wwPDB EM Validation Summary Report

Oct 16, 2024 – 12:53 AM JST

PDB ID	:	8XSE
EMDB ID	:	EMD-38617
Title	:	SARS-CoV-2 RBD + IMCAS-123 + IMCAS-72 Fab
Authors	:	Tong, Z.; Cui, Y.; Xie, Y.; Tong, J.; Gao, G.F.; Qi, J.
Deposited on	:	2024-01-09
Resolution	:	2.50 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
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:

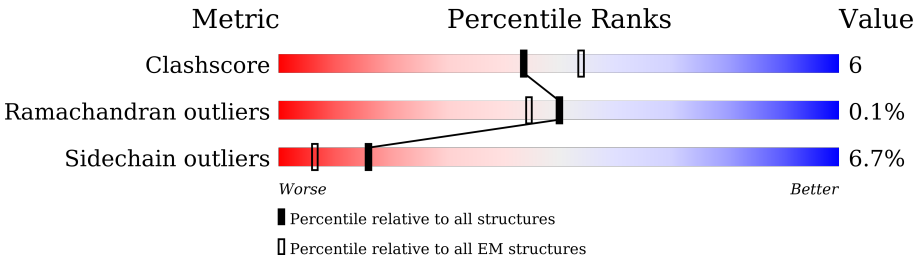
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 2.50 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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	209	80% 12% 7%
2	H	226	77% 20% ••
3	L	215	78% 18% ••
4	M	226	76% 22% ••
5	N	215	82% 15% ••

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8120 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 protein S1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	195	1543	989	257	289	8	0	0

- Molecule 2 is a protein called IMCAS-123 H chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	H	223	1658	1044	281	326	7	0	0

- Molecule 3 is a protein called IMCAS-123 L chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	L	212	1611	1008	272	326	5	0	0

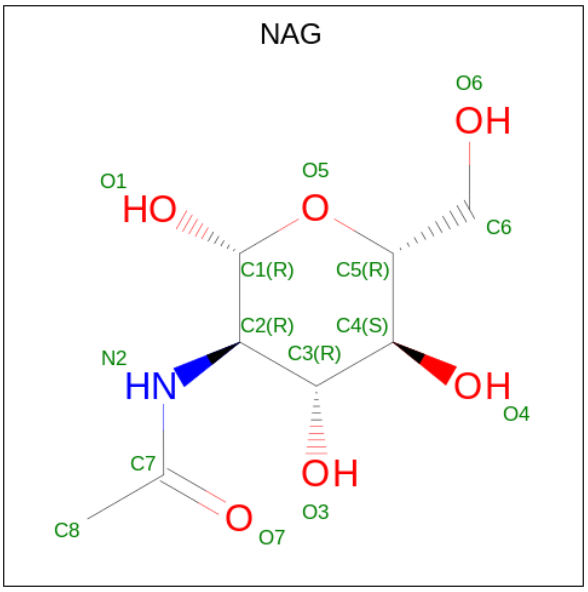
- Molecule 4 is a protein called IMCAS-72 H chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	M	223	1673	1053	287	327	6	0	0

- Molecule 5 is a protein called IMCAS-72 L chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	N	212	1621	1015	270	331	5	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



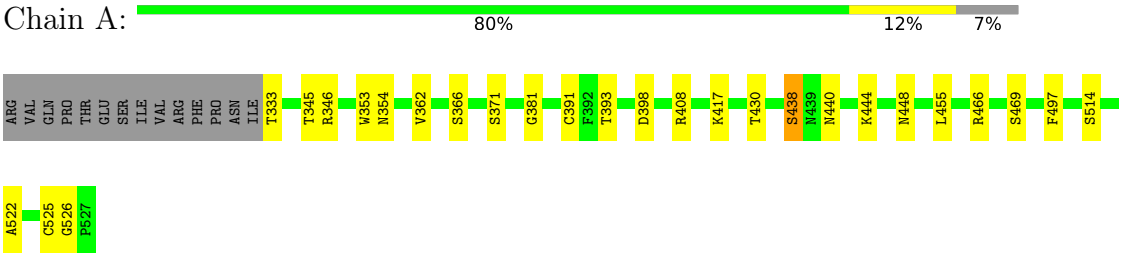
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
6	A	1	14	8	1	5	0

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.

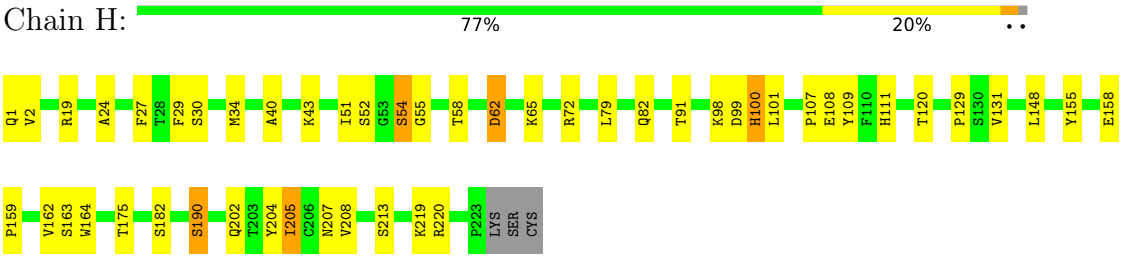
- Molecule 1: Spike protein S1

Chain A:



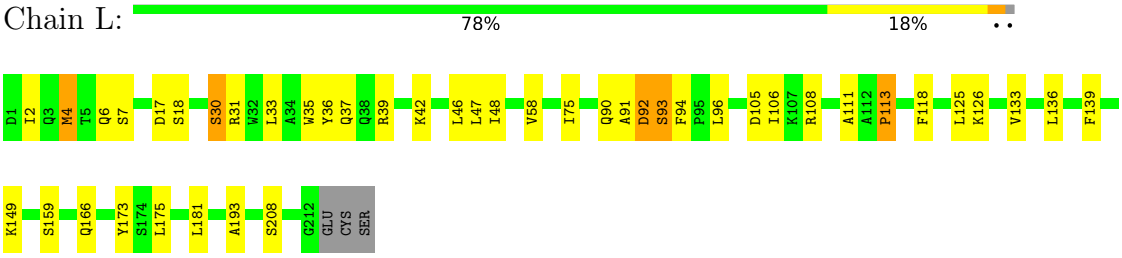
- Molecule 2: IMCAS-123 H chain

Chain H:



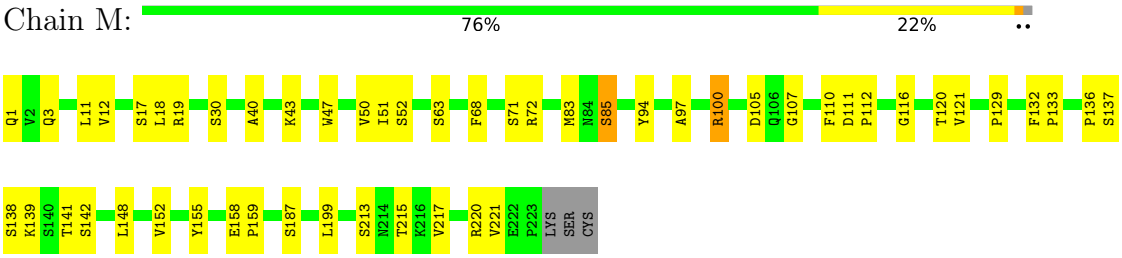
- Molecule 3: IMCAS-123 L chain

Chain L:



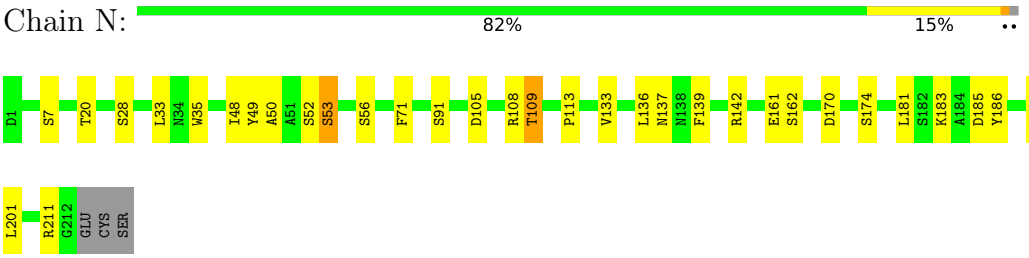
- Molecule 4: IMCAS-72 H chain

Chain M:



• Molecule 5: IMCAS-72 L chain

Chain N:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	583803	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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	A	0.27	0/1587	0.47	0/2161
2	H	0.26	0/1699	0.52	0/2314
3	L	0.28	0/1645	0.56	2/2234 (0.1%)
4	M	0.26	0/1715	0.51	0/2334
5	N	0.26	0/1657	0.50	0/2251
All	All	0.27	0/8303	0.51	2/11294 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	L	113	PRO	CA-N-CD	-8.54	99.55	111.50
3	L	113	PRO	N-CD-CG	-5.11	95.54	103.20

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	1543	0	1459	15	0
2	H	1658	0	1617	31	0
3	L	1611	0	1563	22	0
4	M	1673	0	1620	23	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	N	1621	0	1576	14	0
6	A	14	0	13	0	0
All	All	8120	0	7848	100	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 100 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)
4:M:213:SER:HG	4:M:215:THR:HG1	1.34	0.76
4:M:111:ASP:HB3	4:M:112:PRO:HD3	1.75	0.68
1:A:346:ARG:NH1	2:H:108:GLU:OE1	2.28	0.67
2:H:19:ARG:HG3	2:H:82:GLN:HG2	1.75	0.67
2:H:131:VAL:O	2:H:219:LYS:NZ	2.31	0.63

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 PDB entries followed by that with respect to all EM entries.

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	193/209 (92%)	186 (96%)	7 (4%)	0	100	100
2	H	221/226 (98%)	205 (93%)	15 (7%)	1 (0%)	25	44
3	L	210/215 (98%)	199 (95%)	11 (5%)	0	100	100
4	M	221/226 (98%)	205 (93%)	16 (7%)	0	100	100
5	N	210/215 (98%)	200 (95%)	10 (5%)	0	100	100
All	All	1055/1091 (97%)	995 (94%)	59 (6%)	1 (0%)	50	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	54	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 PDB entries followed by that with respect to all EM entries.

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	168/182 (92%)	162 (96%)	6 (4%)	30 56
2	H	186/189 (98%)	174 (94%)	12 (6%)	14 29
3	L	180/184 (98%)	168 (93%)	12 (7%)	13 28
4	M	184/188 (98%)	168 (91%)	16 (9%)	8 17
5	N	187/190 (98%)	172 (92%)	15 (8%)	10 20
All	All	905/933 (97%)	844 (93%)	61 (7%)	16 28

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	L	159	SER
5	N	161	GLU
4	M	63	SER
5	N	142	ARG
5	N	191	VAL

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 oligosaccharides 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	$\#\|Z\| > 2$
6	NAG	A	601	1	14,14,15	0.21	0	17,19,21	0.40	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
6	NAG	A	601	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	601	NAG	O5-C5-C6-O6
6	A	601	NAG	C4-C5-C6-O6

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.