W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Mar 31, 2025 – 10:12 PM JST

PDB ID	:	9LUG / pdb_00009lug
Title	:	Crystal structure of SARS-Cov-2 main protease E166V mutant in complex with Bofutrelvir
Authors	:	Zhou, X.L.; Li, J.
Deposited on	:	2025-02-08
Resolution	:	2.13 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

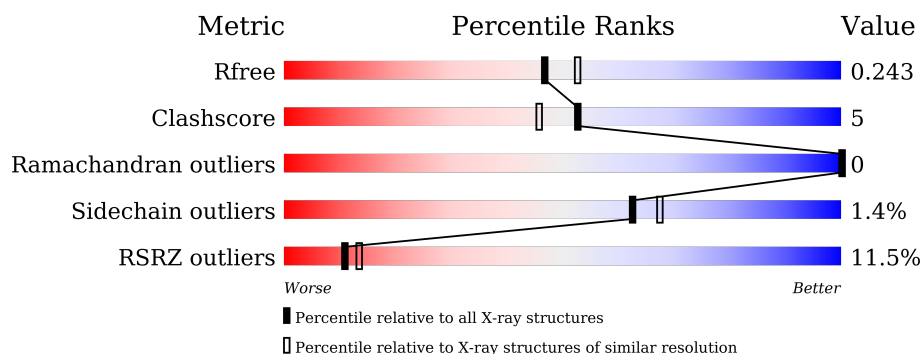|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 2.13 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		0.243
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		1.4%
RSRZ outliers		11.5%

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}	164625	3336 (2.16-2.12)
Clashscore	180529	3585 (2.16-2.12)
Ramachandran outliers	177936	3554 (2.16-2.12)
Sidechain outliers	177891	3553 (2.16-2.12)
RSRZ outliers	164620	3337 (2.16-2.12)

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	A	302	17% 88% 11%
1	B	302	6% 89% 10% ·

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5028 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 ORF1a polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	0	0
			2334	1478	399	435	22			
1	B	299	Total	C	N	O	S	0	0	0
			2319	1469	396	432	22			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	166	VAL	GLU	engineered mutation	UNP A0A7U3EXT3
B	166	VAL	GLU	engineered mutation	UNP A0A7U3EXT3

- Molecule 2 is {N}-[(2 {S})-3-cyclohexyl-1-oxidanylidene-1-[[(2 {S})-1-oxidanylidene-3-[(3 {S})-2-oxidanylidenepyrrolidin-3-yl]propan-2-yl]amino]propan-2-yl]-1 {H}-indole-2-carboxamide (CCD ID: FHR) (formula: $C_{25}H_{32}N_4O_4$) (labeled as "Ligand of Interest" by depositor).

FHR
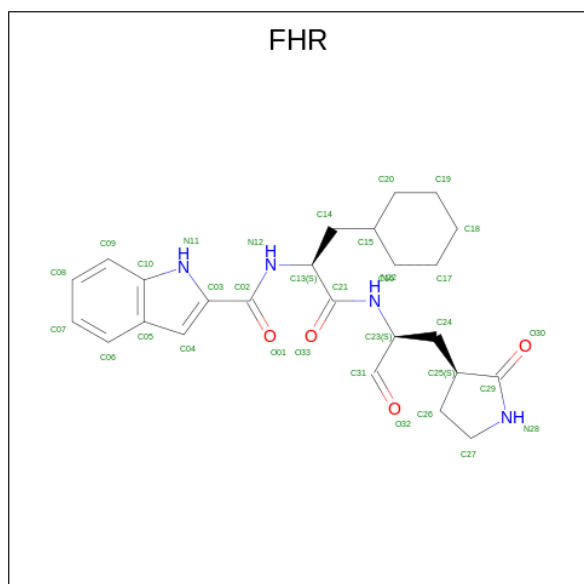
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 33	C 25	N 4	O 4	33	0
2	B	1	Total 33	C 25	N 4	O 4	33	0

• Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	121	Total 121	O 121	0	0
3	B	188	Total 188	O 188	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: ORF1a polyprotein

Chain A:

• Molecule 1: ORF1a polyprotein

Chain B:
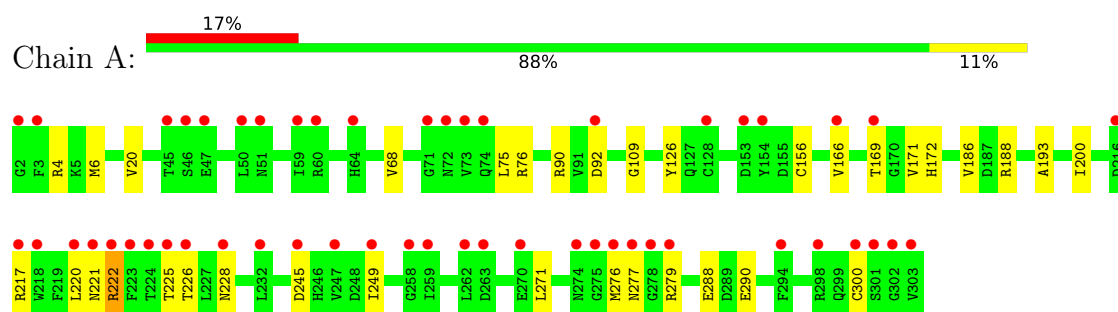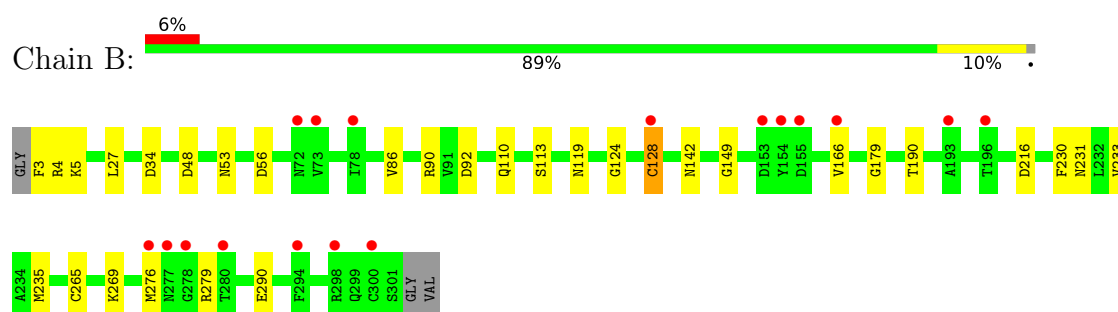
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	68.28Å 101.73Å 104.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.91 – 2.13 72.91 – 2.13	Depositor EDS
% Data completeness (in resolution range)	99.9 (72.91-2.13) 99.9 (72.91-2.13)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 2.12Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.210 , 0.243 0.211 , 0.243	Depositor DCC
R_{free} test set	2033 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.432	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 49.5	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5028	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: FHR

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	0.43	0/2386	0.60	0/3244
1	B	0.45	1/2371 (0.0%)	0.61	1/3224 (0.0%)
All	All	0.44	1/4757 (0.0%)	0.60	1/6468 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	128	CYS	CB-SG	-5.20	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	166	VAL	CG1-CB-CG2	5.05	118.99	110.90

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	2334	0	2284	26	0
1	B	2319	0	2269	21	0
2	A	33	0	0	0	0
2	B	33	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	121	0	0	10	1
3	B	188	0	0	8	1
All	All	5028	0	4553	45	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 5.

The worst 5 of 45 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:156:CYS:O	3:A:501:HOH:O	1.82	0.97
1:B:3:PHE:N	3:B:501:HOH:O	2.02	0.93
1:B:276:MET:HG2	1:B:279:ARG:O	1.79	0.83
1:A:288:GLU:OE1	3:A:502:HOH:O	1.98	0.79
1:B:110:GLN:NE2	3:B:504:HOH:O	2.16	0.77

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)
3:A:567:HOH:O	3:B:675:HOH:O[2_454]	2.05	0.15

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	A	300/302 (99%)	295 (98%)	5 (2%)	0	100	100
1	B	297/302 (98%)	290 (98%)	7 (2%)	0	100	100
All	All	597/604 (99%)	585 (98%)	12 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	259/259 (100%)	256 (99%)	3 (1%)	67	72
1	B	258/259 (100%)	254 (98%)	4 (2%)	58	63
All	All	517/518 (100%)	510 (99%)	7 (1%)	62	67

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	27	LEU
1	B	128	CYS
1	B	216	ASP
1	B	190	THR
1	A	222	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	53	ASN
1	B	110	GLN
1	B	189	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 ligands are 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	#$|Z| > 2$
2	FHR	A	401	1	33,36,36	5.99	16 (48%)	40,49,49	1.93	10 (25%)
2	FHR	B	401	1	33,36,36	6.33	17 (51%)	40,49,49	2.13	8 (20%)

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
2	FHR	A	401	1	-	1/22/44/44	0/4/4/4
2	FHR	B	401	1	-	4/22/44/44	0/4/4/4

The worst 5 of 33 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	FHR	C29-N28	26.58	1.62	1.33
2	A	401	FHR	C29-N28	24.69	1.59	1.33
2	B	401	FHR	C26-C27	11.74	1.71	1.53
2	A	401	FHR	C26-C27	9.87	1.68	1.53
2	B	401	FHR	C09-C10	9.36	1.58	1.41

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	FHR	C26-C27-N28	7.03	113.74	103.43
2	B	401	FHR	C27-N28-C29	-6.31	101.45	113.84
2	B	401	FHR	C26-C25-C29	5.62	110.19	102.88
2	A	401	FHR	C15-C14-C13	-5.07	107.70	114.52
2	A	401	FHR	C03-C02-N12	4.26	123.11	115.20

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	401	FHR	C13-C14-C15-C20
2	B	401	FHR	N12-C13-C14-C15
2	B	401	FHR	C21-C13-C14-C15
2	B	401	FHR	C13-C14-C15-C16
2	A	401	FHR	O01-C02-C03-C04

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand FHR A 401

Bond lengths

Bond angles

Torsions

Rings

Ligand FHR B 401

Bond lengths

Bond angles

Torsions

Rings

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 (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(Å²)	Q<0.9
1	A	302/302 (100%)	0.73	52 (17%)	5	7	17, 33, 52, 63	0
1	B	299/302 (99%)	0.12	17 (5%)	30	36	17, 26, 42, 63	0
All	All	601/604 (99%)	0.42	69 (11%)	11	13	17, 28, 50, 63	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	223	PHE	5.3
1	B	154	TYR	4.8
1	A	225	THR	4.4
1	A	278	GLY	3.9
1	A	302	GLY	3.9

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)

LIGAND-RSR INFOmissingINFO

6.5   Other polymers (i)

There are no such residues in this entry.