Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2023 – 02:25 PM JST

PDB ID	:	7VJZ
Title	:	Crystal Structure of SARS-CoV-2 Mpro at 1.90 A resolution-7
Authors	:	DeMirci, H.; Tokay, N.
Deposited on	:	2021-09-29
Resolution	:	1.90 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:

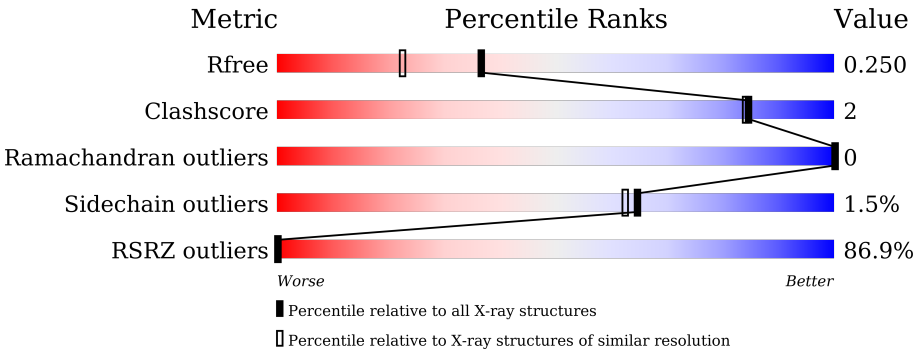
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	87%, 93%, 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2502 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 3C-like proteinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	306	2447	1545	416	462	24	0	10	0

- Molecule 2 is water.

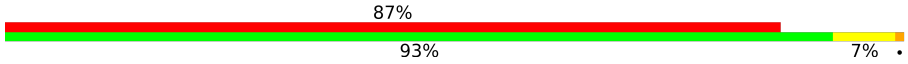
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	55	55	55	0	0

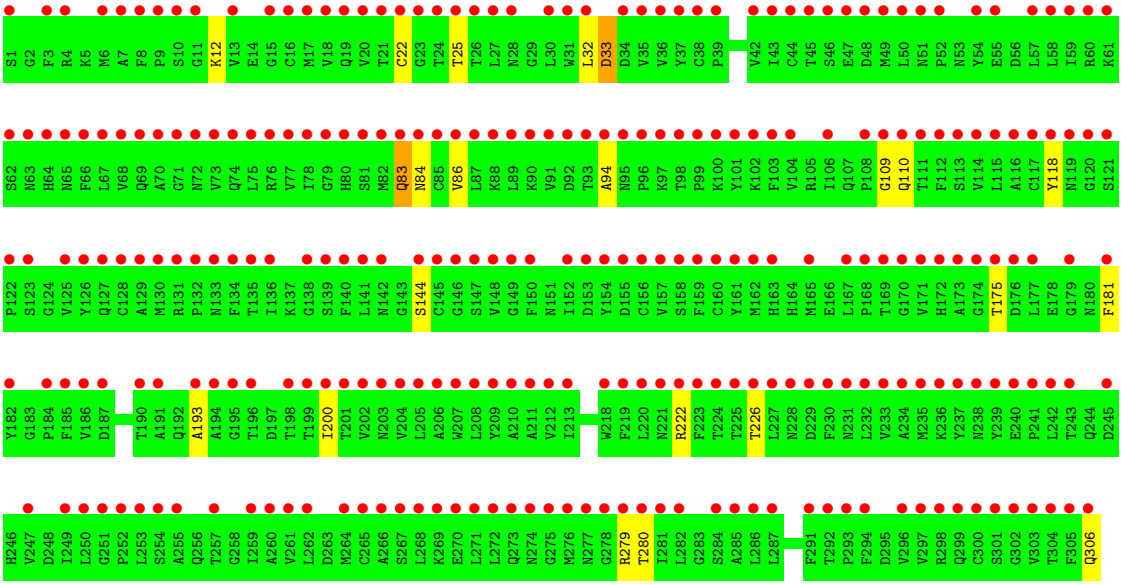
3 Residue-property plots (i)

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

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	115.70Å 55.20Å 45.60Å 90.00° 101.30° 90.00°	Depositor
Resolution (Å)	34.74 – 1.90 34.74 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (34.74-1.90) 100.0 (34.74-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 1.89Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.217 , 0.250 0.217 , 0.250	Depositor DCC
R_{free} test set	2000 reflections (8.94%)	wwPDB-VP
Wilson B-factor (Å^2)	36.2	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 66.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2502	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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	0.32	0/2500	0.53	0/3395

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	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	32	LEU	Peptide
1	A	83	GLN	Peptide

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	2447	0	2382	11	0
2	A	55	0	0	4	0
All	All	2502	0	2382	11	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 2.

