



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2023 – 12:14 AM EDT

PDB ID : 2GHW  
Title : Crystal structure of SARS spike protein receptor binding domain in complex with a neutralizing antibody, 80R  
Authors : Hwang, W.C.; Lin, Y.; Santelli, E.; Sui, J.; Jaroszewski, L.; Stec, B.; Farzan, M.; Marasco, W.A.; Liddington, R.C.  
Deposited on : 2006-03-27  
Resolution : 2.30 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](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  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

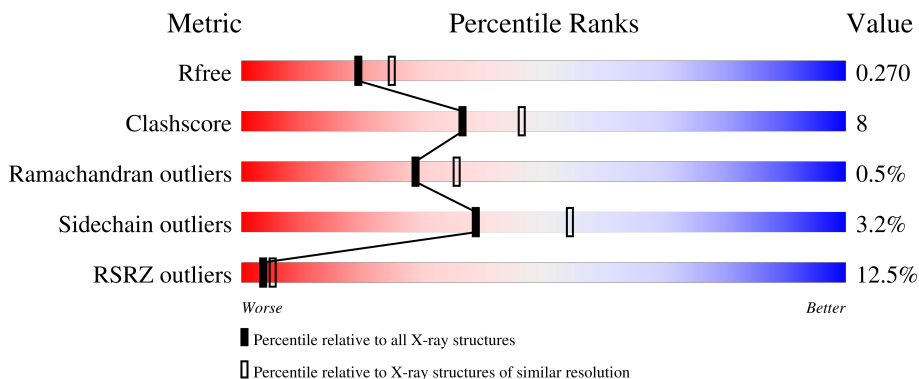
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	203	 17% 81% 12% 6%
1	C	203	 13% 78% 13% 7%
2	B	247	 10% 76% 16% 6%
2	D	247	 8% 78% 14% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
3	CL	A	517	-	-	X	-
3	CL	A	88	-	-	X	-

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7089 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 glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	191	1528	989	248	283	8	0	0	0
1	C	188	1512	978	246	280	8	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	314	MET	-	cloning artifact	UNP P59594
C	315	ALA	-	cloning artifact	UNP P59594
C	316	ASP	-	cloning artifact	UNP P59594
C	511	SER	-	cloning artifact	UNP P59594
C	512	GLY	-	cloning artifact	UNP P59594
C	513	LEU	-	cloning artifact	UNP P59594
C	514	VAL	-	cloning artifact	UNP P59594
C	515	PRO	-	cloning artifact	UNP P59594
C	516	ARG	-	cloning artifact	UNP P59594
A	314	MET	-	cloning artifact	UNP P59594
A	315	ALA	-	cloning artifact	UNP P59594
A	316	ASP	-	cloning artifact	UNP P59594
A	511	SER	-	cloning artifact	UNP P59594
A	512	GLY	-	cloning artifact	UNP P59594
A	513	LEU	-	cloning artifact	UNP P59594
A	514	VAL	-	cloning artifact	UNP P59594
A	515	PRO	-	cloning artifact	UNP P59594
A	516	ARG	-	cloning artifact	UNP P59594

- Molecule 2 is a protein called anti-sars scFv antibody, 80R.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	232	1785	1118	312	349	6	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	231	1789	1120	314	349	6	0	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	Cl	0	0
			4	4		
3	C	1	Total	Cl	0	0
			1	1		

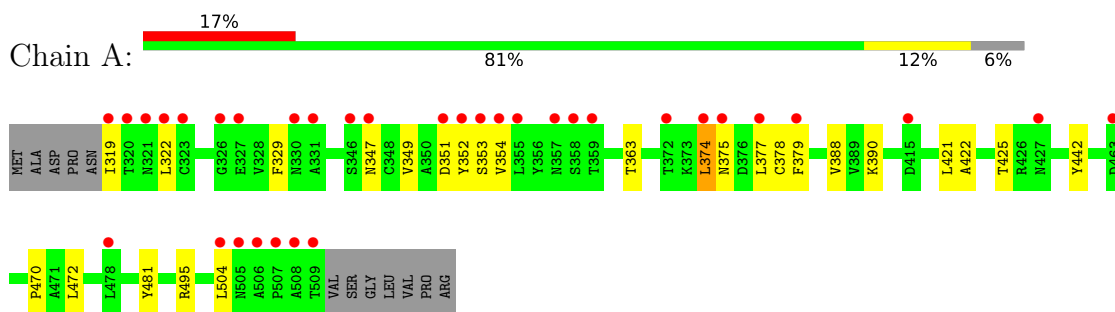
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	108	Total	O	0	0
			108	108		
4	B	118	Total	O	0	0
			118	118		
4	C	115	Total	O	0	0
			115	115		
4	D	129	Total	O	0	0
			129	129		

