



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 3, 2020 – 12:03 PM EST

PDB ID : 6W52
Title : Prefusion RSV F bound by neutralizing antibody RSB1
Authors : Harshbarger, W.; Chandramouli, S.; Malito, M.
Deposited on : 2020-03-12
Resolution : 3.74 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 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	:	2.14.6
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)	:	2.14.6

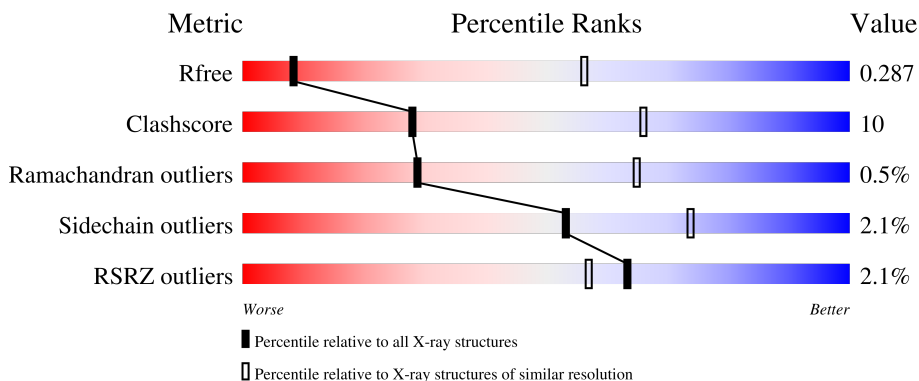
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.74 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	1001 (3.90-3.58)
Clashscore	141614	1063 (3.90-3.58)
Ramachandran outliers	138981	1027 (3.90-3.58)
Sidechain outliers	138945	1023 (3.90-3.58)
RSRZ outliers	127900	1006 (3.92-3.56)

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 on 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	A	82	<div> <div>54%</div> <div>32%</div> <div>•</div> <div>12%</div> </div>
2	B	414	<div> <div>70%</div> <div>20%</div> <div>10%</div> </div>
3	H	283	<div> <div>6%</div> <div>55%</div> <div>19%</div> <div>•</div> <div>24%</div> </div>
4	L	236	<div> <div>%</div> <div>72%</div> <div>18%</div> <div>10%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6632 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 Fusion glycoprotein F0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	72	Total	C	N	O	S	0	0	0
			571	361	93	114	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	PRO	conflict	UNP P03420

- Molecule 2 is a protein called Fusion glycoprotein F1 fused with Fibrin trimerization domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	373	Total	C	N	O	S	0	0	0
			2878	1820	476	562	20			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	152	VAL	ILE	conflict	UNP A0A2H4WLA4
B	155	CYS	SER	conflict	UNP A0A2H4WLA4
B	190	PHE	SER	conflict	UNP A0A2H4WLA4
B	207	LEU	VAL	conflict	UNP A0A2H4WLA4
B	290	CYS	SER	conflict	UNP A0A2H4WLA4
B	514	SER	-	linker	UNP A0A2H4WLA4
B	515	ALA	-	linker	UNP A0A2H4WLA4
B	516	ILE	-	linker	UNP A0A2H4WLA4
B	517	GLY	-	linker	UNP A0A2H4WLA4
B	545	GLY	-	expression tag	UNP A0A2Z5WL46
B	546	GLY	-	expression tag	UNP A0A2Z5WL46
B	547	LEU	-	expression tag	UNP A0A2Z5WL46
B	548	VAL	-	expression tag	UNP A0A2Z5WL46
B	549	PRO	-	expression tag	UNP A0A2Z5WL46

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	550	ARG	-	expression tag	UNP A0A2Z5WL46

- Molecule 3 is a protein called RSB1 Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	214	Total	C	N	O	S	0	0	0
			1598	1012	263	316	7			

- Molecule 4 is a protein called RSB1 Fab Light Chain.

