wwPDB X-ray Structure Validation Summary Report (i)

Mar 15, 2023 – 06:15 pm GMT

PDB ID	:	8BBO
Title	:	SARS-CoV-2 Delta-RBD complexed with BA.2-36 Fab
Authors	:	Zhou, D.; Ren, J.; Stuart, D.I.
Deposited on	:	2022-10-14
Resolution	:	2.75 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 (i) 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 (1)) were used in the production of this report:

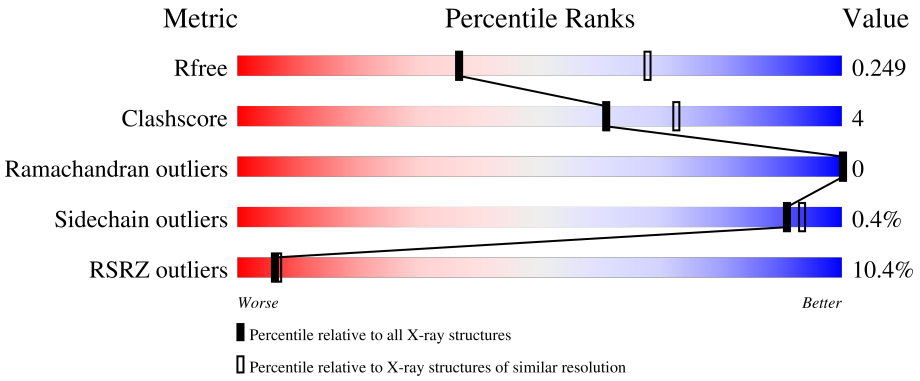
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 <=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	R	202	28% 76% 14% 10%
2	H	224	3% 87% 12%
3	L	213	92% 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4752 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
1	R	181	Total	C	N	O	S	0	0	0
			1450	930	244	269	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	327	HIS	-	expression tag	UNP A0A8B6RM54
R	328	HIS	-	expression tag	UNP A0A8B6RM54
R	329	HIS	-	expression tag	UNP A0A8B6RM54
R	330	HIS	-	expression tag	UNP A0A8B6RM54
R	331	HIS	-	expression tag	UNP A0A8B6RM54
R	332	HIS	-	expression tag	UNP A0A8B6RM54
R	478	LYS	THR	conflict	UNP A0A8B6RM54
R	527	LYS	PRO	conflict	UNP A0A8B6RM54

- Molecule 2 is a protein called IGH@ protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	222	Total	C	N	O	S	0	0	0
			1656	1039	281	331	5			

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	5	VAL	GLN	conflict	UNP Q6GMX6
H	11	ARG	LEU	conflict	UNP Q6GMX6
H	16	GLN	GLU	conflict	UNP Q6GMX6
H	27	ASP	GLY	conflict	UNP Q6GMX6
H	30	ASN	-	insertion	UNP Q6GMX6
H	33	ILE	TYR	conflict	UNP Q6GMX6
H	34	ASN	-	insertion	UNP Q6GMX6
H	37	ASN	SER	conflict	UNP Q6GMX6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	46	GLU	GLY	conflict	UNP Q6GMX6
H	54	PHE	TYR	conflict	UNP Q6GMX6
H	58	THR	SER	conflict	UNP Q6GMX6
H	60	HIS	ASN	conflict	UNP Q6GMX6
H	71	ILE	MET	conflict	UNP Q6GMX6
H	75	ARG	THR	conflict	UNP Q6GMX6
H	79	GLU	GLN	conflict	UNP Q6GMX6
H	83	THR	LYS	conflict	UNP Q6GMX6
H	85	ASN	SER	conflict	UNP Q6GMX6
H	96	PHE	TYR	conflict	UNP Q6GMX6
H	98	GLY	ALA	conflict	UNP Q6GMX6
H	101	GLY	ARG	conflict	UNP Q6GMX6
H	102	THR	PHE	conflict	UNP Q6GMX6
H	103	ASP	THR	conflict	UNP Q6GMX6
H	104	ASP	-	insertion	UNP Q6GMX6
H	106	VAL	PHE	conflict	UNP Q6GMX6
H	216	ARG	LYS	conflict	UNP Q6GMX6

