



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 23, 2023 – 02:53 AM JST

PDB ID : 7YTN
Title : Crystal structure of SARS-CoV-2 Alpha RBD in complex with the D27LEY neutralizing antibody Fab fragment
Authors : Jeong, B.S.; Jeon, J.Y.; Oh, B.H.
Deposited on : 2022-08-15
Resolution : 3.51 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) : 1.13
EDS : 2.36
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) : 2.36

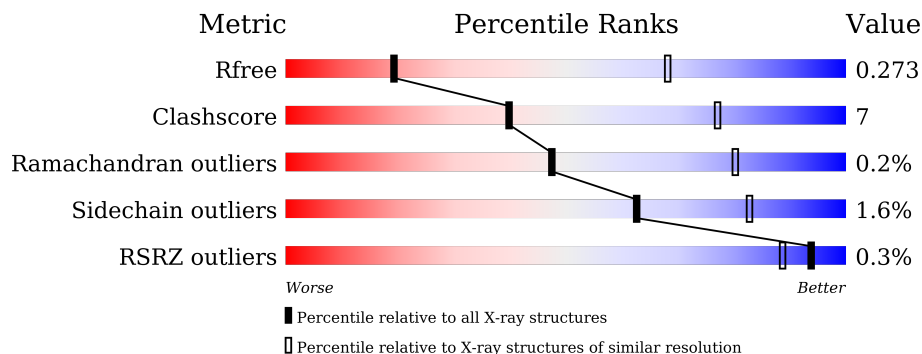
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.51 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	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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	H	239	 71% 17% 11%
2	L	239	 71% 17% 12%
3	C	262	 58% 11% 31%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4638 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 Heavy chain of Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	212	1568	985	263	313	7	0	0	0

- Molecule 2 is a protein called Light chain of Fab.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	211	1620	1009	277	330	4	0	0	0

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	181	1450	932	239	272	7	1	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	280	MET	-	initiating methionine	UNP P0DTC2
C	281	VAL	-	expression tag	UNP P0DTC2
C	282	LEU	-	expression tag	UNP P0DTC2
C	283	VAL	-	expression tag	UNP P0DTC2
C	284	ASN	-	expression tag	UNP P0DTC2
C	285	GLN	-	expression tag	UNP P0DTC2
C	286	SER	-	expression tag	UNP P0DTC2
C	287	HIS	-	expression tag	UNP P0DTC2
C	288	GLN	-	expression tag	UNP P0DTC2
C	289	GLY	-	expression tag	UNP P0DTC2
C	290	PHE	-	expression tag	UNP P0DTC2
C	291	ASN	-	expression tag	UNP P0DTC2
C	292	LYS	-	expression tag	UNP P0DTC2
C	293	GLU	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	294	HIS	-	expression tag	UNP P0DTC2
C	295	THR	-	expression tag	UNP P0DTC2
C	296	SER	-	expression tag	UNP P0DTC2
C	297	LYS	-	expression tag	UNP P0DTC2
C	298	MET	-	expression tag	UNP P0DTC2
C	299	VAL	-	expression tag	UNP P0DTC2
C	300	SER	-	expression tag	UNP P0DTC2
C	301	ALA	-	expression tag	UNP P0DTC2
C	302	ILE	-	expression tag	UNP P0DTC2
C	303	VAL	-	expression tag	UNP P0DTC2
C	304	LEU	-	expression tag	UNP P0DTC2
C	305	TYR	-	expression tag	UNP P0DTC2
C	306	VAL	-	expression tag	UNP P0DTC2
C	307	LEU	-	expression tag	UNP P0DTC2
C	308	LEU	-	expression tag	UNP P0DTC2
C	309	ALA	-	expression tag	UNP P0DTC2
C	310	ALA	-	expression tag	UNP P0DTC2
C	311	ALA	-	expression tag	UNP P0DTC2
C	312	ALA	-	expression tag	UNP P0DTC2
C	313	HIS	-	expression tag	UNP P0DTC2
C	314	SER	-	expression tag	UNP P0DTC2
C	315	ALA	-	expression tag	UNP P0DTC2
C	316	PHE	-	expression tag	UNP P0DTC2
C	317	ALA	-	expression tag	UNP P0DTC2
C	318	ALA	-	expression tag	UNP P0DTC2
C	501	TYR	ASN	variant	UNP P0DTC2

