



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 22, 2023 – 09:58 PM JST

PDB ID : 7XEG  
Title : SARS-CoV-2-Beta-RBD and CB6-092-Fab complex  
Authors : Wang, Y.; Feng, Y.  
Deposited on : 2022-03-31  
Resolution : 2.69 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 : 2.36  
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

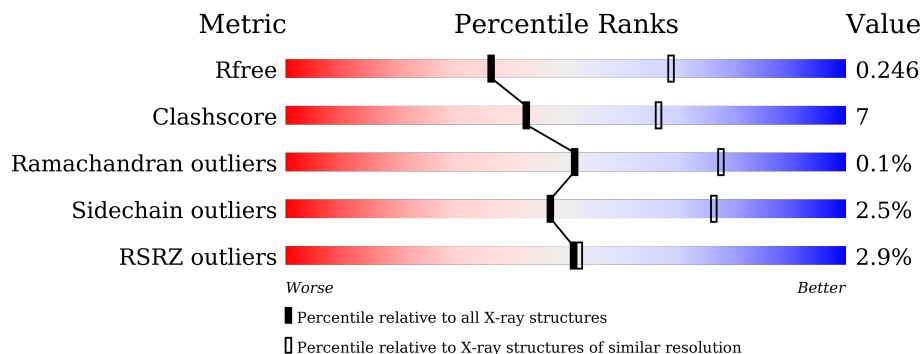
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.69 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	227	
1	B	227	
2	C	233	
2	E	233	
3	D	216	
3	F	216	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9667 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
			Total	C	N	O	S			
1	A	196	1555	999	259	289	8	0	0	0
1	B	195	1546	993	257	288	8	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	ASN	LYS	variant	UNP P0DTC2
A	484	LYS	GLU	variant	UNP P0DTC2
A	501	TYR	ASN	variant	UNP P0DTC2
A	538	HIS	-	expression tag	UNP P0DTC2
A	539	HIS	-	expression tag	UNP P0DTC2
A	540	HIS	-	expression tag	UNP P0DTC2
A	541	HIS	-	expression tag	UNP P0DTC2
A	542	HIS	-	expression tag	UNP P0DTC2
A	543	HIS	-	expression tag	UNP P0DTC2
A	544	HIS	-	expression tag	UNP P0DTC2
A	545	HIS	-	expression tag	UNP P0DTC2
B	417	ASN	LYS	variant	UNP P0DTC2
B	484	LYS	GLU	variant	UNP P0DTC2
B	501	TYR	ASN	variant	UNP P0DTC2
B	538	HIS	-	expression tag	UNP P0DTC2
B	539	HIS	-	expression tag	UNP P0DTC2
B	540	HIS	-	expression tag	UNP P0DTC2
B	541	HIS	-	expression tag	UNP P0DTC2
B	542	HIS	-	expression tag	UNP P0DTC2
B	543	HIS	-	expression tag	UNP P0DTC2
B	544	HIS	-	expression tag	UNP P0DTC2
B	545	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called CB6-092-Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	218	Total 1632	C 1035	N 271	O 318	S 8	0	0	0
2	E	218	Total 1632	C 1035	N 271	O 318	S 8	0	0	0

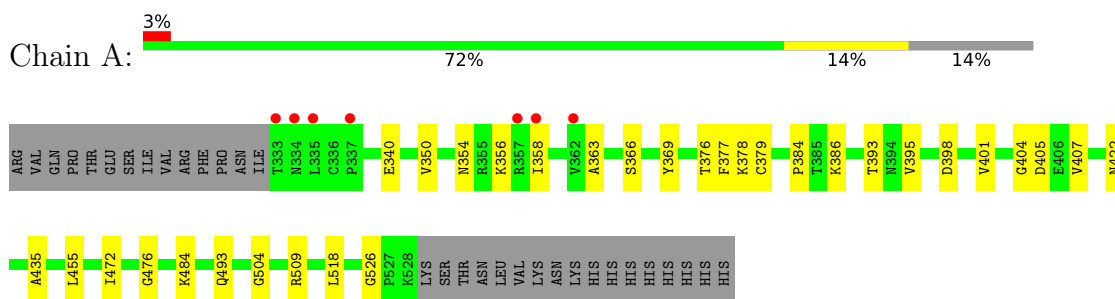
- Molecule 3 is a protein called CB6-092-Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	215	Total 1651	C 1030	N 277	O 339	S 5	0	0	0
3	F	215	Total 1651	C 1030	N 277	O 339	S 5	0	0	0