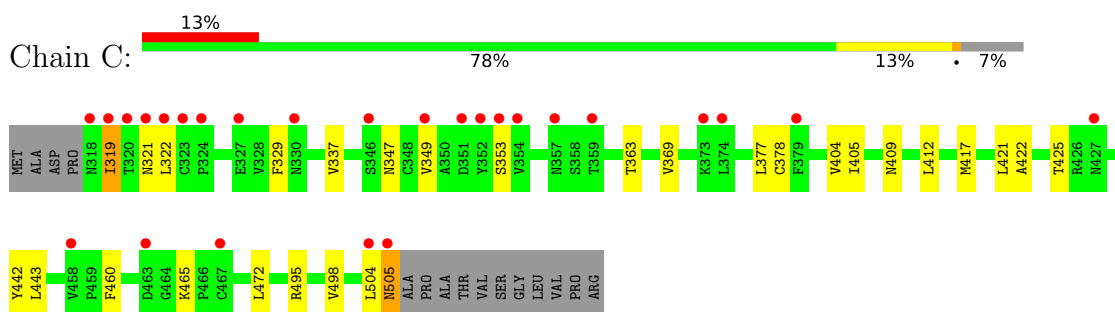
### 3 Residue-property plots [i](#)

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.

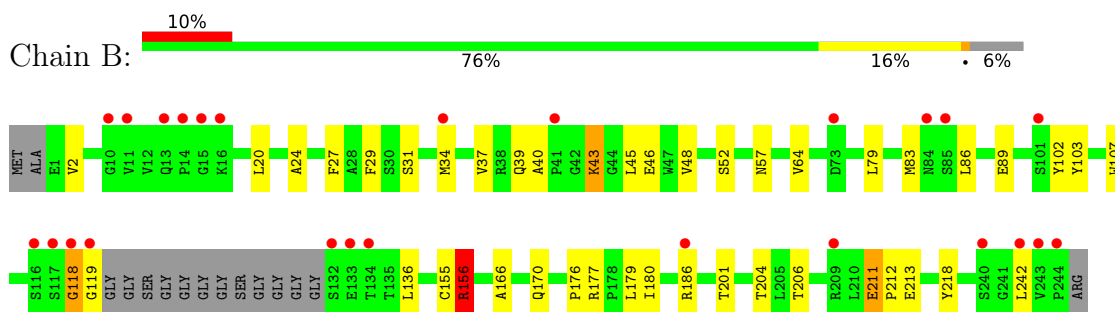
- Molecule 1: Spike glycoprotein



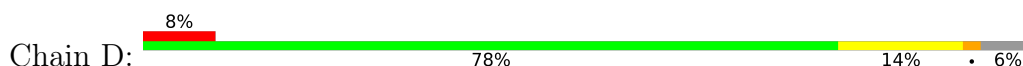
- Molecule 1: Spike glycoprotein

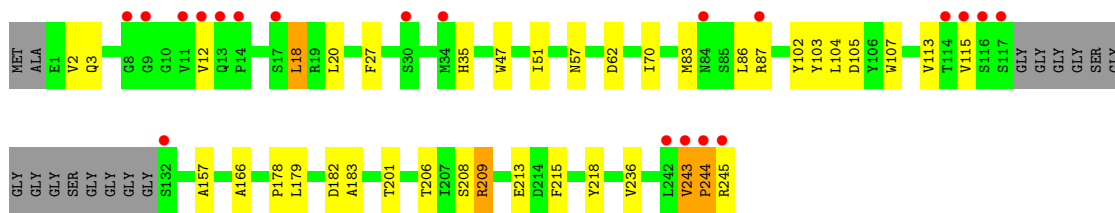


- Molecule 2: anti-sars scFv antibody, 80R



- Molecule 2: anti-sars scFv antibody, 80R





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.45Å 175.90Å 67.56Å 90.00° 96.55° 90.00°	Depositor
Resolution (Å)	45.55 – 2.30 47.14 – 2.19	Depositor EDS
% Data completeness (in resolution range)	93.8 (45.55-2.30) 91.5 (47.14-2.19)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.18Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.248 , 0.295 0.263 , 0.270	Depositor DCC
$R_{free}$ test set	2617 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.7	Xtrriage
Anisotropy	1.024	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 30.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	7089	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.61% 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: CL

