



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 30, 2023 – 09:11 am GMT

PDB ID : 8BG4  
Title : Crystal structure of the SARS-CoV-2 S RBD in complex with pT1611 scFV  
Authors : Hansen, G.; Ssebyatika, G.L.; Krey, T.  
Deposited on : 2022-10-27  
Resolution : 1.60 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 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  
Mogul : 1.8.4, CSD as541be (2020)  
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

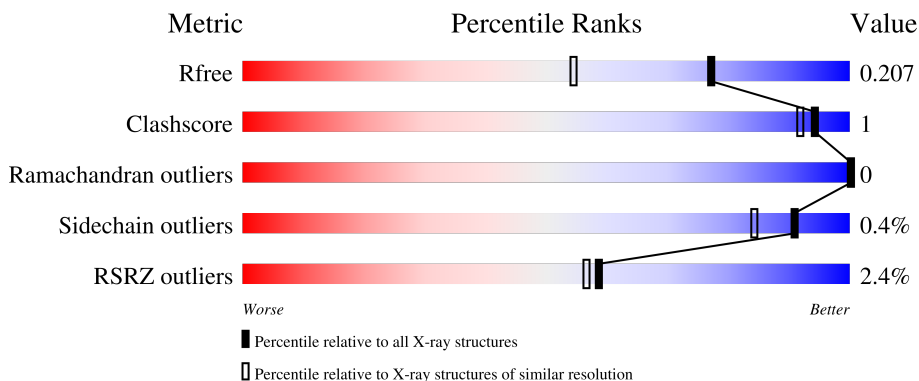
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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	249	 3% 84% 5% 11%
1	C	249	 2% 86% 11%
2	B	199	 2% 95% 11%
2	D	199	 3% 93% 11%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7239 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 pT1611 single-chain Fv.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	222	Total	C	N	O	S	0	0	0
			1688	1062	288	331	7			
1	C	222	Total	C	N	O	S	0	0	0
			1687	1061	287	331	8			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	195	Total	C	N	O	S	0	0	0
			1544	989	257	290	8			
2	D	194	Total	C	N	O	S	0	0	0
			1536	985	256	287	8			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	528	ASP	-	expression tag	UNP P0DTC2
B	529	ASP	-	expression tag	UNP P0DTC2
B	530	ASP	-	expression tag	UNP P0DTC2
B	531	ASP	-	expression tag	UNP P0DTC2
B	532	LYS	-	expression tag	UNP P0DTC2
D	528	ASP	-	expression tag	UNP P0DTC2
D	529	ASP	-	expression tag	UNP P0DTC2
D	530	ASP	-	expression tag	UNP P0DTC2
D	531	ASP	-	expression tag	UNP P0DTC2
D	532	LYS	-	expression tag	UNP P0DTC2

- 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
			Total	C	N	O		
3	B	1	14	8	1	5	0	0
3	D	1	14	8	1	5	0	0

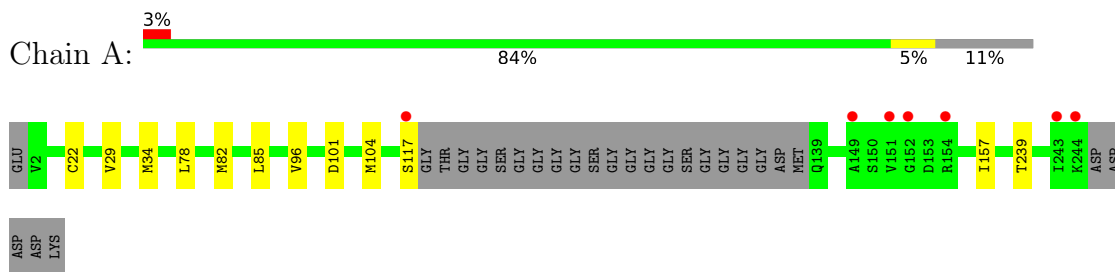
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	190	190	190	0	0
4	B	176	176	176	0	0
4	C	206	206	206	0	0
4	D	184	184	184	0	0

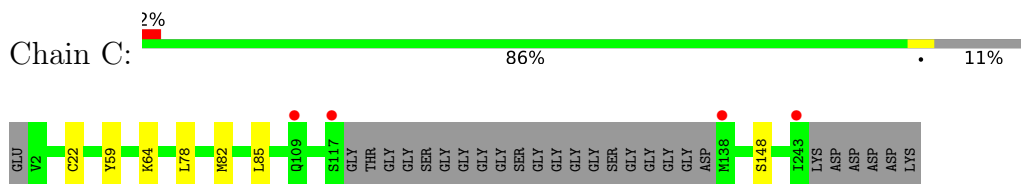
### 3 Residue-property plots [i](#)

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.

