Full wwPDB X-ray Structure Validation Report ⓘ

Feb 24, 2026 – 06:09 PM JST

PDB ID	:	9M29 / pdb_00009m29
Title	:	Crystal Structure of the SARS-CoV-2 (COVID-19) main protease with inhibitor AD05
Authors	:	Yin, W.; Kong, W.P.; Wong, K.Y.
Deposited on	:	2025-02-27
Resolution	:	1.97 Å(reported)

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

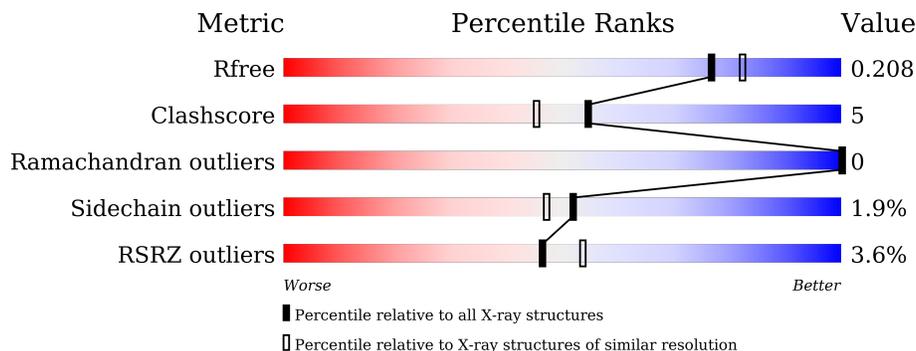|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.48.1

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 1.97 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		0.208
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		1.9%
RSRZ outliers		3.6%

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$_{free}$	164625	3187 (1.96-1.96)
Clashscore	180529	3412 (1.96-1.96)
Ramachandran outliers	177936	3390 (1.96-1.96)
Sidechain outliers	177891	3390 (1.96-1.96)
RSRZ outliers	164620	3186 (1.96-1.96)

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	310	4% / 81% / 14%
2	B	3	100%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4901 atoms, of which 2353 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 nsp5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	301	Total 4705	C 1501	H 2330	N 401	O 449	S 24	78	11	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	ILE	-	expression tag	UNP P0DTD1
A	308	ASN	-	expression tag	UNP P0DTD1
A	309	PRO	-	expression tag	UNP P0DTD1
A	310	SEC	-	expression tag	UNP P0DTD1

- Molecule 2 is a protein called inhibitor AD05.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
2	B	3	Total 50	C 17	Cl 1	H 23	N 4	O 4	S 1	1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	146	Total 146	O 146	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: 3C-like proteinase nsp5
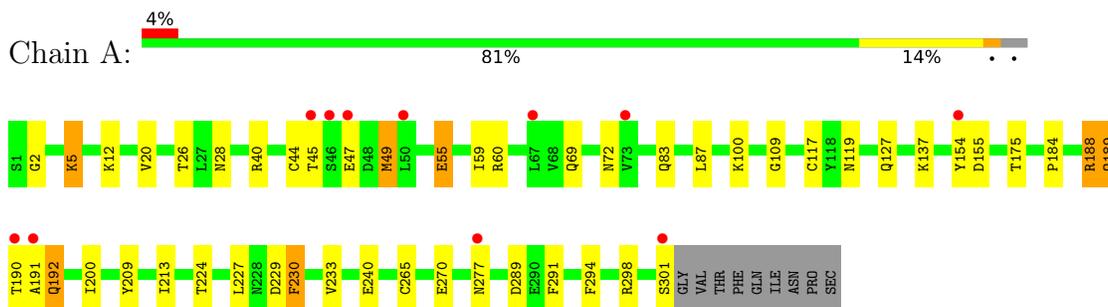
Chain A:

• Molecule 2: inhibitor AD05

Chain B:

There are no outlier residues recorded for this chain.

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	51.52Å 80.14Å 88.97Å 90.00° 97.47° 90.00°	Depositor
Resolution (Å)	23.68 – 1.97 23.68 – 1.97	Depositor EDS
% Data completeness (in resolution range)	99.3 (23.68-1.97) 98.7 (23.68-1.97)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.06 (at 1.96Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.176 , 0.216 0.173 , 0.208	Depositor DCC
R_{free} test set	1280 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	16.9	Xtriage
Anisotropy	0.497	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 41.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4901	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.58% 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: A1L8H, A1L8I

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	0.96	0/2464	1.53	32/3350 (1.0%)
2	B	1.12	0/7	1.22	0/8
All	All	0.96	0/2471	1.53	32/3358 (1.0%)

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
1	A	0	3

There are no bond length outliers.