All (11) 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:25:THR:N	2:A:401:HOH:O	2.18	0.77
1:A:22:CYS:O	2:A:401:HOH:O	2.02	0.76
1:A:12:LYS:NZ	1:A:306:GLN:O	2.21	0.70
1:A:110:GLN:OE1	2:A:402:HOH:O	2.15	0.64
1:A:279:ARG:HE	1:A:280:THR:H	1.52	0.56
1:A:109:GLY:HA2	1:A:200:ILE:HD13	1.91	0.52
1:A:193:ALA:O	2:A:403:HOH:O	2.19	0.51
1:A:118:TYR:CE1	1:A:144:SER:HB3	2.49	0.48
1:A:33:ASP:O	1:A:94:ALA:HA	2.15	0.46
1:A:175:THR:HG22	1:A:181:PHE:HA	1.98	0.46
1:A:83:GLN:O	1:A:86:VAL:HG12	2.16	0.46

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	314/306 (103%)	308 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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 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	273/263 (104%)	268 (98%)	5 (2%)	59 55

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	33	ASP
1	A	84	ASN
1	A	222[A]	ARG
1	A	222[B]	ARG
1	A	226	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	151	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.

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	306/306 (100%)	3.66	266 (86%)	0	0	24, 42, 71, 91	0

All (266) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	72	ASN	9.8
1	A	277	ASN	9.6
1	A	306	GLN	9.3
1	A	78	ILE	9.3
1	A	154	TYR	9.3
1	A	222[A]	ARG	8.3
1	A	77	VAL	8.2
1	A	153	ASP	8.1
1	A	303	VAL	7.6
1	A	223	PHE	7.5
1	A	93	THR	7.4
1	A	262[A]	LEU	6.8
1	A	24	THR	6.8
1	A	226	THR	6.6
1	A	79	GLY	6.6
1	A	75	LEU	6.6
1	A	62	SER	6.3
1	A	74	GLN	6.3
1	A	68	VAL	6.3
1	A	91	VAL	6.3
1	A	234	ALA	6.2
1	A	42	VAL	6.2
1	A	22	CYS	6.1
1	A	155	ASP	5.9
1	A	92	ASP	5.9
1	A	80	HIS	5.7
1	A	135[A]	THR	5.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	125	VAL	5.7
1	A	73	VAL	5.7
1	A	171	VAL	5.6
1	A	67	LEU	5.5
1	A	230	PHE	5.4
1	A	85	CYS	5.4
1	A	279	ARG	5.3
1	A	160	CYS	5.3
1	A	27	LEU	5.2
1	A	114	VAL	5.2
1	A	25	THR	5.2
1	A	233	VAL	5.1
1	A	225	THR	5.1
1	A	116	ALA	5.1
1	A	13	VAL	5.1
1	A	305	PHE	5.0
1	A	46[A]	SER	5.0
1	A	232	LEU	5.0
1	A	296	VAL	5.0
1	A	280	THR	4.9
1	A	247	VAL	4.9
1	A	136	ILE	4.9
1	A	31	TRP	4.9
1	A	271	LEU	4.9
1	A	235	MET	4.9
1	A	207	TRP	4.9
1	A	117	CYS	4.8
1	A	64	HIS	4.8
1	A	70	ALA	4.8
1	A	221	ASN	4.8
1	A	38	CYS	4.7
1	A	224	THR	4.7
1	A	126	TYR	4.7
1	A	98	THR	4.6
1	A	229	ASP	4.6
1	A	35	VAL	4.6
1	A	157	VAL	4.6
1	A	52	PRO	4.6
1	A	259	ILE	4.6
1	A	213	ILE	4.6
1	A	8	PHE	4.5
1	A	297	VAL	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	148	VAL	4.5
1	A	36	VAL	4.5
1	A	194	ALA	4.5
1	A	18	VAL	4.4
1	A	294	PHE	4.4
1	A	115	LEU	4.4
1	A	69	GLN	4.4
1	A	301[A]	SER	4.4
1	A	150	PHE	4.4
1	A	274	ASN	4.4
1	A	60	ARG	4.3
1	A	145	CYS	4.3
1	A	99	PRO	4.3
1	A	65[A]	ASN	4.3
1	A	111	THR	4.3
1	A	241	PRO	4.2
1	A	66	PHE	4.2
1	A	212	VAL	4.2
1	A	142	ASN	4.2
1	A	261	VAL	4.2
1	A	238	ASN	4.2
1	A	32	LEU	4.2
1	A	191	ALA	4.2
1	A	193	ALA	4.2
1	A	34	ASP	4.1
1	A	156	CYS	4.1
1	A	211	ALA	4.1
1	A	242	LEU	4.1
1	A	281	ILE	4.1
1	A	128	CYS	4.1
1	A	165[A]	MET	4.1
