



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 22, 2023 – 10:41 PM JST

PDB ID : 7XSA
Title : Crystal structure of SARS-CoV-2 spike receptor binding domain bound with P2S-2E9 Fab
Authors : Wang, X.; Wang, Z.
Deposited on : 2022-05-13
Resolution : 2.20 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

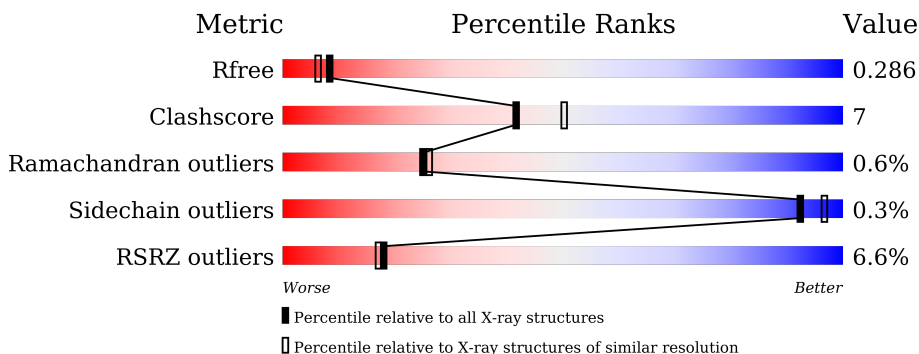
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	B	218	 2% (poor fit), 82% (0-1 outliers), 16% (2-3 outliers), 0% (4+ outliers)
1	H	218	 8% (poor fit), 77% (0-1 outliers), 20% (2-3 outliers), 0% (4+ outliers)
2	C	215	 0% (poor fit), 88% (0-1 outliers), 9% (2-3 outliers), 0% (4+ outliers)
2	I	215	 23% (poor fit), 73% (0-1 outliers), 24% (2-3 outliers), 0% (4+ outliers)
3	E	242	 0% (poor fit), 70% (0-1 outliers), 7% (2-3 outliers), 23% (4+ outliers)
3	J	242	 3% (poor fit), 65% (0-1 outliers), 12% (2-3 outliers), 22% (4+ outliers)

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9284 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 P2S-2E9 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	215	Total 1627	C 1035	N 267	O 319	S 6	0	0	0
1	H	212	Total 1610	C 1028	N 263	O 313	S 6	0	0	0

- Molecule 2 is a protein called P2S-2E9 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	209	Total 1532	C 954	N 252	O 320	S 6	0	0	0
2	I	209	Total 1529	C 950	N 252	O 321	S 6	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	E	187	Total 1490	C 957	N 247	O 279	S 7	0	0	0
3	J	188	Total 1496	C 960	N 248	O 280	S 8	0	0	0

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	292	MET	-	initiating methionine	UNP P0DTC2
E	293	LEU	-	expression tag	UNP P0DTC2
E	294	LEU	-	expression tag	UNP P0DTC2
E	295	VAL	-	expression tag	UNP P0DTC2
E	296	ASN	-	expression tag	UNP P0DTC2
E	297	GLN	-	expression tag	UNP P0DTC2
E	298	SER	-	expression tag	UNP P0DTC2
E	299	HIS	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	300	GLN	-	expression tag	UNP P0DTC2
E	301	GLY	-	expression tag	UNP P0DTC2
E	302	PHE	-	expression tag	UNP P0DTC2
E	303	ASN	-	expression tag	UNP P0DTC2
E	304	LYS	-	expression tag	UNP P0DTC2
E	305	GLU	-	expression tag	UNP P0DTC2
E	306	HIS	-	expression tag	UNP P0DTC2
E	307	THR	-	expression tag	UNP P0DTC2
E	308	SER	-	expression tag	UNP P0DTC2
E	309	LYS	-	expression tag	UNP P0DTC2
E	310	MET	-	expression tag	UNP P0DTC2
E	311	VAL	-	expression tag	UNP P0DTC2
E	312	SER	-	expression tag	UNP P0DTC2
E	313	ALA	-	expression tag	UNP P0DTC2
E	314	ILE	-	expression tag	UNP P0DTC2
E	315	VAL	-	expression tag	UNP P0DTC2
E	316	LEU	-	expression tag	UNP P0DTC2
E	317	TYR	-	expression tag	UNP P0DTC2
E	318	VAL	-	expression tag	UNP P0DTC2
E	319	LEU	-	expression tag	UNP P0DTC2
E	320	LEU	-	expression tag	UNP P0DTC2
E	321	ALA	-	expression tag	UNP P0DTC2
E	322	ALA	-	expression tag	UNP P0DTC2
E	323	ALA	-	expression tag	UNP P0DTC2
E	324	ALA	-	expression tag	UNP P0DTC2
E	325	HIS	-	expression tag	UNP P0DTC2
E	326	SER	-	expression tag	UNP P0DTC2
E	327	ALA	-	expression tag	UNP P0DTC2
E	328	PHE	-	expression tag	UNP P0DTC2
E	329	ALA	-	expression tag	UNP P0DTC2
E	330	ALA	-	expression tag	UNP P0DTC2
E	331	ASP	-	expression tag	UNP P0DTC2
E	332	PRO	-	expression tag	UNP P0DTC2
E	417	ASN	LYS	variant	UNP P0DTC2
E	484	LYS	GLU	variant	UNP P0DTC2
E	501	TYR	ASN	variant	UNP P0DTC2
E	528	HIS	-	expression tag	UNP P0DTC2
E	529	HIS	-	expression tag	UNP P0DTC2
E	530	HIS	-	expression tag	UNP P0DTC2
E	531	HIS	-	expression tag	UNP P0DTC2
E	532	HIS	-	expression tag	UNP P0DTC2
E	533	HIS	-	expression tag	UNP P0DTC2