- Molecule 3 is a protein called Immunoglobulin kappa light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	211	Total	C	N	O	S	0	0	0
			1632	1022	278	327	5			

There are 17 discrepancies between the modelled and reference sequences:

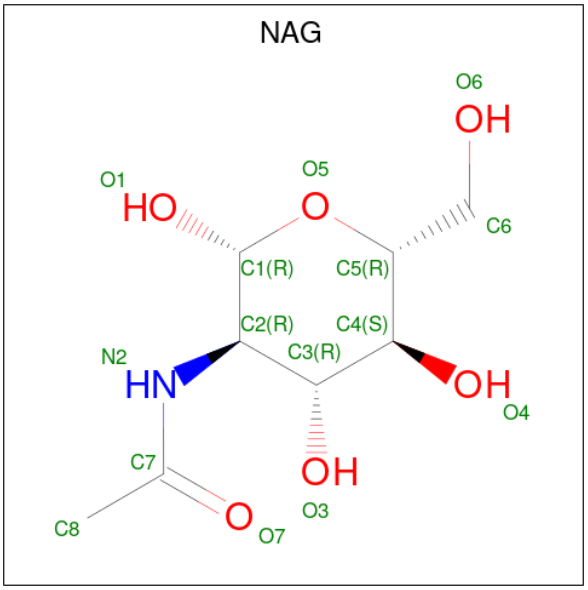
Chain	Residue	Modelled	Actual	Comment	Reference
L	1	ALA	-	expression tag	UNP P0DOX7
L	28	ASP	SER	conflict	UNP P0DOX7
L	31	SER	THR	conflict	UNP P0DOX7
L	48	ILE	MET	conflict	UNP P0DOX7
L	50	ASP	LYS	conflict	UNP P0DOX7
L	55	HIS	GLU	conflict	UNP P0DOX7
L	60	THR	SER	conflict	UNP P0DOX7
L	63	SER	ILE	conflict	UNP P0DOX7
L	85	SER	THR	conflict	UNP P0DOX7
L	92	LYS	ASN	conflict	UNP P0DOX7
L	?	-	ASP	deletion	UNP P0DOX7
L	94	TYR	SER	conflict	UNP P0DOX7
L	95	ARG	LYS	conflict	UNP P0DOX7
L	96	THR	MET	conflict	UNP P0DOX7
L	99	ARG	GLN	conflict	UNP P0DOX7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	105	ILE	VAL	conflict	UNP P0DOX7
L	107	ARG	GLY	conflict	UNP P0DOX7

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



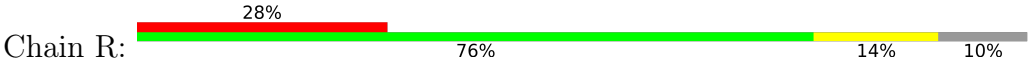
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	R	1	Total	C	N	O	0	0
			14	8	1	5		

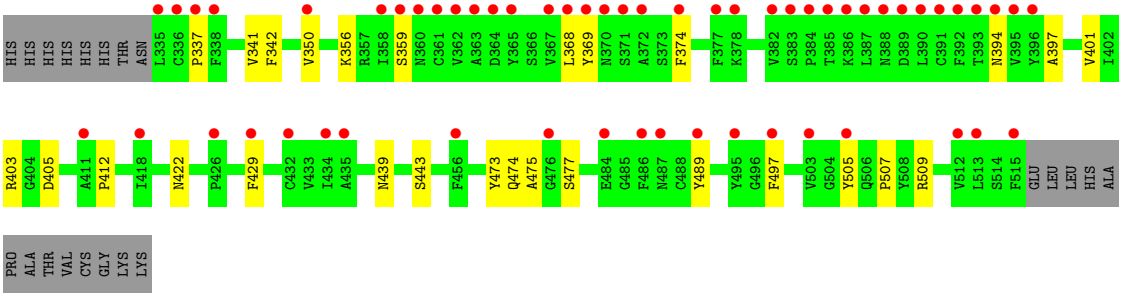
3 Residue-property plots (i)