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	0.47	0/1576	0.60	0/2152
1	C	0.47	0/1559	0.61	0/2127
2	B	0.42	0/1827	0.65	1/2481 (0.0%)
2	D	0.44	0/1831	0.61	0/2485
All	All	0.45	0/6793	0.62	1/9245 (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	B	0	1
2	D	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	156	ARG	NE-CZ-NH1	7.00	123.80	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	102	TYR	Peptide
2	D	102	TYR	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	1528	0	1455	26	0
1	C	1512	0	1439	20	0
2	B	1785	0	1722	27	0
2	D	1789	0	1729	29	0
3	A	4	0	0	5	0
3	C	1	0	0	1	0
4	A	108	0	0	4	0
4	B	118	0	0	2	0
4	C	115	0	0	0	0
4	D	129	0	0	1	0
All	All	7089	0	6345	102	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 8.

The worst 5 of 102 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:39:GLN:HE22	2:B:170:GLN:HE22	1.31	0.77
1:A:472:LEU:HD13	2:B:206:THR:HG21	1.67	0.75
3:A:88:CL:CL	4:B:411:HOH:O	2.42	0.75
2:B:40:ALA:HB3	2:B:43:LYS:HG2	1.71	0.72
1:A:322:LEU:HD12	1:A:349:VAL:CG2	2.20	0.72

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	189/203 (93%)	178 (94%)	11 (6%)	0	100	100
1	C	186/203 (92%)	171 (92%)	15 (8%)	0	100	100
2	B	228/247 (92%)	211 (92%)	16 (7%)	1 (0%)	34	42
2	D	227/247 (92%)	214 (94%)	10 (4%)	3 (1%)	12	12
All	All	830/900 (92%)	774 (93%)	52 (6%)	4 (0%)	29	35

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	209	ARG
2	D	243	VAL
2	B	118	GLY
2	D	244	PRO

### 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	167/177 (94%)	164 (98%)	3 (2%)	59	75
1	C	166/177 (94%)	161 (97%)	5 (3%)	41	57
2	B	192/196 (98%)	183 (95%)	9 (5%)	26	37
2	D	193/196 (98%)	187 (97%)	6 (3%)	40	55
All	All	718/746 (96%)	695 (97%)	23 (3%)	39	54

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

Mol	Chain	Res	Type
1	C	353	SER
2	D	3	GLN
1	C	505	ASN
2	D	18	LEU
2	B	177	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	57	ASN
2	D	174	GLN
2	B	174	GLN
1	C	321	ASN
1	C	347	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](#)

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 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, 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	191/203 (94%)	1.28	34 (17%) <b>1</b> <b>1</b>	15, 20, 32, 42	1 (0%)
1	C	188/203 (92%)	1.14	26 (13%) <b>2</b> <b>4</b>	15, 21, 30, 41	0
2	B	232/247 (93%)	0.88	25 (10%) <b>5</b> <b>8</b>	14, 20, 26, 38	2 (0%)
2	D	231/247 (93%)	0.92	20 (8%) <b>10</b> <b>14</b>	15, 20, 25, 44	2 (0%)
All	All	842/900 (93%)	1.04	105 (12%) <b>3</b> <b>5</b>	14, 20, 29, 44	5 (0%)

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	244	PRO	10.5
1	A	507	PRO	8.8
1	A	506	ALA	8.4
1	A	319	ILE	8.3
2	B	244	PRO	7.1

### 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	CL	A	246	1/1	0.93	0.29	14,14,14,14	0
3	CL	A	88	1/1	0.95	0.24	3,3,3,3	0
3	CL	A	517	1/1	0.95	0.27	17,17,17,17	0
3	CL	A	142	1/1	0.96	0.21	8,8,8,8	0
3	CL	C	20	1/1	0.98	0.15	11,11,11,11	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.