wwPDB EM Validation Summary Report (i)

Sep 18, 2021 – 08:03 AM BST

PDB ID	:	7OZU
EMDB ID	:	EMD-13135
Title	:	SARS-CoV-2 RdRp with Molnupiravir/ NHC in the template strand base-paired with A
Authors	:	Kabinger, F.; Stiller, C.; Schmitzova, J.; Dienemann, C.; Kokic, G.; Hillen, H.S.; Hoebartner, C.; Cramer, P.
Deposited on	:	2021-06-28
Resolution	:	3.30 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

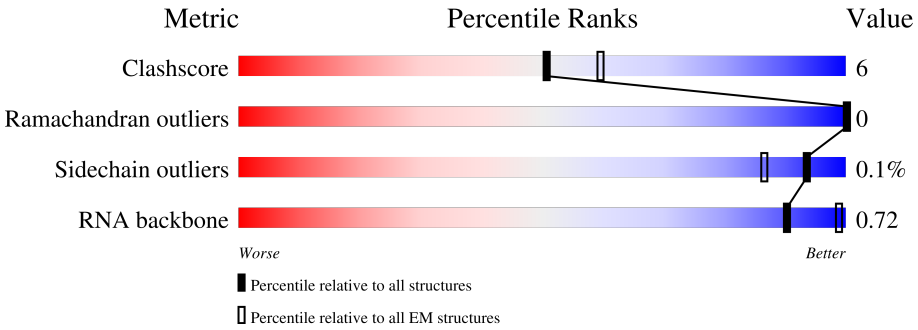
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.30 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	A	932	74% 13% 13%
2	B	217	44% 9% 47%
3	C	84	57% 17% 26%
4	P	32	25% 6% 69%
5	T	33	33% • 64%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8408 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Replicase polyprotein 1ab.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	814	Total	C	N	O	S	0	0
			6565	4209	1092	1219	45		

- Molecule 2 is a protein called Non-structural protein 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	115	Total	C	N	O	S	0	0
			891	562	149	173	7		

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	MET	-	initiating methionine	UNP P0DTD1
B	-17	GLY	-	expression tag	UNP P0DTD1
B	-16	SER	-	expression tag	UNP P0DTD1
B	-15	SER	-	expression tag	UNP P0DTD1
B	-14	HIS	-	expression tag	UNP P0DTD1
B	-13	HIS	-	expression tag	UNP P0DTD1
B	-12	HIS	-	expression tag	UNP P0DTD1
B	-11	HIS	-	expression tag	UNP P0DTD1
B	-10	HIS	-	expression tag	UNP P0DTD1
B	-9	HIS	-	expression tag	UNP P0DTD1
B	-8	GLU	-	expression tag	UNP P0DTD1
B	-7	ASN	-	expression tag	UNP P0DTD1
B	-6	LEU	-	expression tag	UNP P0DTD1
B	-5	TYR	-	expression tag	UNP P0DTD1
B	-4	PHE	-	expression tag	UNP P0DTD1
B	-3	GLN	-	expression tag	UNP P0DTD1
B	-2	SER	-	expression tag	UNP P0DTD1
B	-1	ASN	-	expression tag	UNP P0DTD1
B	0	ALA	-	expression tag	UNP P0DTD1

- Molecule 3 is a protein called Non-structural protein 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	62	Total	C	N	O	S	0	0
			478	303	78	92	5		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	SER	-	expression tag	UNP P0DTD1
C	-1	ASN	-	expression tag	UNP P0DTD1
C	0	ALA	-	expression tag	UNP P0DTD1

- Molecule 4 is a RNA chain called Product RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	P	10	Total	C	N	O	P	0	0
			212	95	38	69	10		

- Molecule 5 is a RNA chain called Template RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	T	12	Total	C	N	O	P	0	0
			260	115	48	85	12		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

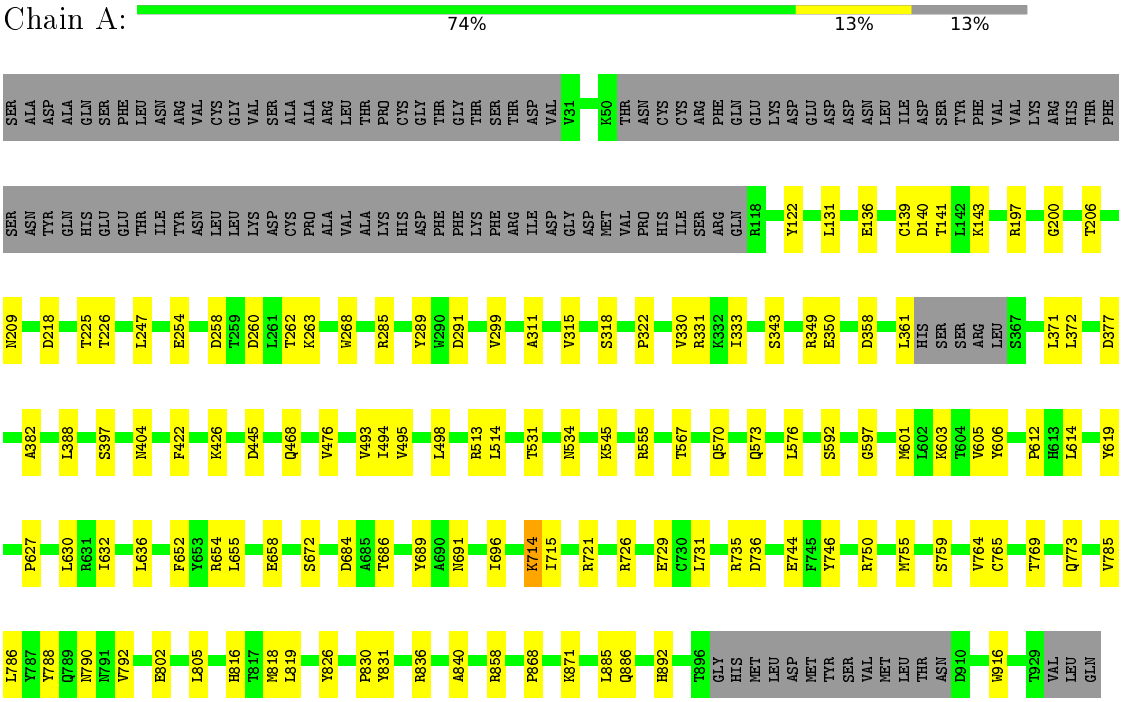
Mol	Chain	Residues	Atoms		AltConf
6	A	2	Total	Zn	0
			2	2	

