wwPDB X-ray Structure Validation Summary Report (i)

Nov 20, 2023 – 11:34 AM JST

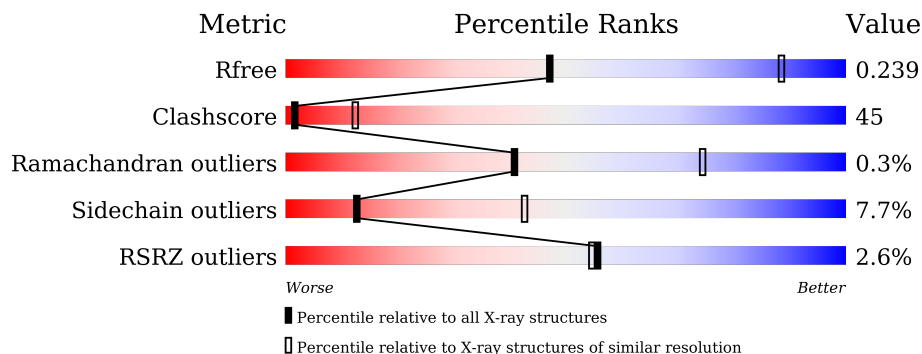|		
PDB ID	:	7CHE
Title	:	Crystal structure of the SARS-CoV-2 RBD in complex with BD-236 Fab and BD-368-2 Fab
Authors	:	Xiao, J.; Zhu, Q.
Deposited on	:	2020-07-05
Resolution	:	3.42 Å(reported)

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 3.42 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		0.239
Clashscore		45
Ramachandran outliers		0.3%
Sidechain outliers		7.7%
RSRZ outliers		2.6%

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	1486 (3.50-3.34)
Clashscore	141614	1572 (3.50-3.34)
Ramachandran outliers	138981	1534 (3.50-3.34)
Sidechain outliers	138945	1535 (3.50-3.34)
RSRZ outliers	127900	1395 (3.50-3.34)

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	H	223	.% 43% 44% 8% 5%
2	L	214	51% 40% 7% •
3	R	223	7% 38% 39% 7% 15%
4	A	230	41% 42% 11% • 5%
5	B	219	4% 42% 49% 7% •

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 7956 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 BD-236 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	212	Total 1559	C 977	N 262	O 313	S 7	0	0	0

- Molecule 2 is a protein called BD-236 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	211	Total 1604	C 1007	N 269	O 324	S 4	0	0	0

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	189	Total 1498	C 958	N 249	O 283	S 8	0	0	0

- Molecule 4 is a protein called BD-368-2 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	219	Total 1636	C 1034	N 278	O 318	S 6	0	0	0

- Molecule 5 is a protein called BD-368-2 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	216	Total 1659	C 1038	N 281	O 334	S 6	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 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: BD-236 Fab heavy chain

Chain H:

• Molecule 2: BD-236 Fab light chain

Chain L:

• Molecule 3: Spike protein S1

Chain R:

N388 D389 L390 C391 F392 T393 N394 V395 Y396 F400 V401 I402 R403 G404 D405 E406 V407 R408 Q409 I410 Q414 T415 G416 K417 I418 A419 D420 Y421 N422 Y423 K424 L425 P426 D427 T430 G431 C432 V433 I434 N437 S438 D442 S443 K444 G447 N448 Y451 L452 Y453 R454 L455 F456 R457

