



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2021 – 12:40 PM EST

PDB ID : 7N0R  
Title : Structure of the SARS-CoV-2 N protein RNA-binding domain bound to single-domain antibody C2  
Authors : Ye, Q.; Corbett, K.D.  
Deposited on : 2021-05-25  
Resolution : 1.42 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 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 : 2.23.2  
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) : 2.23.2

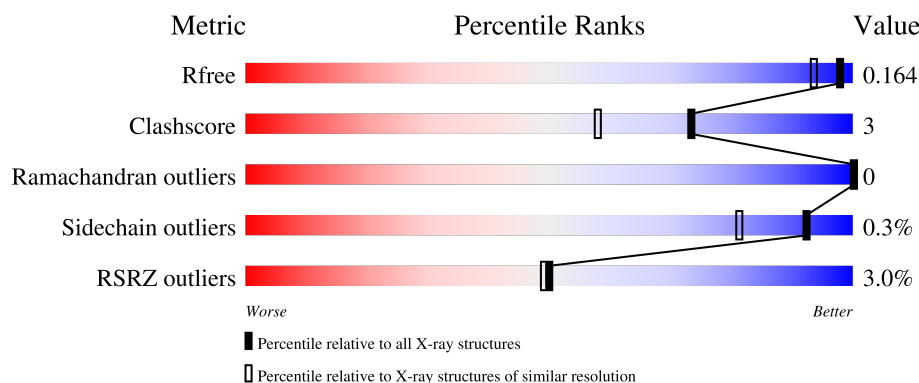
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.42 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	2579 (1.44-1.40)
Clashscore	141614	2696 (1.44-1.40)
Ramachandran outliers	138981	2632 (1.44-1.40)
Sidechain outliers	138945	2631 (1.44-1.40)
RSRZ outliers	127900	2528 (1.44-1.40)

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  $\geq 3$ , 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	129	<div> <div>2%</div> <div>95%</div> <div>5%</div> <div>0%</div> </div>
1	B	129	<div> <div>2%</div> <div>94%</div> <div>5%</div> <div>0%</div> </div>
2	C	134	<div> <div>5%</div> <div>84%</div> <div>7%</div> <div>8%</div> </div>
2	D	134	<div> <div>0%</div> <div>83%</div> <div>9%</div> <div>8%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8308 atoms, of which 3734 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
1	A	127	Total	C	H	N	O	S	0	2	0
			1960	629	965	179	186	1			
1	B	127	Total	C	H	N	O	S	0	1	0
			1962	630	967	178	185	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	SER	-	expression tag	UNP P0DTC9
A	47	ASN	-	expression tag	UNP P0DTC9
A	48	ALA	-	expression tag	UNP P0DTC9
B	46	SER	-	expression tag	UNP P0DTC9
B	47	ASN	-	expression tag	UNP P0DTC9
B	48	ALA	-	expression tag	UNP P0DTC9

- Molecule 2 is a protein called Single-domain antibody C2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	123	Total	C	H	N	O	S	0	6	0
			1841	584	901	164	185	7			
2	D	123	Total	C	H	N	O	S	0	6	0
			1841	584	901	164	185	7			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	201	Total	O	0	0
			201	201		
4	B	209	Total	O	0	0
			209	209		
4	C	122	Total	O	0	0
			122	122		
4	D	157	Total	O	0	0
			157	157		



- Molecule 1: Nucleoprotein



- 

- MET  
 ALA  
 E1  
 V2  
 Q3  
 L4  
 V12  
 R13  
 C22  
 A23  
 A24  
 T28  
 L45  
 L79  
 L86  
 Q105  
 Y110  
 W111  
 S121  
 A122  
 A123  
 ALA  
 LEU  
 GLU  
 HIS  
 HIS  
 HIS  
 HIS  
 HIS  
 HIS  
 HIS

- |      |
|------|
| MET  |
| E1   |
| R13  |
| R19  |
| Q39  |
| A40  |
| P41  |
| G44  |
| L45  |
| N77  |
| C96  |
| Q100 |
| E106 |
| X107 |
| E108 |
| A122 |
| A123 |
| ALA  |
| LEU  |
| GLU  |
| HIS  |
| HIS  |
| HIS  |
| HIS  |
| HIS  |
| HIS  |

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	46.67Å 77.12Å 71.93Å 90.00° 95.14° 90.00°	Depositor
Resolution (Å)	52.49 – 1.42 52.49 – 1.42	Depositor EDS
% Data completeness (in resolution range)	99.0 (52.49-1.42) 99.1 (52.49-1.42)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.59 (at 1.42Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, $R_{free}$	0.141 , 0.166 0.138 , 0.164	Depositor DCC
$R_{free}$ test set	4631 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.1	Xtriage
Anisotropy	0.529	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 43.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	8308	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.37% of the height of the origin peak. No significant pseudotranslation is detected.*

<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.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.42	0/1032	0.67	0/1401
1	B	0.42	0/1023	0.65	0/1389
2	C	0.48	2/974 (0.2%)	0.67	0/1317
2	D	0.47	0/974	0.73	0/1317
All	All	0.45	2/4003 (0.0%)	0.68	0/5424

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	22[A]	CYS	CB-SG	-5.20	1.73	1.81
2	C	22[B]	CYS	CB-SG	-5.20	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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

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	995	965	951	3	0
1	B	995	967	966	4	0
2	C	940	901	885	8	1
2	D	940	901	885	10	0
3	A	5	0	0	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	10	0	0	0	0
4	A	201	0	0	1	0
4	B	209	0	0	2	1
4	C	122	0	0	1	1
4	D	157	0	0	6	2
All	All	4574	3734	3687	25	3

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 3.

