wwPDB X-ray Structure Validation Summary Report (i)

May 30, 2022 – 04:05 pm BST

PDB ID	:	7ZF4
Title	:	SARS-CoV-2 Omicron RBD in complex with Omi-9 Fab and nanobody F2
Authors	:	Zhou, D.; Huo, J.; Ren, J.; Stuart, D.I.
Deposited on	:	2022-04-01
Resolution	:	4.18 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 following versions of software and data (see references (1)) were used in the production of this report:

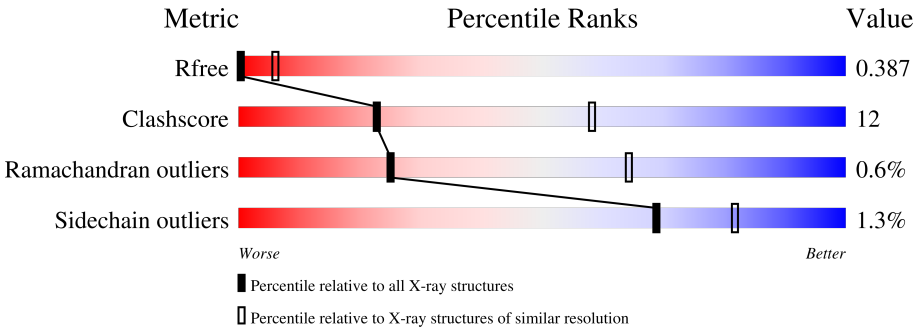
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 4.18 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	1034 (4.60-3.76)
Clashscore	141614	1030 (4.54-3.80)
Ramachandran outliers	138981	1006 (4.58-3.78)
Sidechain outliers	138945	1037 (4.60-3.76)

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%

Mol	Chain	Length	Quality of chain
1	F	132	82% 12% 6%
2	H	227	76% 21% .
3	L	217	87% 8% 5%
4	E	202	77% 18% 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5716 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 Nanobody F2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	124	Total	C	N	O	S	0	0	0
			981	621	170	186	4			

- Molecule 2 is a protein called Omi-9 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	219	Total	C	N	O	S	0	3	0
			1623	1024	275	318	6			

- Molecule 3 is a protein called Omi-9 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	207	Total	C	N	O	S	0	2	0
			1569	982	256	327	4			

- Molecule 4 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	192	Total	C	N	O	S	0	0	0
			1543	993	261	281	8			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	327	HIS	-	expression tag	UNP P0DTC2
E	328	HIS	-	expression tag	UNP P0DTC2
E	329	HIS	-	expression tag	UNP P0DTC2
E	330	HIS	-	expression tag	UNP P0DTC2
E	331	HIS	-	expression tag	UNP P0DTC2
E	332	HIS	-	expression tag	UNP P0DTC2
E	339	ASP	GLY	variant	UNP P0DTC2
E	371	LEU	SER	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	373	PRO	SER	variant	UNP P0DTC2
E	375	PHE	SER	variant	UNP P0DTC2
E	417	ASN	LYS	variant	UNP P0DTC2
E	440	LYS	ASN	variant	UNP P0DTC2
E	446	SER	GLY	variant	UNP P0DTC2
E	477	ASN	SER	variant	UNP P0DTC2
E	478	LYS	THR	variant	UNP P0DTC2
E	484	ALA	GLU	variant	UNP P0DTC2
E	493	ARG	GLN	variant	UNP P0DTC2
E	496	SER	GLY	variant	UNP P0DTC2
E	498	ARG	GLN	conflict	UNP P0DTC2
E	501	TYR	ASN	variant	UNP P0DTC2
E	505	HIS	TYR	variant	UNP P0DTC2
E	527	LYS	-	expression tag	UNP P0DTC2
E	528	LYS	-	expression tag	UNP P0DTC2

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. 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. 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: Nanobody F2

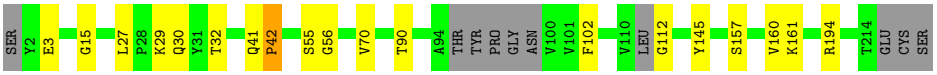
Chain F:

- Molecule 2: Omi-9 heavy chain

Chain H:

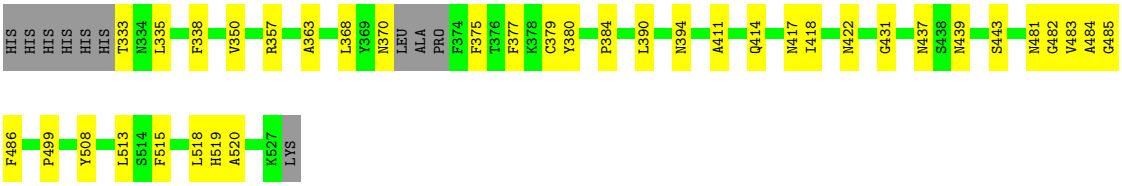
- Molecule 3: Omi-9 light chain

Chain L:



- Molecule 4: Spike protein S1

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	86.57Å 205.15Å 123.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	61.53 – 4.18 66.93 – 4.18	Depositor EDS
% Data completeness (in resolution range)	96.7 (61.53-4.18) 96.7 (66.93-4.18)	Depositor EDS
R_{merge}	0.73	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 4.14Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.369 , 0.385 0.367 , 0.387	Depositor DCC
R_{free} test set	417 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	150.7	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.26$, $< L^2 > = 0.12$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	5716	wwPDB-VP
Average B, all atoms (Å^2)	195.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.84% 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

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	F	0.67	0/1006	0.86	1/1363 (0.1%)
2	H	0.40	0/1669	0.69	3/2276 (0.1%)
3	L	0.45	0/1607	0.70	1/2199 (0.0%)
4	E	0.26	0/1587	0.48	0/2157
All	All	0.44	0/5869	0.68	5/7995 (0.1%)

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
1	F	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	122	SER	O-C-N	11.03	140.35	122.70
2	H	122	SER	CA-C-N	-8.19	99.19	117.20
3	L	42	PRO	CA-N-CD	-8.13	100.12	111.50
2	H	122	SER	C-N-CA	-8.11	101.41	121.70
1	F	100	ARG	NE-CZ-NH2	-5.08	117.76	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	55	SER	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	F	981	0	935	19	5
2	H	1623	0	1585	88	4
3	L	1569	0	1504	26	15
4	E	1543	0	1465	49	5
All	All	5716	0	5489	129	21

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 12.

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:H:34:VAL:CG1	2:H:79:LEU:HD22	1.37	1.54
2:H:49:ALA:CB	2:H:70:LEU:HD11	1.41	1.50
2:H:45:PRO:CD	3:L:102:PHE:CB	1.82	1.40
2:H:108:ALA:CB	4:E:485:GLY:HA2	1.47	1.40
2:H:49:ALA:CB	2:H:70:LEU:CD1	1.99	1.39

The worst 5 of 21 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:3:GLN:NE2	3:L:15:GLY:CA[3_654]	1.38	0.82
3:L:55:SER:O	3:L:55:SER:O[3_654]	1.38	0.82
3:L:3:GLU:N	3:L:3:GLU:OE2[4_554]	1.50	0.70
3:L:3:GLU:CB	3:L:3:GLU:OE1[4_554]	1.51	0.69
3:L:56:GLY:O	3:L:56:GLY:O[3_654]	1.52	0.68

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	F	122/132 (92%)	117 (96%)	5 (4%)	0	100 100
2	H	218/227 (96%)	208 (95%)	8 (4%)	2 (1%)	17 55
3	L	203/217 (94%)	195 (96%)	8 (4%)	0	100 100
4	E	188/202 (93%)	175 (93%)	11 (6%)	2 (1%)	14 51
All	All	731/778 (94%)	695 (95%)	32 (4%)	4 (0%)	25 68

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	E	519	HIS
4	E	520	ALA
2	H	28	THR
2	H	98	SER

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	F	101/109 (93%)	100 (99%)	1 (1%)	76 86
2	H	180/190 (95%)	179 (99%)	1 (1%)	86 92
3	L	179/187 (96%)	178 (99%)	1 (1%)	86 92
4	E	167/177 (94%)	162 (97%)	5 (3%)	41 63
All	All	627/663 (95%)	619 (99%)	8 (1%)	69 82

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	518	LEU
4	E	513	LEU
4	E	335	LEU
4	E	333	THR
4	E	368	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	39	GLN
4	E	370	ASN

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.