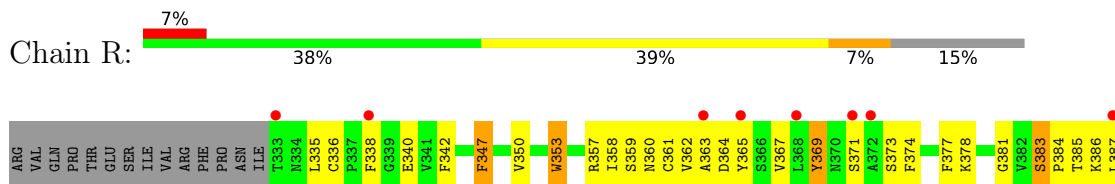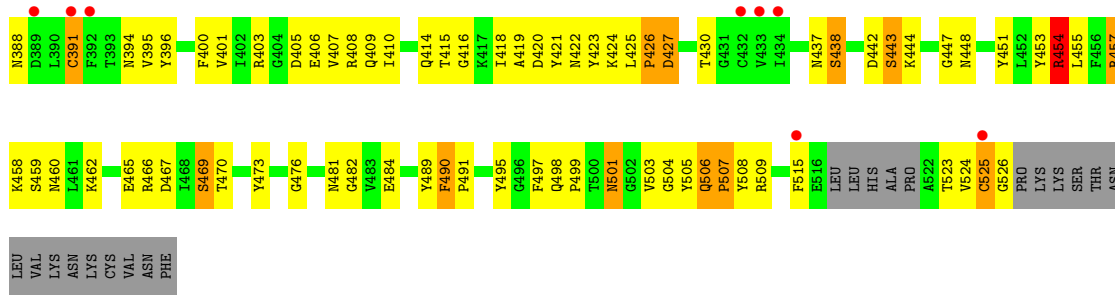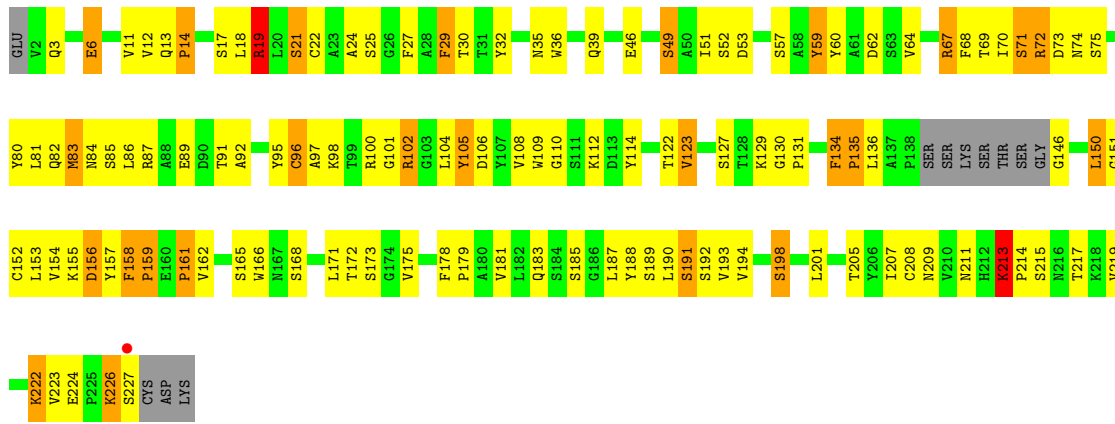K458 S459 N460 L461 K462 E465 R466 D467 I468 S469 T470 Y473 G476 N481 G482 V483 E484 Y489 F490 P491 Y495 G496 F497 Q498 P499 T500 N501 G502 V503 G504 Y505 Q506 P507 Y508 R509 F515 E516 LEU LEU HIS ALA PRO A522 T523 V524 C525 G526 PRO LYS LYS SER THR ASN

LEU VAL LYS ASN LYS CYS VAL ASN PHE

• Molecule 4: BD-368-2 Fab heavy chain

Chain A:

41%   42%   11%   • 5%

GLU V2 Q3 E6 V11 V12 Q13 P14 S17 L18 R19 L20 S21 C22 A23 A24 S25 G26 F27 A28 F29 T30 T31 Y32 N35 W36 Q39 E46 S49 A50 I51 S52 D53 S57 A58 Y59 Y60 A61 D62 S63 V64 R67 F68 T69 I70 S71 R72 D73 N74 S75

Y80 L81 Q82 M83 N84 S85 L86 R87 A88 E89 D90 T91 A92 Y95 C96 A97 K98 T99 R100 G101 R102 G103 L104 Y105 D106 Y107 V108 W109 G110 S111 K112 Y114 D113 T122 V123 S127 T128 K129 G130 P131 F134 P135 L136 A137 P138 SER SER LYS SER THR SER GLY G146 L150 G151

C152 L153 V154 K155 D156 Y157 F158 P159 E160 P161 V162 S165 W166 N167 S168 L171 T172 S173 G174 V175 F178 P179 A180 V181 L182 Q183 S184 S185 G186 L187 Y188 S189 L190 S191 V192 V193 V194 S198 L201 T205 Y206 I207 C208 N209 V210 N211 H212 K213 P214 S215 N216 T217 K218 V219

K222 V223 E224 P225 K226 S227 CYS ASP LYS

• Molecule 5: BD-368-2 Fab light chain

4%

Chain B:

42%   49%   7%   •

D1 I2 V3 M4 T5 Q6 S7 P8 L9 S10 L11 P12 V13 T14 E17 P18 A19 S20 I21 S22 C23 R24 S25 S26 Q27 S28 L29 L30 H31 S32 N33 Y37 L38 D39 Q43 K44 P45 L52 I53 Y54 N58 V63 P64 D65 R66 F67 S68 S72 G73 T74 D75 F76 T77

L78 K79 I80 S81 R82 V83 E84 A85 E86 D87 V88 Y92 C93 M94 Q95 A96 L97 Q98 T99 P100 G101 T102 F103 G104 T107 R108 L109 K112 R113 P118 F121 I122 F123 P124 P125 S126 D127 E128 Q129 L130 K131 S132 S136 L140 L141 F144 Y145 P146 K150 V151 Q152