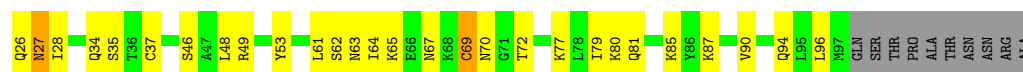
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	213	Total	C	N	O	S	0	0	0
			1585	992	264	323	6			

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.

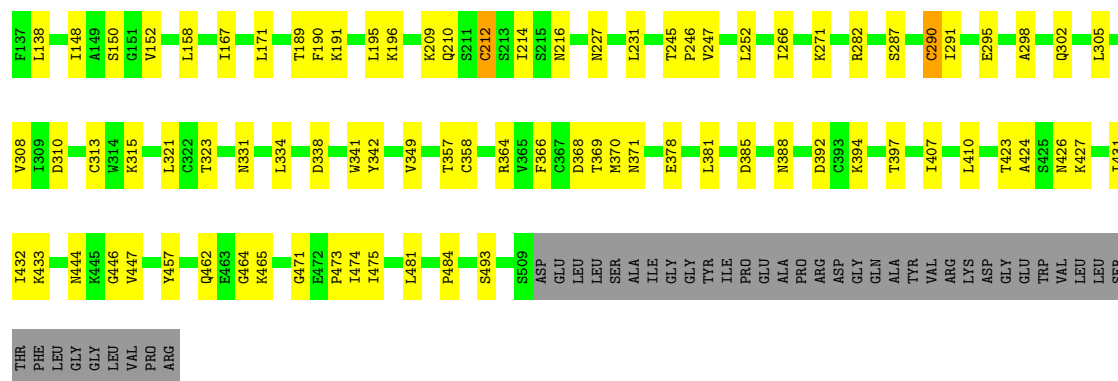
- Molecule 1: Fusion glycoprotein F0

Chain A: 



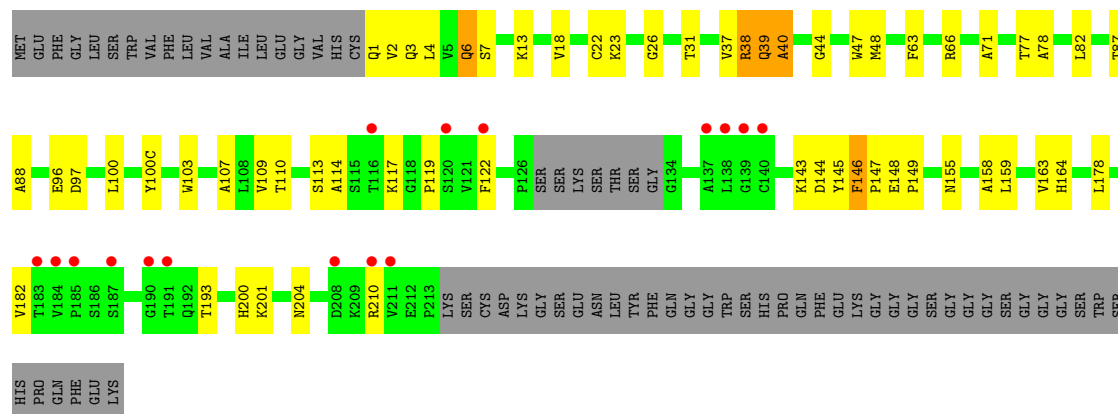
- Molecule 2: Fusion glycoprotein F1 fused with Fibrin trimerization domain

Chain B: 

