



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 12, 2024 – 10:57 am BST

PDB ID : 8QH1
Title : Crystal structure of the SARS-CoV-2 RBD from the Omicron BA4 variant with the antibody Cv2.3194
Authors : Fernandez, I.; Rey, F.A.
Deposited on : 2023-09-06
Resolution : 2.65 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
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36.2
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) : 2.36.2

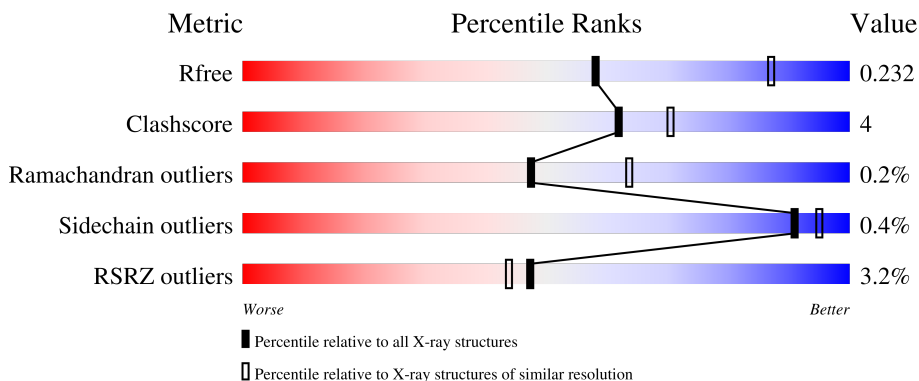
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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	E	243	
2	H	229	
3	L	211	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4753 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	E	192	1526	981	259	278	8	0	1	0

There are 59 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	339	ASP	GLY	conflict	UNP A0A8A5XRG7
E	371	PHE	SER	conflict	UNP A0A8A5XRG7
E	373	PRO	SER	conflict	UNP A0A8A5XRG7
E	375	PHE	SER	conflict	UNP A0A8A5XRG7
E	376	ALA	THR	conflict	UNP A0A8A5XRG7
E	405	ASN	ASP	conflict	UNP A0A8A5XRG7
E	408	SER	ARG	conflict	UNP A0A8A5XRG7
E	440	LYS	ASN	conflict	UNP A0A8A5XRG7
E	452	ARG	LEU	conflict	UNP A0A8A5XRG7
E	477	ASN	SER	conflict	UNP A0A8A5XRG7
E	478	LYS	THR	conflict	UNP A0A8A5XRG7
E	484	ALA	LYS	conflict	UNP A0A8A5XRG7
E	486	VAL	PHE	conflict	UNP A0A8A5XRG7
E	498	ARG	GLN	conflict	UNP A0A8A5XRG7
E	505	HIS	TYR	conflict	UNP A0A8A5XRG7
E	529	GLY	LYS	conflict	UNP A0A8A5XRG7
E	531	GLY	-	expression tag	UNP A0A8A5XRG7
E	532	LEU	-	expression tag	UNP A0A8A5XRG7
E	533	VAL	-	expression tag	UNP A0A8A5XRG7
E	534	PRO	-	expression tag	UNP A0A8A5XRG7
E	535	ARG	-	expression tag	UNP A0A8A5XRG7
E	536	GLY	-	expression tag	UNP A0A8A5XRG7
E	537	SER	-	expression tag	UNP A0A8A5XRG7
E	538	HIS	-	expression tag	UNP A0A8A5XRG7
E	539	HIS	-	expression tag	UNP A0A8A5XRG7
E	540	HIS	-	expression tag	UNP A0A8A5XRG7
E	541	HIS	-	expression tag	UNP A0A8A5XRG7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	542	HIS	-	expression tag	UNP A0A8A5XRG7
E	543	HIS	-	expression tag	UNP A0A8A5XRG7
E	544	HIS	-	expression tag	UNP A0A8A5XRG7
E	545	HIS	-	expression tag	UNP A0A8A5XRG7
E	546	SER	-	expression tag	UNP A0A8A5XRG7
E	547	ALA	-	expression tag	UNP A0A8A5XRG7
E	548	TRP	-	expression tag	UNP A0A8A5XRG7
E	549	SER	-	expression tag	UNP A0A8A5XRG7
E	550	HIS	-	expression tag	UNP A0A8A5XRG7
E	551	PRO	-	expression tag	UNP A0A8A5XRG7
E	552	GLN	-	expression tag	UNP A0A8A5XRG7
E	553	PHE	-	expression tag	UNP A0A8A5XRG7
E	554	GLU	-	expression tag	UNP A0A8A5XRG7
E	555	LYS	-	expression tag	UNP A0A8A5XRG7
E	556	GLY	-	expression tag	UNP A0A8A5XRG7
E	557	THR	-	expression tag	UNP A0A8A5XRG7
E	558	GLY	-	expression tag	UNP A0A8A5XRG7
E	559	GLY	-	expression tag	UNP A0A8A5XRG7
E	560	LEU	-	expression tag	UNP A0A8A5XRG7
E	561	ASN	-	expression tag	UNP A0A8A5XRG7
E	562	ASP	-	expression tag	UNP A0A8A5XRG7
E	563	ILE	-	expression tag	UNP A0A8A5XRG7
E	564	PHE	-	expression tag	UNP A0A8A5XRG7
E	565	GLU	-	expression tag	UNP A0A8A5XRG7
E	566	ALA	-	expression tag	UNP A0A8A5XRG7
E	567	GLN	-	expression tag	UNP A0A8A5XRG7
E	568	LYS	-	expression tag	UNP A0A8A5XRG7
E	569	ILE	-	expression tag	UNP A0A8A5XRG7
E	570	GLU	-	expression tag	UNP A0A8A5XRG7
E	571	TRP	-	expression tag	UNP A0A8A5XRG7
E	572	HIS	-	expression tag	UNP A0A8A5XRG7
E	573	GLU	-	expression tag	UNP A0A8A5XRG7