1	A	7	ALA	4.1
1	A	272	LEU	4.0
1	A	265	CYS	4.0
1	A	47[A]	GLU	4.0
1	A	268	LEU	4.0
1	A	86	VAL	4.0
1	A	202	VAL	4.0
1	A	132	PRO	4.0
1	A	123	SER	4.0
1	A	106	ILE	4.0
1	A	76	ARG	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	139	SER	3.9
1	A	57	LEU	3.9
1	A	130	MET	3.9
1	A	21	THR	3.9
1	A	30	LEU	3.8
1	A	204	VAL	3.8
1	A	161	TYR	3.8
1	A	293	PRO	3.8
1	A	266	ALA	3.8
1	A	237	TYR	3.8
1	A	94	ALA	3.8
1	A	227	LEU	3.8
1	A	147	SER	3.8
1	A	63	ASN	3.8
1	A	228	ASN	3.8
1	A	206	ALA	3.8
1	A	300	CYS	3.8
1	A	81	SER	3.8
1	A	177	LEU	3.8
1	A	43	ILE	3.8
1	A	175	THR	3.8
1	A	200	ILE	3.7
1	A	291	PHE	3.7
1	A	275	GLY	3.7
1	A	44	CYS	3.7
1	A	26	THR	3.7
1	A	218	TRP	3.7
1	A	10	SER	3.7
1	A	103	PHE	3.7
1	A	181	PHE	3.7
1	A	169	THR	3.6
1	A	201	THR	3.6
1	A	129	ALA	3.6
1	A	20	VAL	3.6
1	A	104	VAL	3.6
1	A	198	THR	3.6
1	A	168	PRO	3.6
1	A	118	TYR	3.6
1	A	276	MET	3.6
1	A	250	LEU	3.6
1	A	257	THR	3.6
1	A	304	THR	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	101	TYR	3.5
1	A	209	TYR	3.5
1	A	89	LEU	3.5
1	A	1	SER	3.5
1	A	292	THR	3.5
1	A	16	CYS	3.5
1	A	95	ASN	3.4
1	A	112	PHE	3.4
1	A	152	ILE	3.4
1	A	255	ALA	3.4
1	A	220	LEU	3.4
1	A	173	ALA	3.3
1	A	140	PHE	3.3
1	A	45	THR	3.3
1	A	59	ILE	3.3
1	A	282	LEU	3.3
1	A	196	THR	3.3
1	A	119	ASN	3.2
1	A	195	GLY	3.2
1	A	251	GLY	3.2
1	A	287	LEU	3.2
1	A	122	PRO	3.2
1	A	144	SER	3.2
1	A	186	VAL	3.2
1	A	278	GLY	3.2
1	A	113	SER	3.2
1	A	185	PHE	3.2
1	A	260	ALA	3.2
1	A	50	LEU	3.2
1	A	231	ASN	3.2
1	A	149	GLY	3.2
1	A	285	ALA	3.2
1	A	162	MET	3.2
1	A	141	LEU	3.1
1	A	163	HIS	3.1
1	A	167	LEU	3.1
1	A	6[A]	MET	3.1
1	A	184	PRO	3.1
1	A	90	LYS	3.0
1	A	252	PRO	3.0
1	A	61	LYS	3.0
1	A	17	MET	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	219	PHE	3.0
1	A	49	MET	3.0
1	A	210	ALA	3.0
1	A	9	PRO	3.0
1	A	108	PRO	3.0
1	A	286	LEU	3.0
1	A	190	THR	3.0
1	A	71	GLY	3.0
1	A	270	GLU	2.9
1	A	208	LEU	2.9
1	A	240	GLU	2.8
1	A	51	ASN	2.8
1	A	187	ASP	2.8
1	A	54	TYR	2.8
1	A	15	GLY	2.8
1	A	205	LEU	2.8
1	A	39	PRO	2.8
1	A	249	ILE	2.8
1	A	121	SER	2.8
1	A	182	TYR	2.8
1	A	138	GLY	2.7
1	A	133	ASN	2.7
1	A	159	PHE	2.6
1	A	264	MET	2.6
1	A	302	GLY	2.6
1	A	253	LEU	2.6
1	A	172	HIS	2.6
1	A	239	TYR	2.5
1	A	267	SER	2.5
1	A	127	GLN	2.5
1	A	4	ARG	2.5
1	A	254	SER	2.5
1	A	88	LYS	2.5
1	A	273	GLN	2.5
1	A	97	LYS	2.5
1	A	11	GLY	2.5
1	A	28	ASN	2.5
1	A	58	LEU	2.5
1	A	96	PRO	2.4
1	A	174	GLY	2.4
1	A	245	ASP	2.4
1	A	179	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	203	ASN	2.4
1	A	82	MET	2.4
1	A	84	ASN	2.4
1	A	284	SER	2.4
1	A	102	LYS	2.3
1	A	236	LYS	2.3
1	A	199	THR	2.3
1	A	120	GLY	2.3
1	A	243	THR	2.3
1	A	134	PHE	2.3
1	A	55[A]	GLU	2.3
1	A	131	ARG	2.3
1	A	100	LYS	2.3
1	A	87	LEU	2.3
1	A	48	ASP	2.2
1	A	109	GLY	2.2
1	A	146	GLY	2.2
1	A	37	TYR	2.2
1	A	23	GLY	2.2
1	A	176	ASP	2.2
1	A	19	GLN	2.2
1	A	83	GLN	2.1
1	A	3	PHE	2.1
1	A	158	SER	2.1
1	A	269	LYS	2.1
1	A	299	GLN	2.1
1	A	298	ARG	2.1
1	A	170	GLY	2.1
1	A	110	GLN	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.