.98	0.46
3:E:350:VAL:HG22	3:E:422:ASN:HB3	1.98	0.45
1:H:34:MET:HB3	1:H:78:LEU:HD22	2.00	0.44
3:E:431:GLY:HA2	3:E:515:PHE:HD2	1.83	0.44
1:H:32:ASN:OD1	3:E:475:ALA:HB1	2.18	0.43
3:E:393:THR:HG21	3:E:520:ALA:HB3	2.00	0.43
2:L:109:LYS:HA	2:L:142:TYR:OH	2.19	0.42
3:E:431:GLY:HA2	3:E:515:PHE:CD2	2.54	0.41
3:E:359:SER:HA	3:E:524:VAL:HG22	2.03	0.41
1:H:28:THR:O	1:H:32:ASN:ND2	2.49	0.41
1:H:150:PHE:HA	1:H:151:PRO:HA	1.88	0.41
3:E:379:CYS:HA	3:E:432:CYS:HA	2.02	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	H	212/222 (96%)	206 (97%)	6 (3%)	0	100	100
2	L	214/216 (99%)	203 (95%)	9 (4%)	2 (1%)	17	51
3	E	193/202 (96%)	180 (93%)	9 (5%)	4 (2%)	7	29
All	All	619/640 (97%)	589 (95%)	24 (4%)	6 (1%)	15	48

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	L	96	PRO
3	E	519	HIS
2	L	97	GLY
3	E	374	PHE
3	E	372	ALA
3	E	373	PRO

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 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	H	181/187 (97%)	181 (100%)	0	100	100
2	L	188/188 (100%)	187 (100%)	1 (0%)	88	95
3	E	166/174 (95%)	163 (98%)	3 (2%)	59	82
All	All	535/549 (97%)	531 (99%)	4 (1%)	84	93

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	91	LEU
3	E	371	PHE
3	E	377	PHE
3	E	477	ASN

Sometimes 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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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.

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(Å²)	Q<0.9
1	H	216/222 (97%)	0.15	3 (1%) 75	59	30, 43, 90, 123	0
2	L	215/216 (99%)	0.13	3 (1%) 75	59	29, 47, 75, 107	0
3	E	195/202 (96%)	0.46	14 (7%) 15	8	31, 57, 143, 206	0
All	All	626/640 (97%)	0.24	20 (3%) 47	31	29, 48, 118, 206	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	370	ASN	5.9
1	H	131	SER	4.6
3	E	519	HIS	4.5
3	E	518	LEU	4.3
3	E	369	TYR	4.0
3	E	522	ALA	3.8
3	E	520	ALA	3.5
1	H	220	CYS	3.5
3	E	527	LYS	3.0
3	E	387	LEU	2.9
2	L	96	PRO	2.9
3	E	368	LEU	2.8
2	L	97	GLY	2.5
3	E	390	LEU	2.4
3	E	386	LYS	2.3
1	H	1	GLN	2.3
3	E	372	ALA	2.3
2	L	1	GLU	2.2
3	E	523	THR	2.2
3	E	481	ASN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.