- Molecule 2 is a protein called Cv2.3194 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	217	1608	1012	270	319	7	0	0	0

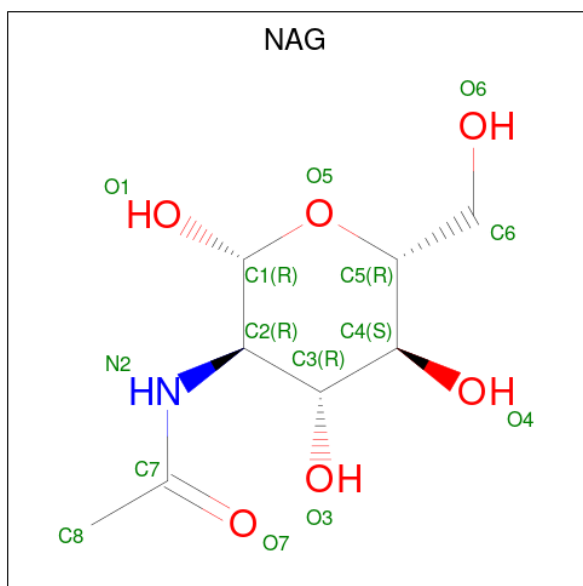
- Molecule 3 is a protein called IGK@ protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	210	1597	998	272	323	4	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	9	GLY	ALA	conflict	UNP Q6PJF2
L	28	SER	ILE	conflict	UNP Q6PJF2
L	32	SER	ALA	conflict	UNP Q6PJF2
L	49	ILE	MET	conflict	UNP Q6PJF2
L	50	TYR	PHE	conflict	UNP Q6PJF2
L	52	ALA	SER	conflict	UNP Q6PJF2
L	61	GLY	ASP	conflict	UNP Q6PJF2
L	86	ILE	VAL	conflict	UNP Q6PJF2
L	?	-	TYR	deletion	UNP Q6PJF2
L	?	-	GLY	deletion	UNP Q6PJF2
L	?	-	SER	deletion	UNP Q6PJF2
L	?	-	SER	deletion	UNP Q6PJF2
L	92	GLY	GLN	conflict	UNP Q6PJF2
L	93	VAL	GLY	conflict	UNP Q6PJF2
L	97	GLY	PRO	conflict	UNP Q6PJF2
L	102	GLU	ASP	conflict	UNP Q6PJF2