All (25) 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)
2:D:44:GLY:N	4:D:201:HOH:O	2.04	0.85
2:C:110:TYR:O	4:C:201:HOH:O	1.99	0.81
1:B:48:ALA:N	4:B:302:HOH:O	2.21	0.74
2:D:106:GLU:OE1	4:D:202:HOH:O	2.10	0.69
2:D:77:ASN:OD1	4:D:203:HOH:O	2.16	0.63
2:C:22[B]:CYS:HB3	2:C:79:LEU:HB3	1.89	0.54
1:B:98:ASP:OD2	4:B:301:HOH:O	2.19	0.52
2:D:44:GLY:HA2	4:D:310:HOH:O	2.11	0.50
2:D:13:ARG:HG2	4:D:304:HOH:O	2.14	0.48
1:B:95:ARG:HG2	1:B:101[A]:MET:SD	2.54	0.47
2:D:100:GLN:NE2	2:D:108:GLU:OE2	2.47	0.47
2:D:19:ARG:HD3	4:D:319:HOH:O	2.15	0.47
2:C:22[A]:CYS:HB3	2:C:79:LEU:HB3	1.96	0.46
1:A:95:ARG:HG3	1:A:101:MET:SD	2.56	0.46
2:D:39:GLN:HB2	2:D:45:LEU:HD12	1.98	0.46
1:A:107:ARG:HD3	4:A:382:HOH:O	2.16	0.45
2:C:45:LEU:HD23	2:C:105:GLN:CG	2.47	0.45
2:D:40:ALA:HB1	2:D:41:PRO:HD2	1.99	0.44
2:C:4:LEU:HD23	2:C:24:ALA:HA	1.99	0.44
2:C:12:VAL:HG21	2:C:86:LEU:HD13	2.00	0.44
2:C:45:LEU:HD21	2:C:111:TRP:CH2	2.53	0.43
1:B:93:ARG:HG2	1:B:103:ASP:OD1	2.19	0.42
2:D:39:GLN:HB2	2:D:45:LEU:CD1	2.49	0.42
2:C:45:LEU:HD21	2:C:111:TRP:HH2	1.86	0.41
1:A:55:ALA:HB2	1:A:109:TYR:CE2	2.57	0.40

All (3) 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)
4:C:289:HOH:O	4:D:209:HOH:O[1_556]	1.69	0.51
4:B:423:HOH:O	4:D:341:HOH:O[1_455]	1.97	0.23
2:C:28:THR:H	3:A:201:SO4:O3[1_655]	1.53	0.07

## 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	127/129 (98%)	124 (98%)	3 (2%)	0	100	100
1	B	126/129 (98%)	125 (99%)	1 (1%)	0	100	100
2	C	127/134 (95%)	125 (98%)	2 (2%)	0	100	100
2	D	127/134 (95%)	125 (98%)	2 (2%)	0	100	100
All	All	507/526 (96%)	499 (98%)	8 (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	101/101 (100%)	101 (100%)	0	100	100
1	B	100/101 (99%)	100 (100%)	0	100	100
2	C	97/103 (94%)	97 (100%)	0	100	100
2	D	97/103 (94%)	95 (98%)	2 (2%)	53	20
All	All	395/408 (97%)	393 (100%)	2 (0%)	92	73

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

Mol	Chain	Res	Type
2	D	96[A]	CYS
2	D	96[B]	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	153	ASN
1	A	154	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](#)

3 ligands are 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	$\# Z  > 2$
3	SO4	A	201	-	4,4,4	0.17	0	6,6,6	0.09	0
3	SO4	B	202	-	4,4,4	0.16	0	6,6,6	0.05	0
3	SO4	B	201	-	4,4,4	0.13	0	6,6,6	0.15	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	201	SO4	0	1

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	127/129 (98%)	-0.13	3 (2%) 59 58	11, 18, 35, 44	0
1	B	127/129 (98%)	-0.17	3 (2%) 59 58	11, 18, 33, 47	0
2	C	123/134 (91%)	0.09	7 (5%) 23 21	12, 20, 41, 50	0
2	D	123/134 (91%)	-0.07	2 (1%) 72 71	11, 19, 33, 43	0
All	All	500/526 (95%)	-0.07	15 (3%) 50 49	11, 19, 37, 50	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	1	GLU	4.5
2	C	2	VAL	3.9
2	C	122	ALA	3.8
2	C	123	ALA	3.3
2	C	110	TYR	3.3
1	B	48	ALA	3.3
1	B	49	THR	3.0
1	A	48	ALA	2.7
1	A	49	THR	2.6
2	C	121	SER	2.4
1	A	152	ALA	2.4
2	D	1	GLU	2.3
2	D	122	ALA	2.1
2	C	13	ARG	2.1
1	B	68	ARG	2.0

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

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	B	202	5/5	0.87	0.14	70,71,71,71	0
3	SO4	A	201	5/5	0.92	0.11	67,67,67,67	0
3	SO4	B	201	5/5	0.94	0.09	40,41,43,43	0

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

There are no such residues in this entry.