Full wwPDB X-ray Structure Validation Report ⓘ

Mar 17, 2026 – 11:39 PM UTC

PDB ID	:	7MB9 / pdb_00007mb9
Title	:	SARS-CoV-2 Main Protease (Mpro) C145A in Complex with Cleavage Site Nsp10/11 (P6-P1)
Authors	:	Lockbaum, G.J.; Schiffer, C.A.
Deposited on	:	2021-03-31
Resolution	:	1.81 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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:

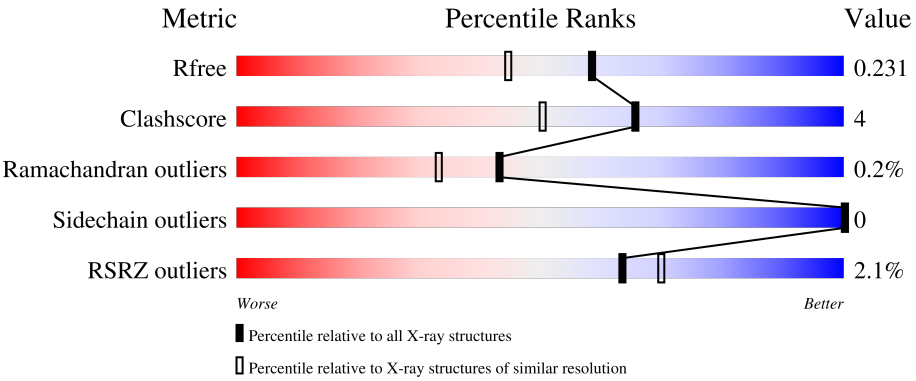
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

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

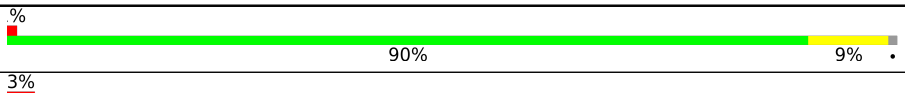
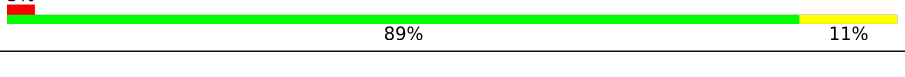
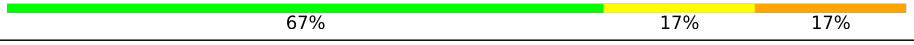
The reported resolution of this entry is 1.81 Å.

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}	180053	1112 (1.82-1.82)
Clashscore	190562	1148 (1.82-1.82)
Ramachandran outliers	187476	1140 (1.82-1.82)
Sidechain outliers	187428	1140 (1.82-1.82)
RSRZ outliers	180081	1112 (1.82-1.82)

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	306	
1	B	306	
2	C	6	
2	D	6	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9691 atoms, of which 4556 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 3C-like proteinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	302	Total	C	H	N	O	S	0	0	0
			4521	1455	2222	391	432	21			
1	B	305	Total	C	H	N	O	S	0	0	0
			4547	1468	2233	390	435	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	145	ALA	CYS	engineered mutation	UNP P0DTD1
B	145	ALA	CYS	engineered mutation	UNP P0DTD1

- Molecule 2 is a protein called ARG-GLU-PRO-MET-LEU-GLN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	6	Total	C	H	N	O	S	0	0	0
			109	32	56	10	10	1			
2	D	6	Total	C	H	N	O		0	0	0
			91	28	45	10	8				

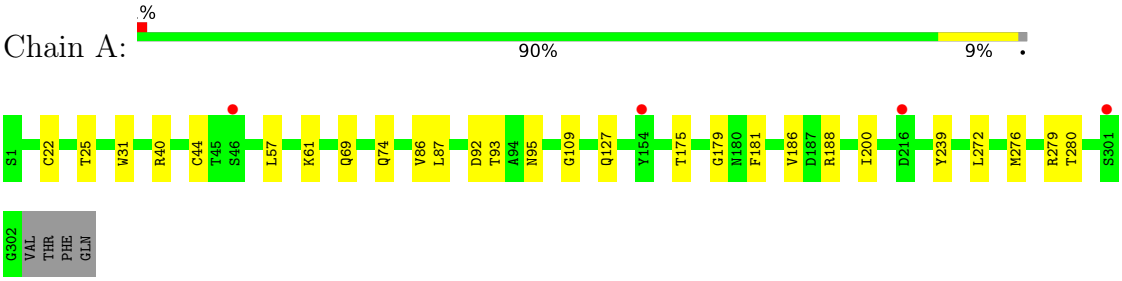
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	225	Total	O	0	0
			225	225		
3	B	191	Total	O	0	0
			191	191		
3	C	2	Total	O	0	0
			2	2		
3	D	5	Total	O	0	0
			5	5		