All (32) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	60	ARG	CD-NE-CZ	10.03	138.44	124.40
1	A	192	GLN	CB-CA-C	-8.63	97.73	110.16
1	A	55	GLU	CB-CG-CD	8.52	127.08	112.60
1	A	127	GLN	N-CA-CB	-8.38	96.69	110.52
1	A	154	TYR	CB-CA-C	8.35	127.04	110.42
1	A	230	PHE	CA-CB-CG	-7.93	105.87	113.80
1	A	154	TYR	N-CA-CB	-7.90	97.14	110.49
1	A	60	ARG	CB-CG-CD	7.86	129.37	111.30
1	A	60	ARG	CB-CA-C	-7.83	95.88	110.36
1	A	119	ASN	CA-CB-CG	-7.71	104.89	112.60
1	A	20	VAL	N-CA-CB	7.48	121.99	111.82
1	A	60	ARG	NE-CZ-NH1	7.10	128.60	121.50
1	A	127	GLN	CB-CA-C	6.93	121.12	109.48
1	A	224	THR	CA-CB-OG1	-6.78	99.42	109.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	294	PHE	CA-CB-CG	-6.77	107.03	113.80
1	A	60	ARG	CG-CD-NE	-6.59	97.50	112.00
1	A	277	ASN	CB-CA-C	-6.56	103.08	111.86
1	A	44	CYS	CB-CA-C	-6.54	98.38	109.83
1	A	49	MET	CG-SD-CE	5.98	114.07	100.90
1	A	60	ARG	CA-CB-CG	5.90	125.91	114.10
1	A	12	LYS	CG-CD-CE	-5.88	97.78	111.30
1	A	26	THR	CA-CB-OG1	-5.78	100.92	109.60
1	A	270	GLU	CB-CA-C	5.78	119.96	110.88
1	A	289[A]	ASP	CA-CB-CG	5.78	118.38	112.60
1	A	289[B]	ASP	CA-CB-CG	5.78	118.38	112.60
1	A	175	THR	OG1-CB-CG2	5.72	120.73	109.30
1	A	100	LYS	N-CA-CB	5.58	118.26	109.83
1	A	5	LYS	CB-CA-C	5.57	119.79	110.88
1	A	2	GLY	N-CA-C	-5.42	106.22	112.29
1	A	47	GLU	CB-CA-C	-5.35	102.48	110.88
1	A	192	GLN	N-CA-CB	5.23	117.77	109.51
1	A	60	ARG	NE-CZ-NH2	-5.19	114.53	119.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	184	PRO	Mainchain
1	A	188	ARG	Sidechain
1	A	298	ARG	Sidechain

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	2375	2330	2296	22	0
2	B	27	23	10	0	0
3	A	146	0	0	5	1
All	All	2548	2353	2306	22	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.

All (22) 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:137:LYS:HE3	3:A:508:HOH:O	1.44	1.18
1:A:45:THR:O	1:A:49:MET:HG3	1.95	0.66
1:A:189:GLN:NE2	3:A:404:HOH:O	2.35	0.58
1:A:188:ARG:C	1:A:189:GLN:OE1	2.49	0.56
1:A:137:LYS:CE	3:A:508:HOH:O	2.21	0.56
1:A:40:ARG:HA	1:A:87:LEU:HG	1.90	0.53
1:A:189:GLN:OE1	1:A:189:GLN:N	2.42	0.53
1:A:55:GLU:O	1:A:59:ILE:HG13	2.11	0.50
1:A:190:THR:HG22	1:A:191:ALA:H	1.80	0.46
1:A:5:LYS:HE3	1:A:291:PHE:CZ	2.51	0.46
1:A:59:ILE:CG1	3:A:443:HOH:O	2.65	0.45
1:A:155:ASP:OD1	1:A:155:ASP:N	2.48	0.44
1:A:190:THR:HB	1:A:192:GLN:HG2	2.00	0.43
1:A:229:ASP:O	1:A:233:VAL:HG23	2.18	0.43
1:A:230:PHE:CD1	1:A:265:CYS:HB3	2.54	0.43
1:A:190:THR:C	1:A:192:GLN:H	2.27	0.43
1:A:5:LYS:HD3	1:A:5:LYS:HA	1.84	0.43
1:A:69:GLN:HG3	1:A:72:ASN:O	2.18	0.42
1:A:109:GLY:HA2	1:A:200:ILE:HD13	2.01	0.42
1:A:28:ASN:ND2	1:A:117[B]:CYS:O	2.52	0.42
1:A:209:TYR:O	1:A:213:ILE:HG13	2.21	0.41
1:A:59:ILE:HG13	3:A:443:HOH:O	2.21	0.40

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:432:HOH:O	3:A:481:HOH:O[2_555]	1.96	0.24

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	310/310 (100%)	297 (96%)	13 (4%)	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	268/266 (101%)	262 (98%)	6 (2%)	47	41
2	B	1/1 (100%)	1 (100%)	0	100	100
All	All	269/267 (101%)	263 (98%)	6 (2%)	52	41

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	GLN
1	A	189	GLN
1	A	227	LEU
1	A	240[A]	GLU
1	A	240[B]	GLU
1	A	301	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	ASN
1	A	74	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 ⓘ

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	A	301/310 (97%)	-0.28	11 (3%)	45	52	6, 20, 50, 83	7 (2%)
2	B	1/3 (33%)	-0.46	0	100	100	24, 24, 24, 24	0
All	All	302/313 (96%)	-0.28	11 (3%)	46	54	6, 20, 50, 83	7 (2%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	154	TYR	5.7
1	A	190	THR	3.4
1	A	191	ALA	3.2
1	A	50	LEU	2.6
1	A	46	SER	2.6
1	A	301	SER	2.4
1	A	47	GLU	2.3
1	A	73	VAL	2.2
1	A	45	THR	2.2
1	A	67	LEU	2.1
1	A	277	ASN	2.0

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 oligosaccharides in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.