W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2022 – 04:13 pm BST

PDB ID	:	7ZFD
Title	:	SARS-CoV-2 Omicron RBD in complex with Omi-25 Fab
Authors	:	Zhou, D.; Huo, J.; Ren, J.; Stuart, D.I.
Deposited on	:	2022-04-01
Resolution	:	3.39 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

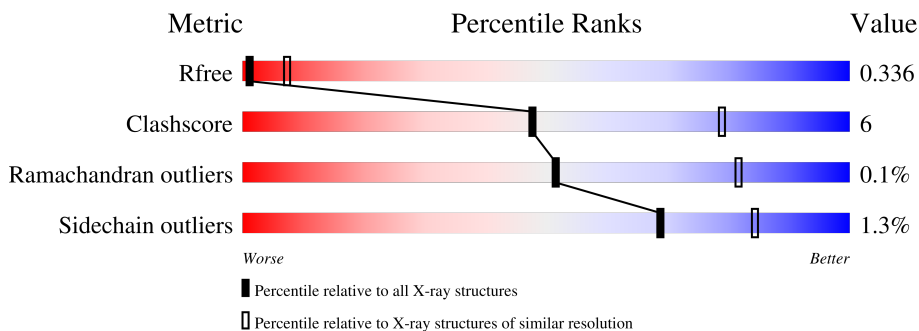|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.39 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	Percentile Ranks	Value
Rfree		0.336
Clashscore		6
Ramachandran outliers		0.1%
Sidechain outliers		1.3%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)

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	A	228	81% 16% ··
1	H	228	78% 19% ··
2	E	209	78% 11% 11%
2	F	209	77% 9% · 13%
3	B	214	90% 9%
3	L	214	85% 14%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9597 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 Omi-25 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	224	Total 1681	C 1062	N 286	O 326	S 7	0	1	0
1	A	224	Total 1679	C 1062	N 284	O 326	S 7	0	1	0

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	185	Total 1496	C 965	N 251	O 272	S 8	0	0	0
2	F	182	Total 1465	C 940	N 248	O 269	S 8	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	339	ASP	GLY	variant	UNP P0DTC2
E	371	LEU	SER	variant	UNP P0DTC2
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	533	HIS	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	534	HIS	-	expression tag	UNP P0DTC2
E	535	HIS	-	expression tag	UNP P0DTC2
E	536	HIS	-	expression tag	UNP P0DTC2
E	537	HIS	-	expression tag	UNP P0DTC2
E	538	HIS	-	expression tag	UNP P0DTC2
F	339	ASP	GLY	variant	UNP P0DTC2
F	371	LEU	SER	variant	UNP P0DTC2
F	373	PRO	SER	variant	UNP P0DTC2
F	375	PHE	SER	variant	UNP P0DTC2
F	417	ASN	LYS	variant	UNP P0DTC2
F	440	LYS	ASN	variant	UNP P0DTC2
F	446	SER	GLY	variant	UNP P0DTC2
F	477	ASN	SER	variant	UNP P0DTC2
F	478	LYS	THR	variant	UNP P0DTC2
F	484	ALA	GLU	variant	UNP P0DTC2
F	493	ARG	GLN	variant	UNP P0DTC2
F	496	SER	GLY	variant	UNP P0DTC2
F	498	ARG	GLN	conflict	UNP P0DTC2
F	501	TYR	ASN	variant	UNP P0DTC2
F	505	HIS	TYR	variant	UNP P0DTC2
F	533	HIS	-	expression tag	UNP P0DTC2
F	534	HIS	-	expression tag	UNP P0DTC2
F	535	HIS	-	expression tag	UNP P0DTC2
F	536	HIS	-	expression tag	UNP P0DTC2
F	537	HIS	-	expression tag	UNP P0DTC2
F	538	HIS	-	expression tag	UNP P0DTC2

- Molecule 3 is a protein called Omi-25 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	213	Total 1638	C 1027	N 273	O 333	S 5	0	0	0
3	B	213	Total 1638	C 1027	N 273	O 333	S 5	0	0	0

