



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2021 – 07:20 PM JST

PDB ID : 7CJD
Title : Crystal structure of the SARS-CoV-2 PLpro C111S mutant
Authors : Gao, X.; Cui, S.
Deposited on : 2020-07-10
Resolution : 2.50 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.18
buster-report : 1.1.7 (2018)
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) : 2.18

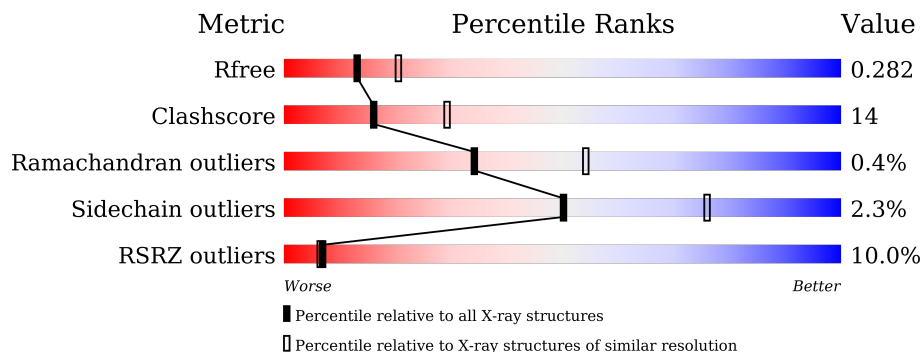
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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	325	<div> <div>5%</div> <div>68%</div> <div>22%</div> <div>• 9%</div> </div>
1	B	325	<div> <div>9%</div> <div>68%</div> <div>27%</div> <div>•</div> </div>
1	C	325	<div> <div>16%</div> <div>60%</div> <div>33%</div> <div>•• 5%</div> </div>
1	D	325	<div> <div>9%</div> <div>75%</div> <div>21%</div> <div>•</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 19318 atoms, of which 9498 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 Non-structural protein 3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	D	313	Total	C	H	N	O	S	0	0	0
			4906	1580	2422	410	476	18			
1	A	297	Total	C	H	N	O	S	0	0	0
			4649	1502	2292	389	450	16			
1	B	311	Total	C	H	N	O	S	0	0	0
			4876	1571	2408	408	471	18			
1	C	310	Total	C	H	N	O	S	0	0	0
			4831	1566	2370	407	470	18			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P0DTD1
D	111	SER	CYS	engineered mutation	UNP P0DTD1
D	319	HIS	-	expression tag	UNP P0DTD1
D	320	HIS	-	expression tag	UNP P0DTD1
D	321	HIS	-	expression tag	UNP P0DTD1
D	322	HIS	-	expression tag	UNP P0DTD1
D	323	HIS	-	expression tag	UNP P0DTD1
D	324	HIS	-	expression tag	UNP P0DTD1
A	0	MET	-	initiating methionine	UNP P0DTD1
A	111	SER	CYS	engineered mutation	UNP P0DTD1
A	319	HIS	-	expression tag	UNP P0DTD1
A	320	HIS	-	expression tag	UNP P0DTD1
A	321	HIS	-	expression tag	UNP P0DTD1
A	322	HIS	-	expression tag	UNP P0DTD1
A	323	HIS	-	expression tag	UNP P0DTD1
A	324	HIS	-	expression tag	UNP P0DTD1
B	0	MET	-	initiating methionine	UNP P0DTD1
B	111	SER	CYS	engineered mutation	UNP P0DTD1
B	319	HIS	-	expression tag	UNP P0DTD1
B	320	HIS	-	expression tag	UNP P0DTD1
B	321	HIS	-	expression tag	UNP P0DTD1

Continued on next page...