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

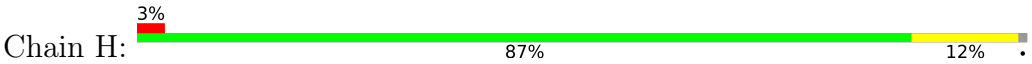
Chain R:





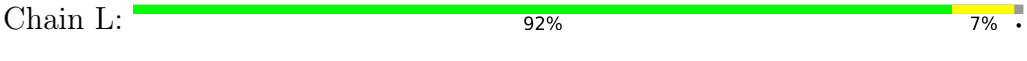
• Molecule 2: IGH@ protein

Chain H:



• Molecule 3: Immunoglobulin kappa light chain

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	99.14Å 99.14Å 86.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.57 – 2.75 54.35 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.57-2.75) 99.8 (54.35-2.75)	Depositor EDS
R_{merge}	0.46	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 2.77Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.224 , 0.252 0.229 , 0.249	Depositor DCC
R_{free} test set	1103 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	64.5	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 35.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.040 for h,-k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4752	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	R	0.25	0/1491	0.47	0/2024
2	H	0.25	0/1695	0.50	0/2316
3	L	0.25	0/1668	0.48	0/2263
All	All	0.25	0/4854	0.48	0/6603

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

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	R	1450	0	1373	17	0
2	H	1656	0	1624	17	0
3	L	1632	0	1592	8	0
4	R	14	0	13	1	0
All	All	4752	0	4602	42	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 4.

The worst 5 of 42 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:H:133:SER:H	2:H:136:SER:HB2	1.50	0.77
2:H:11:ARG:NH1	2:H:122:THR:OG1	2.31	0.64
1:R:341:VAL:HG22	1:R:356:LYS:HD2	1.80	0.64
1:R:403:ARG:HH21	1:R:405:ASP:HB2	1.68	0.58
1:R:474:GLN:NE2	1:R:477:SER:O	2.37	0.58

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	R	179/202 (89%)	172 (96%)	7 (4%)	0	100	100
2	H	220/224 (98%)	212 (96%)	8 (4%)	0	100	100
3	L	209/213 (98%)	198 (95%)	11 (5%)	0	100	100
All	All	608/639 (95%)	582 (96%)	26 (4%)	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	R	157/175 (90%)	156 (99%)	1 (1%)	86	90
2	H	191/194 (98%)	190 (100%)	1 (0%)	88	92
3	L	185/187 (99%)	185 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	533/556 (96%)	531 (100%)	2 (0%)	91 93

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	R	337	PRO
2	H	52	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	37	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

1 ligand is 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	$\#\|Z\| > 2$
4	NAG	R	601	1	14,14,15	0.33	0	17,19,21	0.50	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
4	NAG	R	601	1	-	0/6/23/26	0/1/1/1

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	R	601	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^{th} 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	R	181/202 (89%)	2.04	57 (31%) 0 0	64, 104, 184, 230	0
2	H	222/224 (99%)	0.40	6 (2%) 54 63	47, 62, 87, 104	0
3	L	211/213 (99%)	0.32	1 (0%) 91 94	48, 66, 90, 113	0
All	All	614/639 (96%)	0.85	64 (10%) 6 7	47, 70, 148, 230	0

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	R	389	ASP	12.3
1	R	365	TYR	12.1
1	R	368	LEU	11.1
1	R	391	CYS	10.8
1	R	387	LEU	10.3

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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(Å^2)	Q<0.9
4	NAG	R	601	14/15	0.73	0.30	119,123,128,142	0

6.5 Other polymers (i)

There are no such residues in this entry.