Continued on next page...

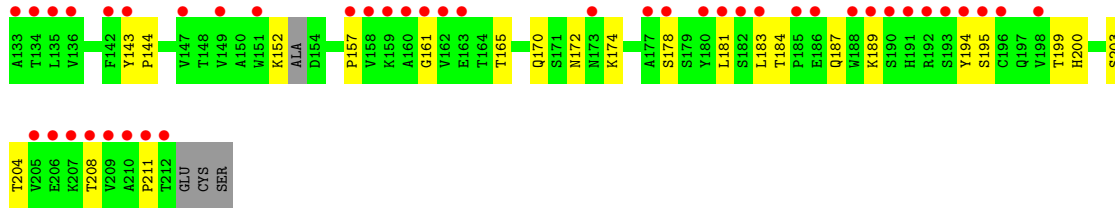
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	292	MET	-	initiating methionine	UNP P0DTC2
J	293	LEU	-	expression tag	UNP P0DTC2
J	294	LEU	-	expression tag	UNP P0DTC2
J	295	VAL	-	expression tag	UNP P0DTC2
J	296	ASN	-	expression tag	UNP P0DTC2
J	297	GLN	-	expression tag	UNP P0DTC2
J	298	SER	-	expression tag	UNP P0DTC2
J	299	HIS	-	expression tag	UNP P0DTC2
J	300	GLN	-	expression tag	UNP P0DTC2
J	301	GLY	-	expression tag	UNP P0DTC2
J	302	PHE	-	expression tag	UNP P0DTC2
J	303	ASN	-	expression tag	UNP P0DTC2
J	304	LYS	-	expression tag	UNP P0DTC2
J	305	GLU	-	expression tag	UNP P0DTC2
J	306	HIS	-	expression tag	UNP P0DTC2
J	307	THR	-	expression tag	UNP P0DTC2
J	308	SER	-	expression tag	UNP P0DTC2
J	309	LYS	-	expression tag	UNP P0DTC2
J	310	MET	-	expression tag	UNP P0DTC2
J	311	VAL	-	expression tag	UNP P0DTC2
J	312	SER	-	expression tag	UNP P0DTC2
J	313	ALA	-	expression tag	UNP P0DTC2
J	314	ILE	-	expression tag	UNP P0DTC2
J	315	VAL	-	expression tag	UNP P0DTC2
J	316	LEU	-	expression tag	UNP P0DTC2
J	317	TYR	-	expression tag	UNP P0DTC2
J	318	VAL	-	expression tag	UNP P0DTC2
J	319	LEU	-	expression tag	UNP P0DTC2
J	320	LEU	-	expression tag	UNP P0DTC2
J	321	ALA	-	expression tag	UNP P0DTC2
J	322	ALA	-	expression tag	UNP P0DTC2
J	323	ALA	-	expression tag	UNP P0DTC2
J	324	ALA	-	expression tag	UNP P0DTC2
J	325	HIS	-	expression tag	UNP P0DTC2
J	326	SER	-	expression tag	UNP P0DTC2
J	327	ALA	-	expression tag	UNP P0DTC2
J	328	PHE	-	expression tag	UNP P0DTC2
J	329	ALA	-	expression tag	UNP P0DTC2
J	330	ALA	-	expression tag	UNP P0DTC2
J	331	ASP	-	expression tag	UNP P0DTC2
J	332	PRO	-	expression tag	UNP P0DTC2
J	417	ASN	LYS	variant	UNP P0DTC2

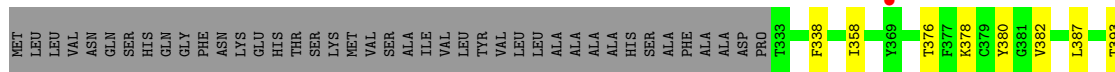
Continued on next page...

Continued from previous page...