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. 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.

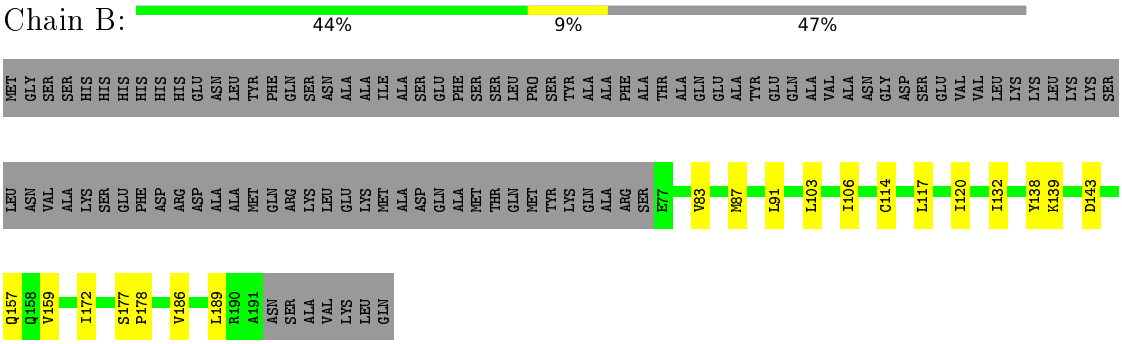
- Molecule 1: Replicase polyprotein 1ab

Chain A:



- Molecule 2: Non-structural protein 8

Chain B:



- Molecule 3: Non-structural protein 7

Chain C:

- Molecule 4: Product RNA

Chain P:

- Molecule 5: Template RNA

Chain T:

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	373938	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	59.6	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 7OK

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	0.27	0/6731	0.46	0/9135
2	B	0.24	0/904	0.48	0/1233
3	C	0.23	0/481	0.40	0/648
4	P	0.18	0/236	0.67	0/365
5	T	0.20	0/266	0.67	0/411
All	All	0.26	0/8618	0.48	0/11792

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	6565	0	6343	78	0
2	B	891	0	906	14	0
3	C	478	0	512	8	0
4	P	212	0	109	3	0
5	T	260	0	120	1	0
6	A	2	0	0	0	0
All	All	8408	0	7990	96	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 6.

The worst 5 of 96 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:330:VAL:HG21	2:B:120:ILE:HD11	1.62	0.81
1:A:358:ASP:HB2	1:A:534:ASN:HD21	1.47	0.80
1:A:426:LYS:NZ	1:A:886:GLN:OE1	2.22	0.73
1:A:514:LEU:HD21	2:B:83:VAL:HG21	1.72	0.71
1:A:531:THR:HG21	1:A:567:THR:HG21	1.76	0.68

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 PDB entries followed by that with respect to all EM entries.

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	806/932 (86%)	778 (96%)	28 (4%)	0	100	100
2	B	113/217 (52%)	111 (98%)	2 (2%)	0	100	100
3	C	60/84 (71%)	60 (100%)	0	0	100	100
All	All	979/1233 (79%)	949 (97%)	30 (3%)	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 PDB entries followed by that with respect to all EM entries.

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	716/823 (87%)	715 (100%)	1 (0%)	93 97
2	B	101/184 (55%)	101 (100%)	0	100 100
3	C	59/77 (77%)	59 (100%)	0	100 100
All	All	876/1084 (81%)	875 (100%)	1 (0%)	93 97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	714	LYS

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	P	9/32 (28%)	0	0
5	T	10/33 (30%)	0	0
All	All	19/65 (29%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue 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	$\#\|Z\| > 2$
5	7OK	T	9	5	15,22,23	5.55	5 (33%)	14,31,34	1.05	1 (7%)

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
5	7OK	T	9	5	-	2/5/27/28	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	T	9	7OK	O4'-C1'	15.73	1.63	1.41
5	T	9	7OK	C2'-C1'	-12.36	1.35	1.53
5	T	9	7OK	O4'-C4'	-6.55	1.30	1.45
5	T	9	7OK	C2-N3	-2.41	1.33	1.38
5	T	9	7OK	C3'-C4'	2.06	1.58	1.53

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	T	9	7OK	C5-C4-N3	-2.68	120.08	122.88

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	T	9	7OK	O4'-C4'-C5'-O5'
5	T	9	7OK	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	T	9	7OK	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.