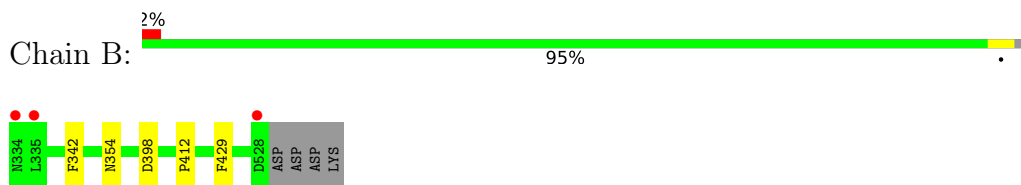
- Molecule 1: pT1611 single-chain Fv



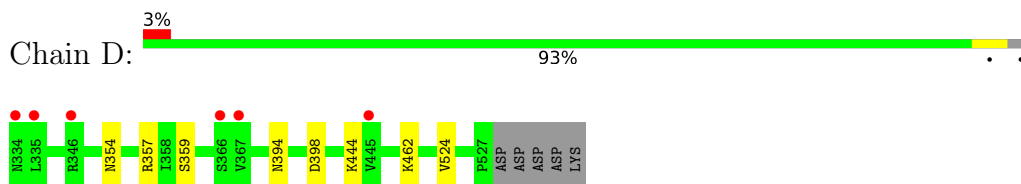
- Molecule 1: pT1611 single-chain Fv



- Molecule 2: Spike protein S1



- Molecule 2: Spike protein S1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.29Å 90.95Å 154.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.00 – 1.60 49.00 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.00-1.60) 99.6 (49.00-1.60)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.41 (at 1.60Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.195 , 0.208 0.194 , 0.207	Depositor DCC
$R_{free}$ test set	1270 reflections (0.78%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.4	Xtrriage
Anisotropy	0.409	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 42.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.009 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7239	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% 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 [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	0.34	0/1723	0.62	0/2337
1	C	0.37	0/1722	0.63	0/2336
2	B	0.38	0/1588	0.60	0/2162
2	D	0.37	0/1580	0.59	0/2151
All	All	0.37	0/6613	0.61	0/8986

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	A	1688	0	1649	5	0
1	C	1687	0	1645	3	0
2	B	1544	0	1456	3	0
2	D	1536	0	1452	6	0
3	B	14	0	13	1	0
3	D	14	0	13	0	0
4	A	190	0	0	0	0
4	B	176	0	0	0	0
4	C	206	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	184	0	0	1	0
All	All	7239	0	6228	17	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 1.

All (17) 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:D:357:ARG:NH2	2:D:394:ASN:OD1	2.25	0.70
2:D:444:LYS:NZ	4:D:1101:HOH:O	2.32	0.59
1:C:82:MET:HE2	1:C:85:LEU:HD21	1.85	0.58
1:A:82:MET:HE2	1:A:85:LEU:HD21	1.95	0.49
2:D:359:SER:HA	2:D:524:VAL:HG22	1.95	0.48
2:B:342:PHE:HB2	3:B:1000:NAG:H82	1.97	0.47
1:C:22:CYS:HB3	1:C:78:LEU:HB3	1.98	0.45
2:D:462:LYS:HB2	2:D:462:LYS:HE2	1.66	0.45
1:A:22:CYS:HB3	1:A:78:LEU:HB3	1.99	0.44
1:A:96:VAL:HG21	1:A:104:MET:HG2	2.01	0.43
2:B:354:ASN:O	2:B:398:ASP:HA	2.18	0.43
2:D:354:ASN:O	2:D:398:ASP:HA	2.19	0.43
2:B:412:PRO:HG3	2:B:429:PHE:HB3	2.00	0.42
1:A:157:ILE:HG12	1:A:239:THR:HG21	2.01	0.42
1:C:59:TYR:HB2	1:C:64:LYS:HD2	2.01	0.41
2:D:357:ARG:HH11	2:D:357:ARG:CG	2.33	0.41
1:A:29:VAL:HG13	1:A:34:MET:HG3	2.03	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	218/249 (88%)	214 (98%)	4 (2%)	0	100	100
1	C	218/249 (88%)	213 (98%)	5 (2%)	0	100	100
2	B	193/199 (97%)	187 (97%)	6 (3%)	0	100	100
2	D	192/199 (96%)	188 (98%)	4 (2%)	0	100	100
All	All	821/896 (92%)	802 (98%)	19 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	184/196 (94%)	182 (99%)	2 (1%)	73	57
1	C	184/196 (94%)	183 (100%)	1 (0%)	88	80
2	B	168/172 (98%)	168 (100%)	0	100	100
2	D	167/172 (97%)	167 (100%)	0	100	100
All	All	703/736 (96%)	700 (100%)	3 (0%)	91	84

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

Mol	Chain	Res	Type
1	A	101	ASP
1	A	117	SER
1	C	148	SER

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

Mol	Chain	Res	Type
2	B	474	GLN
2	B	498	GLN
1	C	3	GLN
2	D	450	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 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
3	NAG	D	1000	2	14,14,15	0.33	0	17,19,21	0.66	1 (5%)
3	NAG	B	1000	2	14,14,15	0.50	0	17,19,21	0.59	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	D	1000	2	-	0/6/23/26	0/1/1/1
3	NAG	B	1000	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	D	1000	NAG	C1-O5-C5	2.40	115.44	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1000	NAG	1	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, 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	222/249 (89%)	0.01	7 (3%) 47 44	20, 27, 48, 62	0
1	C	222/249 (89%)	-0.00	4 (1%) 68 67	19, 26, 42, 51	0
2	B	195/199 (97%)	0.05	3 (1%) 73 73	17, 26, 42, 58	0
2	D	194/199 (97%)	-0.02	6 (3%) 49 46	18, 26, 39, 46	0
All	All	833/896 (92%)	0.01	20 (2%) 59 56	17, 26, 43, 62	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	528	ASP	6.8
1	A	243	ILE	6.2
2	B	334	ASN	5.3
1	C	243	ILE	3.9
1	C	138	MET	3.8
2	B	335	LEU	3.4
2	D	367	VAL	3.3
2	D	445	VAL	3.2
2	D	334	ASN	3.0
2	D	346	ARG	2.9
1	A	154	ARG	2.9
2	D	366	SER	2.9
1	A	244	LYS	2.8
2	D	335	LEU	2.8
1	C	117	SER	2.7
1	A	152	GLY	2.4
1	A	149	ALA	2.2
1	A	151	VAL	2.1
1	C	109	GLN	2.1
1	A	117	SER	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](#)

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	B	1000	14/15	0.82	0.14	32,39,47,49	0
3	NAG	D	1000	14/15	0.88	0.15	35,44,52,54	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.