W153 K154 V155 D156 N157 A158 L159 Q160 N163 S164 Q165 E166 S167 V168 T169 D172 D175 S176 T177 L180 S181 S182 T183 L184 K188 Y191 E192 K193 H194 K195 V196 Y197 A198 V201 T202 L206 S207 V210 T211 K212 S213 F214 N215 R216 GLY GLU CYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.02Å 114.30Å 116.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.62 – 3.42 49.62 – 3.42	Depositor EDS
% Data completeness (in resolution range)	94.6 (49.62-3.42) 94.6 (49.62-3.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.225 , 0.239 0.225 , 0.239	Depositor DCC
R_{free} test set	1771 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å^2)	64.1	Xtriage
Anisotropy	0.172	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 23.5	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.026 for -h,l,k	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	7956	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ 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	H	1.37	11/1591 (0.7%)	1.36	15/2166 (0.7%)
2	L	1.53	12/1639 (0.7%)	1.38	21/2227 (0.9%)
3	R	1.39	13/1538 (0.8%)	1.32	12/2090 (0.6%)
4	A	1.65	17/1675 (1.0%)	1.42	20/2280 (0.9%)
5	B	1.61	14/1695 (0.8%)	1.36	15/2303 (0.7%)
All	All	1.52	67/8138 (0.8%)	1.37	83/11066 (0.8%)

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
2	L	0	1
3	R	0	1
4	A	0	1
5	B	0	1
All	All	0	4

The worst 5 of 67 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	64	PRO	N-CD	-31.65	1.03	1.47
4	A	159	PRO	N-CD	-28.88	1.07	1.47
2	L	113	PRO	N-CD	-27.76	1.08	1.47
4	A	161	PRO	N-CD	-20.10	1.19	1.47
5	B	146	PRO	N-CA	13.56	1.70	1.47

The worst 5 of 83 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	39	ASP	CB-CA-C	-13.10	84.21	110.40
1	H	154	PRO	CA-N-CD	11.89	128.34	111.70

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	153	GLU	C-N-CD	-10.30	97.94	120.60
5	B	12	PRO	CA-N-CD	-10.29	97.10	111.50
2	L	91	LEU	CB-CA-C	9.71	128.66	110.20

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	62	ASP	Mainchain
5	B	32	SER	Mainchain
2	L	140	TYR	Peptide
3	R	501	ASN	Mainchain

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	H	1559	0	1534	158	0
2	L	1604	0	1569	125	1
3	R	1498	0	1408	138	0
4	A	1636	0	1597	172	1
5	B	1659	0	1616	160	0
All	All	7956	0	7724	703	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 45.

The worst 5 of 703 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:A:68:PHE:CD1	4:A:83:MET:HB2	1.44	1.51
4:A:68:PHE:CE1	4:A:83:MET:HB2	1.43	1.51
5:B:153:TRP:CG	5:B:184:LEU:HD11	1.48	1.48
1:H:207:PRO:CA	1:H:207:PRO:N	1.69	1.43
3:R:507:PRO:N	3:R:507:PRO:CA	1.67	1.43

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:L:27:GLN:O	4:A:227:SER:OG[3_645]	1.89	0.31

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	H	208/223 (93%)	199 (96%)	9 (4%)	0	100	100
2	L	209/214 (98%)	198 (95%)	11 (5%)	0	100	100
3	R	185/223 (83%)	167 (90%)	18 (10%)	0	100	100
4	A	215/230 (94%)	198 (92%)	16 (7%)	1 (0%)	29	65
5	B	214/219 (98%)	198 (92%)	14 (6%)	2 (1%)	17	53
All	All	1031/1109 (93%)	960 (93%)	68 (7%)	3 (0%)	41	74

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	B	33	ASN
5	B	31	HIS
4	A	156	ASP

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	H	175/185 (95%)	161 (92%)	14 (8%)	12	41
2	L	182/184 (99%)	168 (92%)	14 (8%)	13	42
3	R	163/196 (83%)	153 (94%)	10 (6%)	18	51
4	A	178/188 (95%)	158 (89%)	20 (11%)	6	26
5	B	190/192 (99%)	180 (95%)	10 (5%)	22	55
All	All	888/945 (94%)	820 (92%)	68 (8%)	13	42

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	222	LYS
5	B	26	SER
5	B	193	LYS
2	L	183	LYS
2	L	182	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	77	ASN
5	B	165	GLN
5	B	203	HIS
4	A	216	ASN
2	L	199	GLN

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 ⓘ

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(Å²)	Q<0.9
1	H	212/223 (95%)	0.26	2 (0%)	84	83	26, 56, 80, 90	0
2	L	211/214 (98%)	0.03	0	100	100	26, 47, 72, 81	0
3	R	189/223 (84%)	0.43	16 (8%)	10	13	20, 50, 107, 123	0
4	A	219/230 (95%)	0.09	1 (0%)	91	90	25, 56, 85, 99	0
5	B	216/219 (98%)	0.36	8 (3%)	41	41	20, 52, 95, 135	0
All	All	1047/1109 (94%)	0.23	27 (2%)	56	54	20, 52, 91, 135	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	333	THR	5.6
4	A	227	SER	5.1
3	R	525	CYS	4.5
3	R	365	TYR	4.1
3	R	391	CYS	4.0

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.