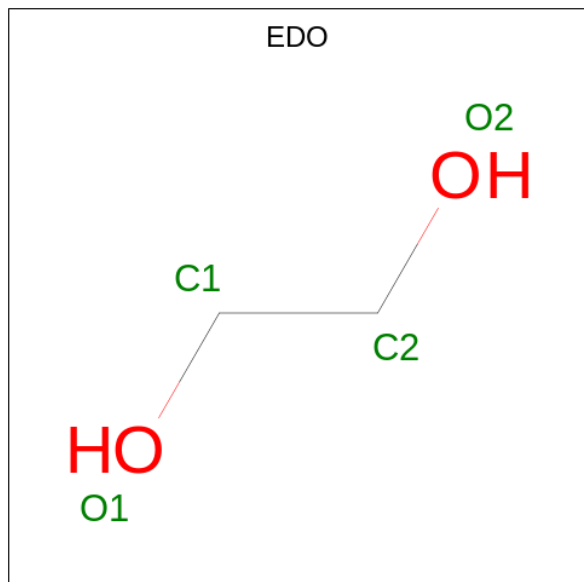
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	322	HIS	-	expression tag	UNP P0DTD1
B	323	HIS	-	expression tag	UNP P0DTD1
B	324	HIS	-	expression tag	UNP P0DTD1
C	0	MET	-	initiating methionine	UNP P0DTD1
C	111	SER	CYS	engineered mutation	UNP P0DTD1
C	319	HIS	-	expression tag	UNP P0DTD1
C	320	HIS	-	expression tag	UNP P0DTD1
C	321	HIS	-	expression tag	UNP P0DTD1
C	322	HIS	-	expression tag	UNP P0DTD1
C	323	HIS	-	expression tag	UNP P0DTD1
C	324	HIS	-	expression tag	UNP P0DTD1

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0
2	B	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	1	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).

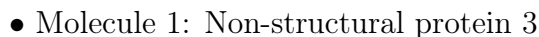
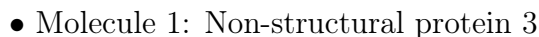


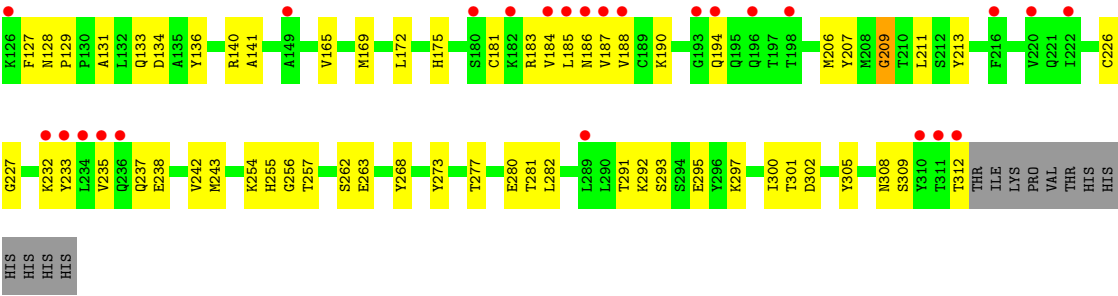
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 4 is water.

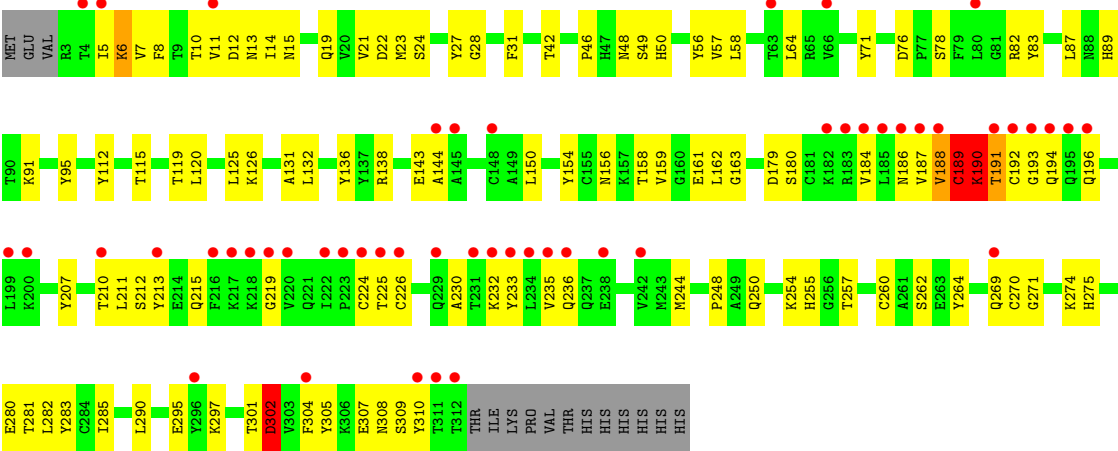
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	14	Total	O	0	0
			14	14		
4	A	15	Total	O	0	0
			15	15		
4	B	5	Total	O	0	0
			5	5		
4	C	8	Total	O	0	0
			8	8		