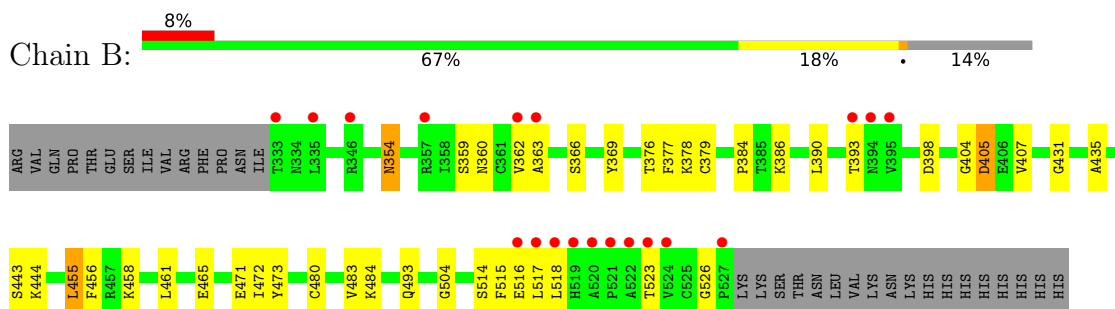
### 3 Residue-property plots [i](#)

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.

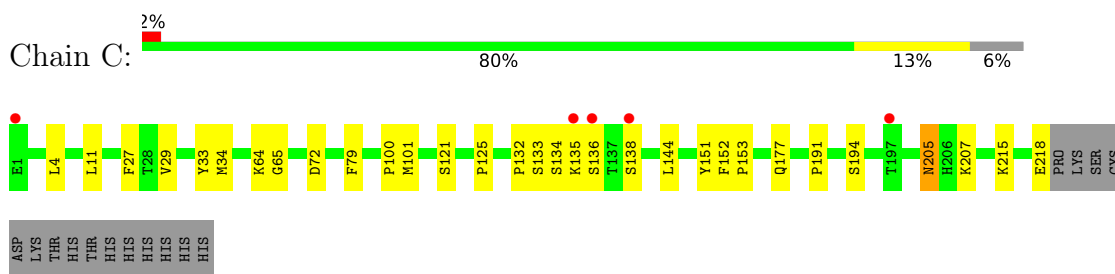
- Molecule 1: Spike protein S1



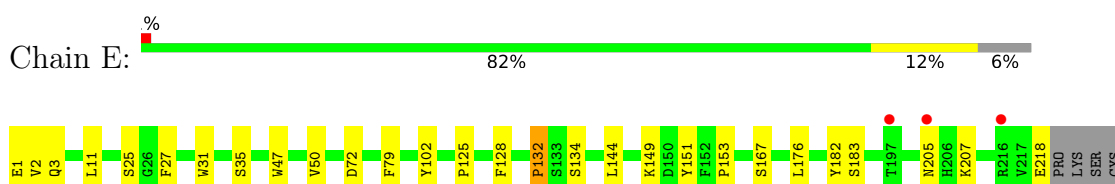
- Molecule 1: Spike protein S1



- Molecule 2: CB6-092-Fab heavy chain




- Molecule 2: CB6-092-Fab heavy chain



ASP  
LYS  
THR  
HIS  
THR  
HIS  
HIS  
HIS  
HIS  
HIS


- Molecule 3: CB6-092-Fab light chain

Chain D:  81% 18%

D1 I2 V3 L11 S12 T22 C23 R24 A25 S26 Q27 R28 L29 L33 Q37 L47 I48 Y49 S53 S63 G64 S65 T69 D70 F71 T72 F83 Q89 Q90 S91 E107 V117 F118 P122 V134 D163 Q157 N160 S161 Q162 E163 Y175

S178 V193 K209 S210 F211 N212 E215 CYS

- Molecule 3: CB6-092-Fab light chain

Chain F:  % 77% 20%

B1 Q6 L11 S12 V19 Q37 L47 S56 G66 F71 F83 Q89 Q90 S91 A92 T104 E107 I108 K109 P122 S123 D124 E125 S129 V134 Y142 P143 B144 E145 A146 K147 V148 Q149 M150 K151 V152 D153 L156 Q157 V165 Q168