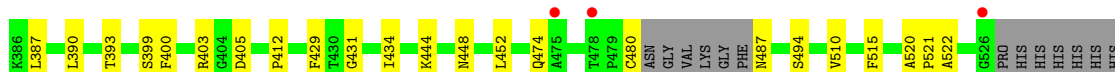
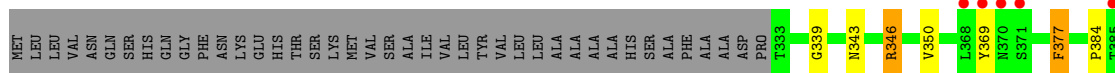
Chain	Residue	Modelled	Actual	Comment	Reference
J	484	LYS	GLU	variant	UNP P0DTC2
J	501	TYR	ASN	variant	UNP P0DTC2
J	528	HIS	-	expression tag	UNP P0DTC2
J	529	HIS	-	expression tag	UNP P0DTC2
J	530	HIS	-	expression tag	UNP P0DTC2
J	531	HIS	-	expression tag	UNP P0DTC2
J	532	HIS	-	expression tag	UNP P0DTC2
J	533	HIS	-	expression tag	UNP P0DTC2



● Molecule 3: Spike protein S1



● Molecule 3: Spike protein S1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	80.95Å 70.80Å 127.80Å 90.00° 95.98° 90.00°	Depositor
Resolution (Å)	32.41 – 2.20 47.89 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (32.41-2.20) 98.9 (47.89-2.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.45 (at 2.20Å)	Xtrriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.231 , 0.283 0.234 , 0.286	Depositor DCC
R_{free} test set	3585 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	49.2	Xtrriage
Anisotropy	0.176	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9284	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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	B	0.45	0/1664	0.65	0/2268
1	H	0.41	0/1649	0.64	0/2252
2	C	0.47	0/1568	0.69	0/2141
2	I	0.43	0/1562	0.62	0/2129
3	E	0.50	0/1532	0.64	0/2086
3	J	0.53	0/1538	0.67	0/2094
All	All	0.47	0/9513	0.65	0/12970

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	B	1627	0	1624	25	0
1	H	1610	0	1607	30	0
2	C	1532	0	1470	15	0
2	I	1529	0	1474	35	0
3	E	1490	0	1409	14	0
3	J	1496	0	1412	18	0
All	All	9284	0	8996	127	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 127 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:I:184:THR:HB	2:I:187:GLN:HB2	1.57	0.86
2:I:53:VAL:HG23	2:I:54:THR:HG23	1.58	0.86
3:J:444:LYS:HG2	3:J:448:ASN:HB2	1.65	0.78
1:H:92:THR:HG22	1:H:117:VAL:H	1.48	0.76
2:C:183:LEU:HD11	2:C:188:TRP:HB2	1.71	0.71

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	B	210/218 (96%)	200 (95%)	9 (4%)	1 (0%)	29	31
1	H	206/218 (94%)	192 (93%)	12 (6%)	2 (1%)	15	14
2	C	207/215 (96%)	195 (94%)	11 (5%)	1 (0%)	29	31
2	I	205/215 (95%)	191 (93%)	12 (6%)	2 (1%)	15	14
3	E	183/242 (76%)	176 (96%)	7 (4%)	0	100	100
3	J	184/242 (76%)	173 (94%)	10 (5%)	1 (0%)	29	31
All	All	1195/1350 (88%)	1127 (94%)	61 (5%)	7 (1%)	25	26

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	207	LYS
2	I	211	PRO
3	J	369	TYR
1	H	152	PHE
2	C	191	HIS

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	B	188/191 (98%)	188 (100%)	0	100	100
1	H	186/191 (97%)	186 (100%)	0	100	100
2	C	174/182 (96%)	174 (100%)	0	100	100
2	I	175/182 (96%)	174 (99%)	1 (1%)	86	93
3	E	162/206 (79%)	162 (100%)	0	100	100
3	J	163/206 (79%)	161 (99%)	2 (1%)	71	83
All	All	1048/1158 (90%)	1045 (100%)	3 (0%)	92	97

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

Mol	Chain	Res	Type
2	I	189	LYS
3	J	346	ARG
3	J	377	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	170	HIS
1	H	210	ASN
3	J	487	ASN
2	I	173	ASN
2	I	197	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, 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	B	215/218 (98%)	0.35	4 (1%) 66 65	36, 53, 83, 92	0
1	H	212/218 (97%)	0.59	17 (8%) 12 11	42, 68, 84, 109	0
2	C	209/215 (97%)	0.27	1 (0%) 91 90	37, 56, 76, 86	0
2	I	209/215 (97%)	1.23	49 (23%) 0 0	40, 67, 103, 111	0
3	E	187/242 (77%)	0.34	2 (1%) 80 79	35, 49, 79, 99	0
3	J	188/242 (77%)	0.45	8 (4%) 35 33	30, 48, 84, 101	0
All	All	1220/1350 (90%)	0.54	81 (6%) 18 17	30, 56, 92, 111	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	158	VAL	8.6
2	I	183	LEU	6.3
2	I	180	TYR	6.2
2	I	188	TRP	5.9
3	J	370	ASN	5.1

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.