



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 18, 2021 – 11:40 AM JST

PDB ID : 7DEU  
Title : Crystal structure of SARS-CoV-2 RBD in complex with a neutralizing anti-body scFv  
Authors : Zhang, Z.; Zhang, G.; Li, X.; Rao, Z.; Guo, Y.  
Deposited on : 2020-11-05  
Resolution : 2.10 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](mailto: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 : 2.18  
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) : 2.18

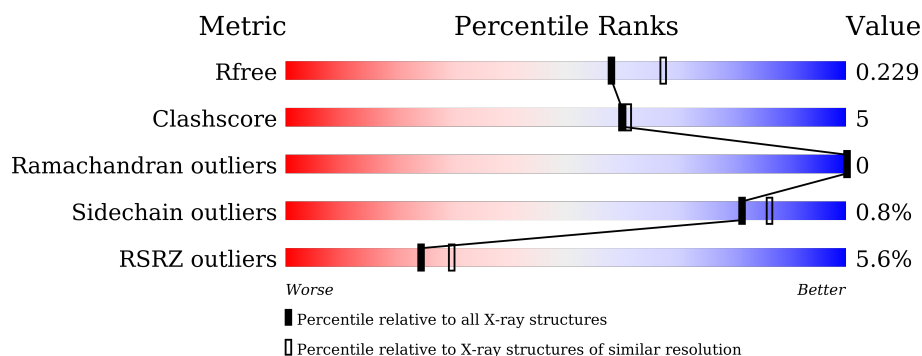
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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  $\geq 3$ , 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	206	<div> <div>11%</div> <div> <div></div> <div>80%</div> <div>15%</div> <div>6%</div> </div> </div>
2	B	253	<div> <div></div> <div>83%</div> <div>8%</div> <div>9%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3560 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	0	0
			1536	985	256	287	8			

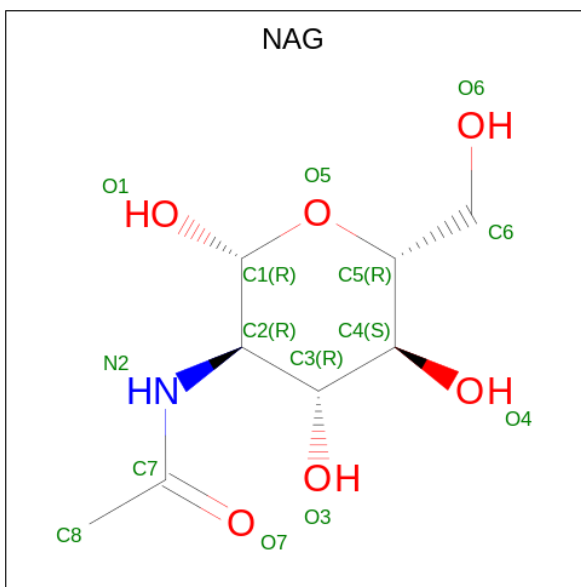
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	531	ALA	-	expression tag	UNP P0DTC2
A	532	ALA	-	expression tag	UNP P0DTC2
A	533	ALA	-	expression tag	UNP P0DTC2
A	534	HIS	-	expression tag	UNP P0DTC2
A	535	HIS	-	expression tag	UNP P0DTC2
A	536	HIS	-	expression tag	UNP P0DTC2
A	537	HIS	-	expression tag	UNP P0DTC2
A	538	HIS	-	expression tag	UNP P0DTC2
A	539	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called antibody scFv.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	231	Total	C	N	O	S	0	0	0
			1781	1118	297	358	8			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



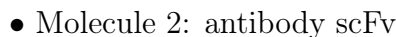
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	100	Total	O	0	0
			100	100		
4	B	129	Total	O	0	0
			129	129		



- Molecule 1: Spike protein S1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.60Å 84.60Å 158.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.70 – 2.10 47.70 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.5 (47.70-2.10) 99.5 (47.70-2.10)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.06 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.14_3228	Depositor
R, $R_{free}$	0.209 , 0.229 0.209 , 0.229	Depositor DCC
$R_{free}$ test set	1990 reflections (5.83%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtriage
Anisotropy	0.388	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 46.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3560	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

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	0.41	0/1580	0.57	0/2151
2	B	0.46	0/1824	0.64	0/2473
All	All	0.44	0/3404	0.61	0/4624

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

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	1536	0	1452	17	0
2	B	1781	0	1694	13	0
3	A	14	0	13	0	0
4	A	100	0	0	4	0
4	B	129	0	0	1	0
All	All	3560	0	3159	30	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 5.