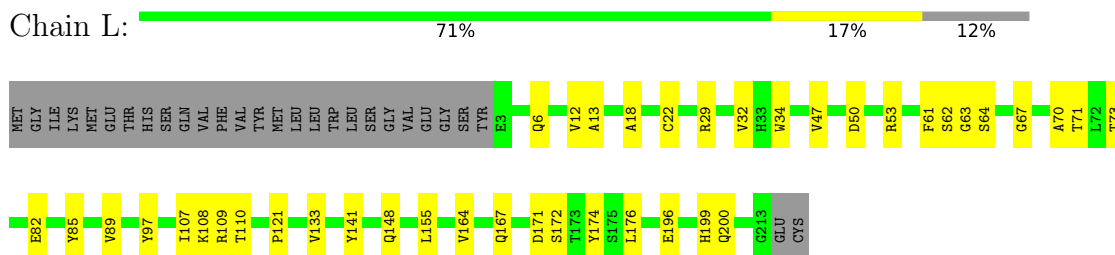
3 Residue-property plots i

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.

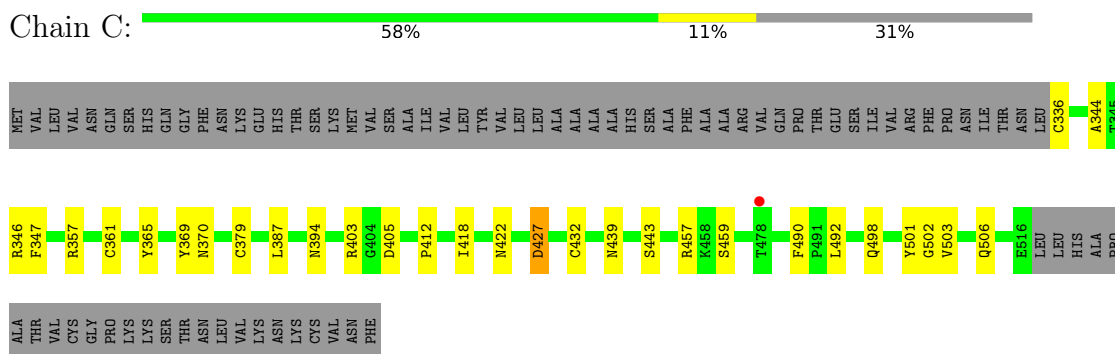
- Molecule 1: Heavy chain of Fab



- Molecule 2: Light chain of Fab



- Molecule 3: Spike protein S1



4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	155.96Å 155.96Å 166.68Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.99 – 3.51 43.46 – 3.51	Depositor EDS
% Data completeness (in resolution range)	91.1 (38.99-3.51) 89.3 (43.46-3.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 3.48Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.237 , 0.274 0.236 , 0.273	Depositor DCC
R_{free} test set	1540 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å ²)	35.2	Xtrriage
Anisotropy	0.005	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 31.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	4638	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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	H	0.25	0/1605	0.45	0/2188
2	L	0.25	0/1656	0.44	0/2253
3	C	0.25	0/1492	0.41	0/2028
All	All	0.25	0/4753	0.43	0/6469

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	1568	0	1522	28	0
2	L	1620	0	1558	23	0
3	C	1450	0	1366	18	0
All	All	4638	0	4446	66	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.

The worst 5 of 66 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)
3:C:336:CYS:N	3:C:361:CYS:HG	1.72	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:46:GLU:OE2	1:H:63:LYS:NZ	2.17	0.75
2:L:29:ARG:HH12	2:L:67:GLY:HA2	1.55	0.72
1:H:6:GLN:NE2	1:H:94:TYR:O	2.25	0.70
2:L:62:SER:HB2	2:L:73:THR:HB	1.74	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 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	208/239 (87%)	194 (93%)	13 (6%)	1 (0%)	29	67
2	L	209/239 (87%)	193 (92%)	16 (8%)	0	100	100
3	C	179/262 (68%)	155 (87%)	24 (13%)	0	100	100
All	All	596/740 (80%)	542 (91%)	53 (9%)	1 (0%)	47	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	27	GLY

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	H	173/199 (87%)	169 (98%)	4 (2%)	50	77
2	L	183/208 (88%)	183 (100%)	0	100	100
3	C	157/226 (70%)	153 (98%)	4 (2%)	47	75
All	All	513/633 (81%)	505 (98%)	8 (2%)	62	83

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	427	ASP
3	C	370	ASN
3	C	346	ARG
1	H	203	CYS
3	C	369	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	6	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](#)

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

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	H	212/239 (88%)	0.00	1 (0%) 91 84	11, 50, 80, 93	0
2	L	211/239 (88%)	-0.36	0 100 100	5, 26, 67, 81	0
3	C	181/262 (69%)	-0.07	1 (0%) 89 81	11, 36, 87, 107	1 (0%)
All	All	604/740 (81%)	-0.15	2 (0%) 94 89	5, 37, 80, 107	1 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	211	ASN	3.4
3	C	478	THR	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.