D169 S170 Y175 S179 T182 L183 S184 D187 Y188 E189 K192 V193 V194 T199 P206 V207 R213 G214 E215 CYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.25Å 113.24Å 102.10Å 90.00° 100.58° 90.00°	Depositor
Resolution (Å)	39.24 – 2.69 49.31 – 2.69	Depositor EDS
% Data completeness (in resolution range)	95.8 (39.24-2.69) 95.9 (49.31-2.69)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.79 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.195 , 0.248 0.196 , 0.246	Depositor DCC
$R_{free}$ test set	2385 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.8	Xtrriage
Anisotropy	0.100	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 39.9	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	0.92	EDS
Total number of atoms	9667	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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](#)

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	0.32	0/1600	0.53	0/2178
1	B	0.33	0/1591	0.58	1/2167 (0.0%)
2	C	0.30	0/1673	0.53	0/2283
2	E	0.29	0/1673	0.54	0/2283
3	D	0.30	0/1685	0.55	1/2287 (0.0%)
3	F	0.36	0/1685	0.59	2/2287 (0.1%)
All	All	0.32	0/9907	0.56	4/13485 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	455	LEU	CA-CB-CG	8.19	134.12	115.30
3	D	153	ASP	CB-CG-OD1	7.45	125.00	118.30
3	F	157	GLN	CA-CB-CG	-6.31	99.52	113.40
3	F	193	VAL	CG1-CB-CG2	6.00	120.49	110.90

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	1555	0	1476	23	0
1	B	1546	0	1463	27	0
2	C	1632	0	1595	22	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	1632	0	1595	20	1
3	D	1651	0	1605	22	0
3	F	1651	0	1605	31	0
All	All	9667	0	9339	129	1

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 129 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:E:47:TRP:HE1	2:E:50:VAL:HG23	1.48	0.79
1:A:455:LEU:HD12	2:C:100:PRO:HG2	1.64	0.79
2:C:132:PRO:HG3	2:C:144:LEU:HB3	1.68	0.76
3:D:193:VAL:HG22	3:D:212:ASN:OD1	1.84	0.76
1:B:362:VAL:HG13	1:B:526:GLY:HA2	1.68	0.75

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:121:SER:OG	2:E:134:SER:O[2_455]	2.05	0.15

## 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	194/227 (86%)	185 (95%)	8 (4%)	1 (0%)	29	54
1	B	193/227 (85%)	180 (93%)	13 (7%)	0	100	100
2	C	216/233 (93%)	207 (96%)	9 (4%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	216/233 (93%)	209 (97%)	7 (3%)	0	100	100
3	D	213/216 (99%)	203 (95%)	10 (5%)	0	100	100
3	F	213/216 (99%)	202 (95%)	11 (5%)	0	100	100
All	All	1245/1352 (92%)	1186 (95%)	58 (5%)	1 (0%)	51	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	386	LYS

### 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	169/200 (84%)	169 (100%)	0	100	100
1	B	168/200 (84%)	161 (96%)	7 (4%)	30	58
2	C	182/197 (92%)	179 (98%)	3 (2%)	62	85
2	E	182/197 (92%)	178 (98%)	4 (2%)	52	79
3	D	189/190 (100%)	183 (97%)	6 (3%)	39	68
3	F	189/190 (100%)	182 (96%)	7 (4%)	34	63
All	All	1079/1174 (92%)	1052 (98%)	27 (2%)	47	76

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	153	ASP
2	E	132	PRO
3	F	129	SER
2	E	102	TYR
2	E	167	SER

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

Mol	Chain	Res	Type
3	D	27	GLN
2	E	111	GLN
2	E	205	ASN
3	F	157	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 [i](#)

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	196/227 (86%)	0.13	7 (3%) 42 42	21, 37, 74, 105	0
1	B	195/227 (85%)	0.52	19 (9%) 7 6	27, 46, 87, 125	0
2	C	218/233 (93%)	-0.14	5 (2%) 60 62	23, 37, 65, 129	0
2	E	218/233 (93%)	-0.12	3 (1%) 75 77	20, 41, 63, 82	0
3	D	215/216 (99%)	-0.22	1 (0%) 91 92	24, 37, 57, 99	0
3	F	215/216 (99%)	0.03	2 (0%) 84 85	30, 46, 84, 97	0
All	All	1257/1352 (92%)	0.02	37 (2%) 51 52	20, 40, 74, 129	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	333	THR	7.2
1	B	521	PRO	7.1
1	B	518	LEU	7.0
1	B	519	HIS	6.9
1	B	523	THR	5.7

### 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.