- 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	E	1	Total	C	N	O	14	0
			14	8	1	5		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	5	Total	O	0	0
			5	5		
5	L	3	Total	O	0	0
			3	3		

4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	57.79Å 85.35Å 193.99Å 90.00° 97.48° 90.00°	Depositor
Resolution (Å)	39.01 – 2.65 39.01 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.5 (39.01-2.65) 99.6 (39.01-2.65)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.16 (at 2.65Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.198 , 0.232 0.197 , 0.232	Depositor DCC
R_{free} test set	1405 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å ²)	64.4	Xtrriage
Anisotropy	0.423	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 43.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4753	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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

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	E	0.26	0/1573	0.47	0/2140
2	H	0.25	0/1643	0.50	0/2241
3	L	0.26	0/1630	0.51	0/2211
All	All	0.25	0/4846	0.49	0/6592

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	E	1526	0	1456	9	0
2	H	1608	0	1583	19	1
3	L	1597	0	1561	14	0
4	E	14	0	13	0	0
5	H	5	0	0	0	0
5	L	3	0	0	0	0
All	All	4753	0	4613	41	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 4.

The worst 5 of 41 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:90:THR:HG22	2:H:115:VAL:H	1.49	0.77
2:H:29:VAL:HG13	2:H:34:MET:HG3	1.72	0.70
1:E:487:ASN:OD1	2:H:97:ARG:NH1	2.25	0.67
1:E:393:THR:HG21	1:E:518:LEU:HB2	1.80	0.62
3:L:29:VAL:HG11	3:L:91:GLN:HG3	1.81	0.62

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:H:139:THR:OG1	2:H:139:THR:OG1[2_656]	2.15	0.05

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	E	189/243 (78%)	178 (94%)	11 (6%)	0	100	100
2	H	213/229 (93%)	208 (98%)	4 (2%)	1 (0%)	29	43
3	L	208/211 (99%)	201 (97%)	7 (3%)	0	100	100
All	All	610/683 (89%)	587 (96%)	22 (4%)	1 (0%)	47	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	100	VAL

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	E	165/207 (80%)	164 (99%)	1 (1%)	86	92
2	H	182/194 (94%)	181 (100%)	1 (0%)	88	94
3	L	179/180 (99%)	179 (100%)	0	100	100
All	All	526/581 (90%)	524 (100%)	2 (0%)	91	95

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

Mol	Chain	Res	Type
1	E	334	ASN
2	H	3	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

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	E	601	1	14,14,15	0.23	0	17,19,21	0.43	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	E	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.

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 [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	E	192/243 (79%)	0.44	17 (8%) 9 8	50, 82, 152, 176	0
2	H	217/229 (94%)	-0.07	3 (1%) 75 73	53, 74, 108, 150	0
3	L	210/211 (99%)	-0.19	0 100 100	52, 70, 99, 135	0
All	All	619/683 (90%)	0.05	20 (3%) 47 44	50, 74, 133, 176	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	527	PRO	5.8
1	E	519	HIS	5.6
1	E	372	ALA	5.3
1	E	375	PHE	5.2
1	E	333	THR	5.2

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, 95th 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(\AA^2)	Q<0.9
4	NAG	E	601	14/15	-	-	111,121,122,122	14

6.5 Other polymers [i](#)

There are no such residues in this entry.