# 3 Residue-property plots ⓘ

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.
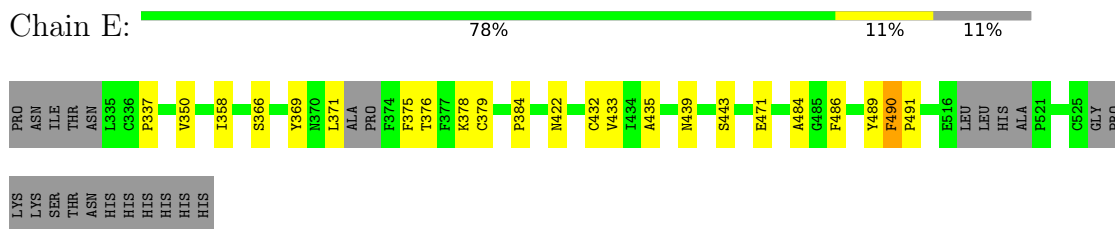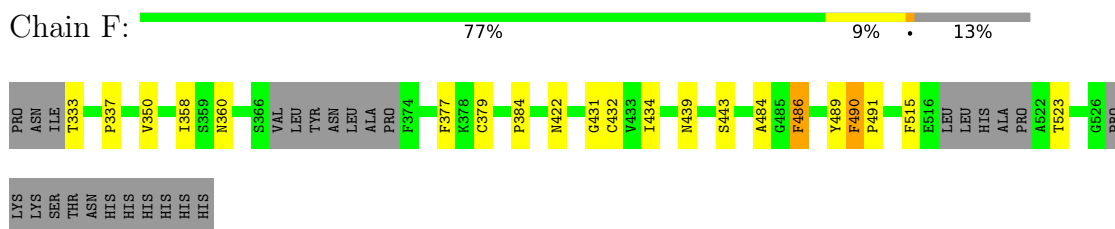
- Molecule 1: Omi-25 heavy chain
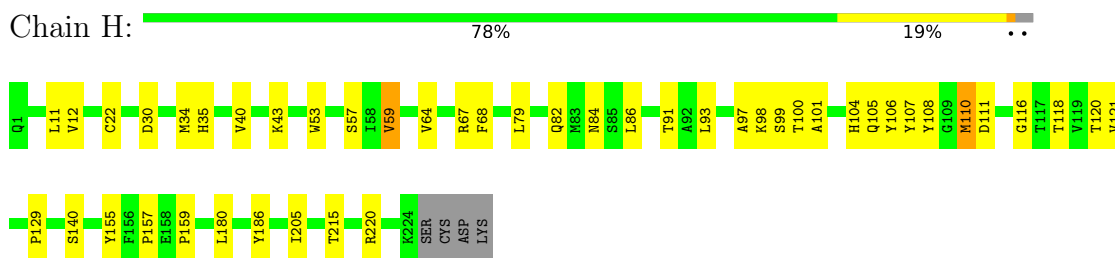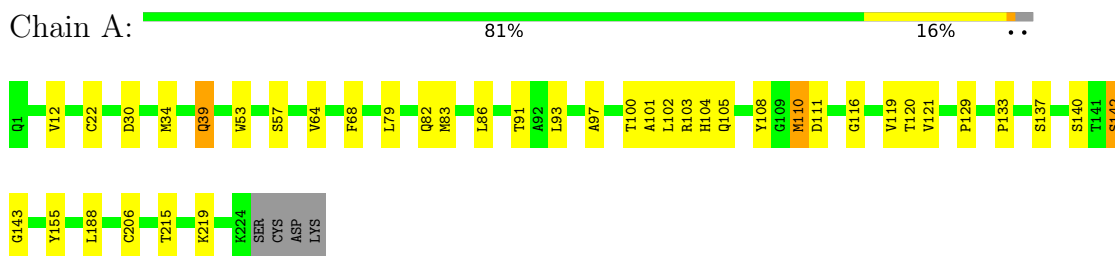
Chain H:

- Molecule 1: Omi-25 heavy chain

Chain A:

- Molecule 2: Spike protein S1

Chain E:

- Molecule 2: Spike protein S1

Chain F:

• Molecule 3: Omi-25 light chain

Chain L:

85%   14%

A1
L11
S12
A13
S31
W35
Y36
Q37
Q38
K39
K42
L47
Y67
L73
A84
C88
Q89
Q90
S91
T94
P95
Y96
G99
V104
E105
I106
K107
P120
V132
P141
R142
V163
Q166
Y173
S174
L175
A184
K188
H198