- Molecule 1: Non-structural protein 3





• Molecule 1: Non-structural protein 3



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	169.02Å 108.63Å 100.27Å 90.00° 118.87° 90.00°	Depositor
Resolution (Å)	43.90 – 2.50 46.58 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.8 (43.90-2.50) 99.5 (46.58-2.50)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.233 , 0.282 0.234 , 0.282	Depositor DCC
R_{free} test set	2727 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	62.0	Xtriage
Anisotropy	0.176	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 45.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for -1/2*h+1/2*k+1,1/2*h-1/2*k+1,1/2*h+1/2*k 0.000 for -1/2*h-1/2*k+1,-1/2*h-1/2*k-1,1/2*h-1/2*k	Xtriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	19318	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: EDO, ZN

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	0.53	2/2412 (0.1%)	0.64	0/3271
1	B	0.44	0/2527	0.61	0/3429
1	C	0.44	0/2520	0.62	0/3419
1	D	0.46	1/2543 (0.0%)	0.63	0/3451
All	All	0.47	3/10002 (0.0%)	0.62	0/13570

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	C	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	226	CYS	CB-SG	-7.66	1.69	1.82
1	A	189	CYS	CB-SG	-6.95	1.70	1.82
1	D	226	CYS	CB-SG	-5.13	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	188	VAL	Peptide
1	C	189	CYS	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	2357	2292	2292	57	0
1	B	2468	2408	2407	64	2
1	C	2461	2370	2400	121	1
1	D	2484	2422	2423	46	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	4	6	6	1	0
4	A	15	0	0	2	0
4	B	5	0	0	0	0
4	C	8	0	0	2	0
4	D	14	0	0	0	0
All	All	9820	9498	9528	280	2

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 14.

The worst 5 of 280 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:C:190:LYS:C	1:C:190:LYS:HE3	1.61	1.17
1:C:190:LYS:HD3	1:C:225:THR:OG1	1.45	1.16
1:C:186:ASN:ND2	1:C:194:GLN:OE1	1.90	1.03
1:C:213:TYR:OH	1:C:307:GLU:OE2	1.78	1.00
1:C:48:ASN:O	1:C:50:HIS:N	1.99	0.96

All (2) 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)
1:B:295:GLU:OE2	1:C:6:LYS:NZ[3_545]	2.10	0.10
1:D:126:LYS:NZ	1:B:141:ALA:O[4_555]	2.15	0.05

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	291/325 (90%)	272 (94%)	19 (6%)	0	100	100
1	B	309/325 (95%)	290 (94%)	18 (6%)	1 (0%)	41	61
1	C	308/325 (95%)	287 (93%)	17 (6%)	4 (1%)	12	21
1	D	311/325 (96%)	297 (96%)	14 (4%)	0	100	100
All	All	1219/1300 (94%)	1146 (94%)	68 (6%)	5 (0%)	34	54

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	49	SER
1	C	190	LYS
1	C	189	CYS
1	B	209	GLY
1	C	302	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	A	258/285 (90%)	252 (98%)	6 (2%)	50	76
1	B	271/285 (95%)	265 (98%)	6 (2%)	52	77
1	C	270/285 (95%)	260 (96%)	10 (4%)	34	60
1	D	273/285 (96%)	270 (99%)	3 (1%)	73	89
All	All	1072/1140 (94%)	1047 (98%)	25 (2%)	50	76

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	297	LYS
1	C	143	GLU
1	C	302	ASP
1	C	14	ILE
1	C	179	ASP

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

Mol	Chain	Res	Type
1	D	146	ASN
1	D	174	GLN
1	B	133	GLN

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](#)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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
3	EDO	A	502	-	3,3,3	0.55	0	2,2,2	0.40	0

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
3	EDO	A	502	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	EDO	1	0

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 [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(Å ²)	Q<0.9
1	A	297/325 (91%)	0.63	15 (5%) 28 29	44, 59, 88, 114	0
1	B	311/325 (95%)	0.76	28 (9%) 9 9	50, 67, 93, 102	0
1	C	310/325 (95%)	1.07	51 (16%) 1 1	47, 71, 113, 129	0
1	D	313/325 (96%)	0.70	29 (9%) 8 8	45, 61, 97, 114	0
All	All	1231/1300 (94%)	0.79	123 (9%) 7 6	44, 65, 100, 129	0

The worst 5 of 123 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	192	CYS	7.9
1	C	224	CYS	7.0
1	D	2	VAL	6.9
1	C	312	THR	6.2
1	B	311	THR	5.9

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.