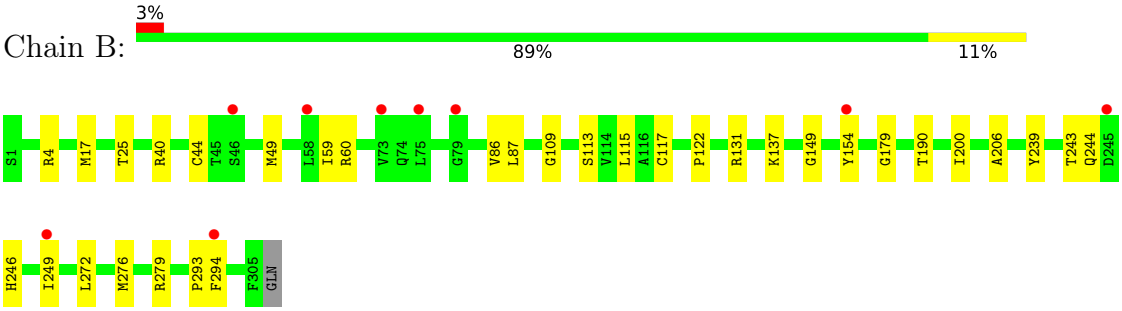
3 Residue-property plots (i)

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 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: 3C-like proteinase



• Molecule 1: 3C-like proteinase



• Molecule 2: ARG-GLU-PRO-MET-LEU-GLN

There are no outlier residues recorded for this chain.

• Molecule 2: ARG-GLU-PRO-MET-LEU-GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	54.22Å 97.89Å 67.51Å 90.00° 102.48° 90.00°	Depositor
Resolution (Å)	29.69 – 1.81 29.69 – 1.81	Depositor EDS
% Data completeness (in resolution range)	98.2 (29.69-1.81) 94.6 (29.69-1.81)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.22 (at 1.82Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.199 , 0.228 0.199 , 0.231	Depositor DCC
R_{free} test set	2000 reflections (3.27%)	wwPDB-VP
Wilson B-factor (Å^2)	28.3	Xtriage
Anisotropy	0.262	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 30.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9691	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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	0.55	0/2350	0.60	0/3200
1	B	0.57	0/2366	0.62	0/3222
2	C	0.83	0/53	0.62	0/68
2	D	0.45	0/46	0.87	1/60 (1.7%)
All	All	0.56	0/4815	0.61	1/6550 (0.0%)

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	D	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	-2	LEU	O-C-N	-5.61	116.12	123.02

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	-2	LEU	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	A	2299	2222	2221	15	0
1	B	2314	2233	2232	21	0
2	C	53	56	56	0	0
2	D	46	45	45	1	0
3	A	225	0	0	1	0
3	B	191	0	0	0	0
3	C	2	0	0	0	0
3	D	5	0	0	0	0
All	All	5135	4556	4554	35	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 4.