E213
CYS

• Molecule 3: Omi-25 light chain

Chain B:

90%   9%

A1
S31
W35
Y36
Q37
Q38
K39
K42
L47
Y67
L73
A84
C88
Q89
T94
P95
Y96
A97
F98
G99
P120
V132
Y186
Y192
E213
CYS
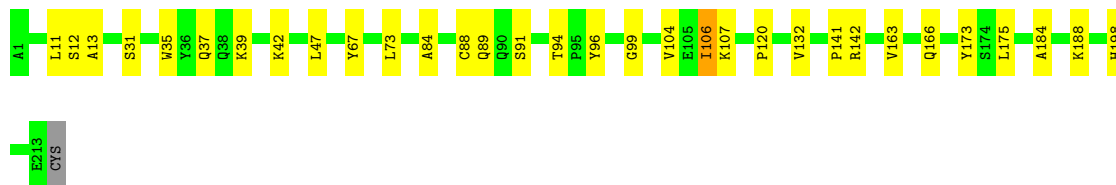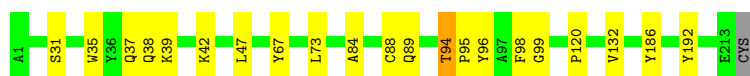
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	123.05Å 123.05Å 223.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.67 – 3.39 68.67 – 3.39	Depositor EDS
% Data completeness (in resolution range)	99.7 (68.67-3.39) 99.7 (68.67-3.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.00 (at 3.41Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.286 , 0.338 0.284 , 0.336	Depositor DCC
R_{free} test set	1216 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	85.8	Xtriage
Anisotropy	0.400	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.32, < L^2 > = 0.16$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.77	EDS
Total number of atoms	9597	wwPDB-VP
Average B, all atoms (Å^2)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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	A	0.26	0/1724	0.51	0/2351
1	H	0.27	0/1726	0.52	0/2354
2	E	0.25	0/1538	0.47	0/2087
2	F	0.26	0/1505	0.48	0/2041
3	B	0.25	0/1674	0.47	0/2274
3	L	0.25	0/1674	0.47	0/2274
All	All	0.26	0/9841	0.49	0/13381

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	A	1679	0	1629	26	0
1	H	1681	0	1629	33	0
2	E	1496	0	1424	15	0
2	F	1465	0	1386	11	0
3	B	1638	0	1588	15	0
3	L	1638	0	1588	22	0
All	All	9597	0	9244	114	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 6.

The worst 5 of 114 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)
1:A:104:HIS:ND1	2:F:484:ALA:O	2.17	0.78
1:H:205:ILE:HG12	1:H:220:ARG:HG2	1.68	0.75
1:A:64:VAL:HG11	1:A:68:PHE:HB2	1.68	0.74
1:H:59:VAL:HG11	3:L:94:THR:HG21	1.70	0.74
1:H:104:HIS:ND1	2:E:484:ALA:O	2.22	0.73

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	223/228 (98%)	210 (94%)	12 (5%)	1 (0%)	34	67
1	H	223/228 (98%)	212 (95%)	11 (5%)	0	100	100
2	E	179/209 (86%)	167 (93%)	12 (7%)	0	100	100
2	F	176/209 (84%)	163 (93%)	13 (7%)	0	100	100
3	B	211/214 (99%)	204 (97%)	7 (3%)	0	100	100
3	L	211/214 (99%)	204 (97%)	7 (3%)	0	100	100
All	All	1223/1302 (94%)	1160 (95%)	62 (5%)	1 (0%)	51	82

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	142	SER

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	185/191 (97%)	182 (98%)	3 (2%)	62	81
1	H	185/191 (97%)	182 (98%)	3 (2%)	62	81
2	E	163/184 (89%)	161 (99%)	2 (1%)	71	85
2	F	159/184 (86%)	155 (98%)	4 (2%)	47	72
3	B	187/189 (99%)	186 (100%)	1 (0%)	88	94
3	L	187/189 (99%)	186 (100%)	1 (0%)	88	94
All	All	1066/1128 (94%)	1052 (99%)	14 (1%)	69	84

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110	MET
1	A	188	LEU
3	B	94	THR
2	F	489	TYR
2	F	490	PHE

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

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 ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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.