All (30) 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:73:HIS:CE1	2:B:75:SER:HB2	2.26	0.69
1:A:450:ASN:HA	4:A:701:HOH:O	1.97	0.64
1:A:445:VAL:HG23	1:A:446:GLY:H	1.64	0.61
1:A:499:PRO:O	1:A:500:THR:OG1	2.17	0.60
1:A:336:CYS:HB2	1:A:363:ALA:HB2	1.85	0.58
1:A:365:TYR:CD2	1:A:387:LEU:HB3	2.39	0.57
1:A:362:VAL:HG13	1:A:526:GLY:HA2	1.89	0.55
1:A:443:SER:OG	1:A:499:PRO:HD3	2.08	0.54
2:B:40:SER:HB2	2:B:43:LYS:HB2	1.90	0.53
1:A:349:SER:HA	4:A:735:HOH:O	2.07	0.52
2:B:169:ALA:HB3	2:B:224:GLN:HG2	1.93	0.51
2:B:12:VAL:O	2:B:118:VAL:HA	2.11	0.51
1:A:403:ARG:HB2	1:A:406:GLU:OE1	2.12	0.49
2:B:12:VAL:HG12	2:B:13:LYS:H	1.78	0.49
1:A:444:LYS:O	1:A:499:PRO:HD2	2.15	0.47
2:B:73:HIS:HE1	2:B:75:SER:HB2	1.74	0.46
2:B:12:VAL:CG1	2:B:16:ALA:HB3	2.45	0.45
2:B:12:VAL:HG21	2:B:18:VAL:CG2	2.46	0.45
1:A:364:ASP:OD2	1:A:367:VAL:N	2.49	0.45
1:A:503:VAL:HA	1:A:506:GLN:CD	2.38	0.44
2:B:11:LEU:HD23	2:B:117:THR:HB	1.99	0.44
2:B:181:PRO:HB2	2:B:190:TYR:CD1	2.53	0.44
1:A:354:ASN:O	1:A:398:ASP:HA	2.18	0.43
2:B:33:THR:OG1	4:B:301:HOH:O	2.22	0.42
1:A:449:TYR:O	4:A:701:HOH:O	2.21	0.42
2:B:110:TRP:CE3	2:B:179:PRO:HD2	2.55	0.41
1:A:347:PHE:CE2	1:A:399:SER:HB2	2.56	0.41
1:A:359:SER:HA	1:A:524:VAL:HG22	2.02	0.40
2:B:183:ILE:HA	2:B:188:SER:O	2.21	0.40
1:A:501:ASN:ND2	4:A:713:HOH:O	2.55	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	192/206 (93%)	182 (95%)	10 (5%)	0	100	100
2	B	227/253 (90%)	219 (96%)	8 (4%)	0	100	100
All	All	419/459 (91%)	401 (96%)	18 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	A	167/176 (95%)	165 (99%)	2 (1%)	71	77
2	B	195/207 (94%)	194 (100%)	1 (0%)	88	92
All	All	362/383 (94%)	359 (99%)	3 (1%)	81	86

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	346	ARG
1	A	386	LYS
2	B	224	GLN

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

Mol	Chain	Res	Type
2	B	5	GLN

### 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

1 ligand is 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	$\# Z  > 2$
3	NAG	A	601	1	14,14,15	0.38	0	17,19,21	0.52	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
3	NAG	A	601	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	601	NAG	C8-C7-N2-C2
3	A	601	NAG	O7-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	194/206 (94%)	0.51	23 (11%) <span style="border: 1px solid red; padding: 0 2px;">4</span> <span style="border: 1px solid red; padding: 0 2px;">5</span>	28, 46, 83, 94	0
2	B	231/253 (91%)	0.10	1 (0%) <span style="border: 1px solid blue; padding: 0 2px;">92</span> <span style="border: 1px solid blue; padding: 0 2px;">93</span>	24, 39, 61, 82	0
All	All	425/459 (92%)	0.29	24 (5%) <span style="border: 1px solid red; padding: 0 2px;">24</span> <span style="border: 1px solid red; padding: 0 2px;">29</span>	24, 41, 79, 94	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	373	SER	7.8
1	A	521	PRO	4.5
1	A	335	LEU	4.1
1	A	520	ALA	3.9
1	A	394	ASN	3.8
1	A	334	ASN	3.6
1	A	362	VAL	3.6
2	B	11	LEU	3.4
1	A	445	VAL	3.4
1	A	518	LEU	3.3
1	A	369	TYR	3.2
1	A	519	HIS	3.0
1	A	372	ALA	3.0
1	A	449	TYR	2.8
1	A	341	VAL	2.7
1	A	517	LEU	2.6
1	A	360	ASN	2.5
1	A	500	THR	2.3
1	A	522	ALA	2.3
1	A	345	THR	2.2
1	A	336	CYS	2.2
1	A	368	LEU	2.1
1	A	346	ARG	2.1
1	A	446	GLY	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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	A	601	14/15	0.77	0.36	79,90,105,106	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.