WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Aug 30, 2023 – 12:06 PM EDT

PDB ID	:	7TX3
Title	:	Neutron crystal structure of SARS-CoV-2 NSP3 macrodomain at 293 K (P43 crystal form)
Authors	:	Correy, G.J.; Fraser, J.S.; Meilleur, F.
Deposited on	:	2022-02-07
Resolution	:	1.60 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 (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:

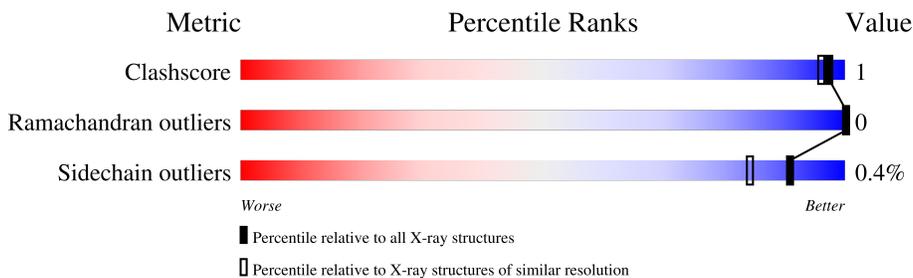|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.60 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
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0.4%

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))
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	169	96% ··
1	B	169	97% ·

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6411 atoms, of which 2572 are hydrogens and 969 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	A	166	Total	C	D	H	N	O	S	79	161	0
			2801	798	279	1265	216	239	4			
1	B	169	Total	C	D	H	N	O	S	69	164	0
			2853	811	284	1291	219	243	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P0DTD1
A	2	MET	-	expression tag	UNP P0DTD1
B	1	SER	-	expression tag	UNP P0DTD1
B	2	MET	-	expression tag	UNP P0DTD1

- Molecule 2 is 2-[N-CYCLOHEXYLAMINO]ETHANE SULFONIC ACID (three-letter code: NHE) (formula: $C_8H_{17}NO_3S$).
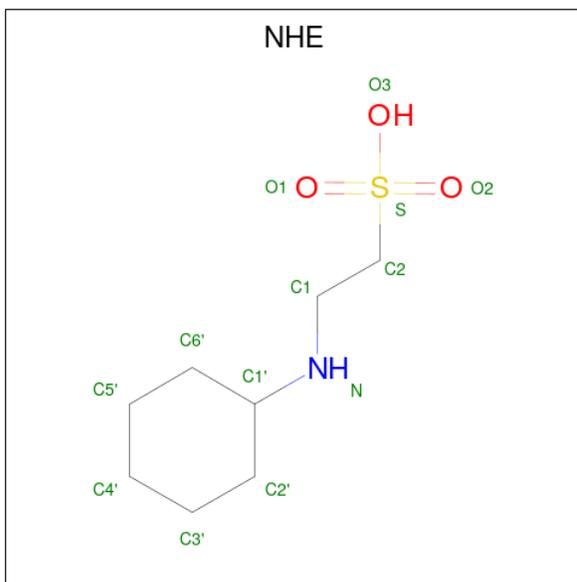
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total C H N O S 29 8 16 1 3 1						0	1

- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	146	Total D O 322 176 146			0	5
3	B	176	Total D O 406 230 176			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. 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. 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.

Note EDS failed to run properly.

• Molecule 1: Non-structural protein 3

Chain A:
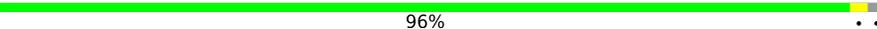
• Molecule 1: Non-structural protein 3

Chain B:

4 Data and refinement statistics ⓘ

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α, β, γ	89.33Å 89.33Å 40.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.14 – 1.60	Depositor
% Data completeness (in resolution range)	99.6 (23.14-1.60)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.92 (at 1.60Å)	Xtriage
Refinement program	PHENIX 1.20_4459	Depositor
R, R_{free}	0.124 , 0.158	Depositor
Wilson B-factor (Å^2)	10.9	Xtriage
Anisotropy	0.243	Xtriage
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
Total number of atoms	6411	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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: NHE

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	0.84	6/2518 (0.2%)	0.83	2/3417 (0.1%)
1	B	0.81	4/2560 (0.2%)	0.87	0/3473
All	All	0.83	10/5078 (0.2%)	0.85	2/6890 (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
1	B	0	2

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	120[A]	GLU	CG-CD	8.19	1.64	1.51
1	A	120[B]	GLU	CG-CD	8.19	1.64	1.51
1	A	26[A]	GLU	CB-CG	-6.68	1.39	1.52
1	A	26[B]	GLU	CB-CG	-6.68	1.39	1.52
1	A	120[A]	GLU	CD-OE1	5.77	1.31	1.25
1	A	120[B]	GLU	CD-OE1	5.77	1.31	1.25
1	B	144[A]	VAL	CB-CG2	-5.51	1.41	1.52
1	B	144[B]	VAL	CB-CG2	-5.51	1.41	1.52
1	B	25[A]	GLU	CG-CD	5.30	1.59	1.51
1	B	25[B]	GLU	CG-CD	5.30	1.59	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126[A]	LEU	CB-CG-CD1	-5.49	101.67	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126[B]	LEU	CB-CG-CD1	-5.49	101.67	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	141[A]	ARG	Sidechain
1	B	141[B]	ARG	Sidechain

CLOSE-CONTACTS INFOmissingINFO

5.2   Torsion angles ⓘ

5.2.1   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	323/169 (191%)	323 (100%)	0	0	100	100
1	B	329/169 (195%)	329 (100%)	0	0	100	100
All	All	652/338 (193%)	652 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

5.2.2   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	271/141 (192%)	271 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	277/141 (196%)	275 (99%)	2 (1%)	84	73
All	All	548/282 (194%)	546 (100%)	2 (0%)	91	84

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	29[A]	LYS
1	B	29[B]	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.2.3 RNA ⓘ

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.5 Ligand geometry ⓘ

1 ligand is 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	NHE	A	201[A]	-	13,13,13	1.49	2 (15%)	16,17,17	2.28	4 (25%)

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	NHE	A	201[A]	-	-	4/7/15/15	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	201[A]	NHE	C2-S	4.18	1.83	1.77
2	A	201[A]	NHE	O1-S	2.03	1.51	1.45

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	201[A]	NHE	O2-S-C2	6.17	114.34	106.92
2	A	201[A]	NHE	O1-S-C2	3.77	111.45	106.92
2	A	201[A]	NHE	C3'-C2'-C1'	3.16	117.05	111.11
2	A	201[A]	NHE	O3-S-O1	-2.95	104.08	111.27

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201[A]	NHE	C1-C2-S-O3
2	A	201[A]	NHE	C1-C2-S-O1
2	A	201[A]	NHE	C1-C2-S-O2
2	A	201[A]	NHE	C2-C1-N-C1'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201[A]	NHE	1	0

5.6 Other polymers ⓘ

There are no such residues in this entry.

5.7 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

6.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.