



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2023 – 06:31 AM EDT

PDB ID : 6VYO  
Title : Crystal structure of RNA binding domain of nucleocapsid phosphoprotein from SARS coronavirus 2  
Authors : Chang, C.; Michalska, K.; Jedrzejczak, R.; Maltseva, N.; Endres, M.; Godzik, A.; Kim, Y.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2020-02-27  
Resolution : 1.70 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](mailto: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 : **FAILED**  
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

## 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 1.70 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4683 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 Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	124	974	618	175	180	1	0	1	0
1	B	125	981	621	177	182	1	0	1	0
1	C	125	1002	633	180	188	1	0	4	0
1	D	125	982	622	176	183	1	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	GLY	-	expression tag	UNP P0DTC9
B	46	GLY	-	expression tag	UNP P0DTC9
C	46	GLY	-	expression tag	UNP P0DTC9
D	46	GLY	-	expression tag	UNP P0DTC9

- Molecule 2 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
3	A	2	Total	Cl	0	0
			2	2		
3	C	2	Total	Cl	0	0
			2	2		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	168	Total O 168 168	0	0
6	B	157	Total O 157 157	0	0
6	C	162	Total O 162 162	0	0
6	D	177	Total O 177 177	0	0

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

### 3 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.24Å 77.43Å 114.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.05 – 1.70	Depositor
% Data completeness (in resolution range)	87.3 (46.05-1.70)	Depositor
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.93 (at 1.70Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.159 , 0.205	Depositor
Wilson B-factor (Å <sup>2</sup> )	13.7	Xtrriage
Anisotropy	0.248	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	4683	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.2808e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

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

### 4.2 Too-close contacts [i](#)

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

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

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

#### 4.3.2 Protein sidechains [i](#)

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

#### 4.3.3 RNA [i](#)

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

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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
2	MES	A	201	-	12,12,12	2.28	1 (8%)	14,16,16	2.23	7 (50%)
2	MES	C	204	-	12,12,12	2.19	1 (8%)	14,16,16	2.20	6 (42%)
5	GOL	C	203	-	5,5,5	0.96	0	5,5,5	0.88	0
5	GOL	D	202	-	5,5,5	0.88	0	5,5,5	0.96	0
5	GOL	B	201	-	5,5,5	0.83	0	5,5,5	1.03	0
2	MES	A	203	-	12,12,12	2.17	1 (8%)	14,16,16	1.89	3 (21%)
5	GOL	D	201	-	5,5,5	0.88	0	5,5,5	1.02	0
2	MES	B	202	-	12,12,12	2.27	1 (8%)	14,16,16	1.88	3 (21%)

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	MES	A	201	-	-	3/6/14/14	0/1/1/1
2	MES	C	204	-	-	5/6/14/14	0/1/1/1
5	GOL	C	203	-	-	4/4/4/4	-
5	GOL	D	202	-	-	0/4/4/4	-
5	GOL	B	201	-	-	2/4/4/4	-
2	MES	A	203	-	-	4/6/14/14	0/1/1/1
5	GOL	D	201	-	-	3/4/4/4	-
2	MES	B	202	-	-	4/6/14/14	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	202	MES	C8-S	-7.62	1.66	1.77
2	A	201	MES	C8-S	-7.61	1.66	1.77
2	C	204	MES	C8-S	-7.33	1.67	1.77
2	A	203	MES	C8-S	-7.27	1.67	1.77

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	201	MES	C5-N4-C3	4.98	120.03	108.83

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	202	MES	C5-N4-C3	4.67	119.35	108.83
2	A	203	MES	C5-N4-C3	4.47	118.90	108.83
2	C	204	MES	C5-N4-C3	4.39	118.71	108.83
2	C	204	MES	C7-N4-C3	3.21	119.43	111.23
2	A	203	MES	O3S-S-C8	3.03	110.67	105.77
2	C	204	MES	O2S-S-C8	3.01	110.54	106.92
2	A	201	MES	C7-N4-C5	2.82	118.45	111.23
2	A	201	MES	C7-N4-C3	2.58	117.82	111.23
2	C	204	MES	O3S-S-C8	2.53	109.86	105.77
2	A	201	MES	O1S-S-C8	2.50	109.92	106.92
2	C	204	MES	C6-C5-N4	-2.49	106.33	110.10
2	A	201	MES	O3S-S-C8	2.45	109.72	105.77
2	C	204	MES	C7-N4-C5	2.40	117.37	111.23
2	B	202	MES	O3S-S-C8	2.40	109.64	105.77
2	A	201	MES	C2-C3-N4	-2.35	106.54	110.10
2	A	203	MES	O2S-S-C8	2.30	109.69	106.92
2	B	202	MES	O2S-S-C8	2.10	109.44	106.92
2	A	201	MES	C6-C5-N4	-2.10	106.93	110.10

There are no chirality outliers.

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	201	MES	C7-C8-S-O2S
2	A	201	MES	C7-C8-S-O3S
2	A	203	MES	C7-C8-S-O1S
2	A	203	MES	C7-C8-S-O3S
2	B	202	MES	C7-C8-S-O1S
2	B	202	MES	C7-C8-S-O3S
2	C	204	MES	C8-C7-N4-C3
2	C	204	MES	C7-C8-S-O2S
5	B	201	GOL	O1-C1-C2-C3
5	C	203	GOL	C1-C2-C3-O3
5	C	203	GOL	O2-C2-C3-O3
2	C	204	MES	C7-C8-S-O3S
5	B	201	GOL	O1-C1-C2-O2
2	B	202	MES	C8-C7-N4-C3
5	C	203	GOL	O1-C1-C2-O2
2	A	201	MES	C7-C8-S-O1S
2	A	203	MES	C7-C8-S-O2S
2	B	202	MES	C7-C8-S-O2S
2	C	204	MES	C7-C8-S-O1S

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	D	201	GOL	C1-C2-C3-O3
2	A	203	MES	C8-C7-N4-C3
2	C	204	MES	C8-C7-N4-C5
5	C	203	GOL	O1-C1-C2-C3
5	D	201	GOL	O1-C1-C2-C3
5	D	201	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data

### 5.1 Protein, DNA and RNA chains

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

### 5.2 Non-standard residues in protein, DNA, RNA chains

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

### 5.3 Carbohydrates

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

### 5.4 Ligands

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

### 5.5 Other polymers

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