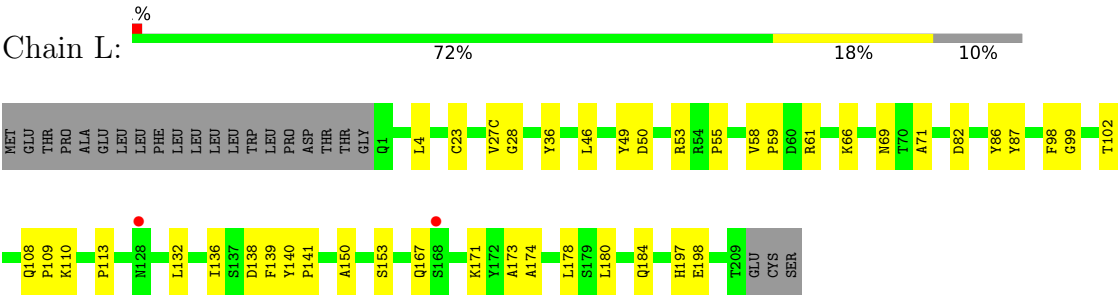


- Molecule 3: RSB1 Fab Heavy Chain

Chain H: 



● Molecule 4: RSB1 Fab Light Chain



4 Data and refinement statistics

Property	Value	Source
Space group	I 21 3	Depositor
Cell constants a, b, c, α , β , γ	190.30Å 190.30Å 190.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.57 – 3.74 44.85 – 3.74	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.57-3.74) 94.6 (44.85-3.74)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.90 (at 3.77Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.218 , 0.286 0.218 , 0.287	Depositor DCC
R_{free} test set	1198 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å ²)	20.0	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 35.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.027 for -l,-k,-h	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	6632	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% 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	A	0.25	0/577	0.49	0/777
2	B	0.25	0/2922	0.42	0/3962
3	H	0.26	0/1636	0.51	0/2232
4	L	0.24	0/1624	0.44	0/2219
All	All	0.25	0/6759	0.46	0/9190

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
3	H	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	H	40	ALA	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	571	0	581	28	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	2878	0	2913	56	0
3	H	1598	0	1563	42	0
4	L	1585	0	1540	30	0
All	All	6632	0	6597	131	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 10.

The worst 5 of 131 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:245:THR:HG23	2:B:246:PRO:HD3	1.62	0.80
2:B:171:LEU:HD13	2:B:191:LYS:HB2	1.72	0.71
3:H:147:PRO:HG3	3:H:200:HIS:NE2	2.08	0.68
1:A:64:ILE:O	1:A:87:LYS:NZ	2.26	0.68
2:B:358:CYS:HA	2:B:366:PHE:O	1.94	0.67

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	70/82 (85%)	67 (96%)	3 (4%)	0	100	100
2	B	371/414 (90%)	356 (96%)	13 (4%)	2 (0%)	29	65
3	H	210/283 (74%)	195 (93%)	13 (6%)	2 (1%)	15	51
4	L	211/236 (89%)	200 (95%)	11 (5%)	0	100	100
All	All	862/1015 (85%)	818 (95%)	40 (5%)	4 (0%)	29	65

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	H	39	GLN
2	B	209	LYS
2	B	212	CYS
3	H	40	ALA

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	65/73 (89%)	61 (94%)	4 (6%)	18	50
2	B	341/373 (91%)	337 (99%)	4 (1%)	71	84
3	H	177/230 (77%)	171 (97%)	6 (3%)	37	64
4	L	178/199 (89%)	176 (99%)	2 (1%)	73	85
All	All	761/875 (87%)	745 (98%)	16 (2%)	53	74

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	481	LEU
3	H	6	GLN
3	H	146	PHE
2	B	290	CYS
3	H	178	LEU

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

Mol	Chain	Res	Type
1	A	81	GLN
3	H	3	GLN
3	H	6	GLN
4	L	197	HIS

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	A	72/82 (87%)	-0.38	0 100 100	14, 47, 115, 133	0
2	B	373/414 (90%)	-0.34	0 100 100	15, 50, 114, 157	0
3	H	214/283 (75%)	0.38	16 (7%) 14 11	36, 104, 154, 170	0
4	L	213/236 (90%)	0.36	2 (0%) 84 81	38, 105, 135, 166	0
All	All	872/1015 (85%)	0.00	18 (2%) 63 57	14, 78, 141, 170	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	139	GLY	4.3
3	H	138	LEU	3.5
4	L	168	SER	3.0
3	H	191	THR	2.7
3	H	120	SER	2.7

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.