All (35) 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:B:59:ILE:HD12	1:B:60:ARG:N	1.98	0.77
1:A:186:VAL:HG23	1:A:188:ARG:HG2	1.77	0.65
1:A:25:THR:HG21	1:A:44:CYS:O	2.07	0.55
1:B:276:MET:HE2	1:B:279:ARG:O	2.07	0.54
1:B:86:VAL:HG13	1:B:179:GLY:HA2	1.90	0.53
1:A:109:GLY:HA2	1:A:200:ILE:HD13	1.90	0.52
1:B:25:THR:HG21	1:B:44:CYS:O	2.10	0.52
1:A:86:VAL:HG13	1:A:179:GLY:HA2	1.94	0.49
1:A:92:ASP:OD1	1:A:93:THR:HG23	2.12	0.49
1:B:113:SER:O	1:B:149:GLY:HA2	2.13	0.49
1:A:57:LEU:HD22	3:A:434:HOH:O	2.12	0.48
1:A:276:MET:HE1	1:A:280:THR:HA	1.95	0.48
1:A:22:CYS:SG	1:A:61:LYS:HE3	2.54	0.48
1:B:109:GLY:HA2	1:B:200:ILE:HD13	1.95	0.48
1:B:44:CYS:HB2	1:B:49:MET:HE2	1.95	0.47
1:B:59:ILE:HD12	1:B:59:ILE:C	2.40	0.47
1:B:244:GLN:O	1:B:244:GLN:HG3	2.15	0.46
1:B:249:ILE:HG22	1:B:293:PRO:HG2	1.99	0.45
1:B:190:THR:O	2:D:-4:PRO:HG2	2.16	0.45
1:A:127:GLN:O	1:B:4:ARG:NH2	2.49	0.44
1:B:40:ARG:HA	1:B:87:LEU:HG	1.99	0.44
1:B:206:ALA:HB2	1:B:293:PRO:HG3	2.00	0.43
1:A:276:MET:HE2	1:A:279:ARG:O	2.19	0.43
1:B:115:LEU:HD11	1:B:122:PRO:HB3	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:239:TYR:CZ	1:B:272:LEU:HD21	2.54	0.42
1:A:69:GLN:NE2	1:A:74:GLN:OE1	2.53	0.42
1:A:175:THR:HG22	1:A:181:PHE:HA	2.01	0.42
1:B:243:THR:O	1:B:246:HIS:HB2	2.19	0.42
1:B:294:PHE:CD1	1:B:294:PHE:N	2.88	0.42
1:B:249:ILE:HD12	1:B:294:PHE:CZ	2.55	0.42
1:B:17:MET:HG3	1:B:117:CYS:SG	2.61	0.41
1:A:40:ARG:HA	1:A:87:LEU:HG	2.02	0.41
1:A:239:TYR:CE1	1:A:272:LEU:HD21	2.56	0.41
1:A:31:TRP:CD2	1:A:95:ASN:HB2	2.57	0.40
1:B:131:ARG:HD3	1:B:137:LYS:HE3	2.03	0.40

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	300/306 (98%)	290 (97%)	10 (3%)	0	100 100
1	B	303/306 (99%)	292 (96%)	10 (3%)	1 (0%)	36 26
2	C	4/6 (67%)	4 (100%)	0	0	100 100
2	D	4/6 (67%)	4 (100%)	0	0	100 100
All	All	611/624 (98%)	590 (97%)	20 (3%)	1 (0%)	43 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	154	TYR

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	250/262 (95%)	250 (100%)	0	100	100
1	B	251/262 (96%)	251 (100%)	0	100	100
2	C	6/6 (100%)	6 (100%)	0	100	100
2	D	4/6 (67%)	4 (100%)	0	100	100
All	All	511/536 (95%)	511 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	69	GLN
1	A	74	GLN
1	A	107	GLN
1	A	180	ASN
1	B	107	GLN
1	B	238	ASN

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

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, 95^{th} 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(Å^2)	Q<0.9
1	A	302/306 (98%)	0.07	4 (1%) 75 80	26, 37, 54, 71	0
1	B	305/306 (99%)	0.29	9 (2%) 52 59	27, 38, 59, 73	0
2	C	6/6 (100%)	0.13	0 100 100	31, 41, 54, 59	0
2	D	6/6 (100%)	0.33	0 100 100	36, 44, 53, 66	0
All	All	619/624 (99%)	0.18	13 (2%) 63 70	26, 38, 57, 73	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	154	TYR	3.5
1	A	46	SER	3.2
1	B	58	LEU	2.9
1	B	154	TYR	2.7
1	B	79	GLY	2.5
1	A	216	ASP	2.5
1	B	245	ASP	2.5
1	A	301	SER	2.4
1	B	73	VAL	2.4
1	B	75	LEU	2.3
1	B	249	ILE	2.2
1	B	294	PHE	2.1
1	B	46	SER	2.1

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 (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.