



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 09:26 AM UTC

PDB ID : 9F83 / pdb\_00009f83  
Title : SARS-CoV-2 Nucleocapsid N-terminal domain (NTD) mutant D63G  
Authors : Dhamotharan, K.; Schlundt, A.; Guenther, S.  
Deposited on : 2024-05-06  
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:

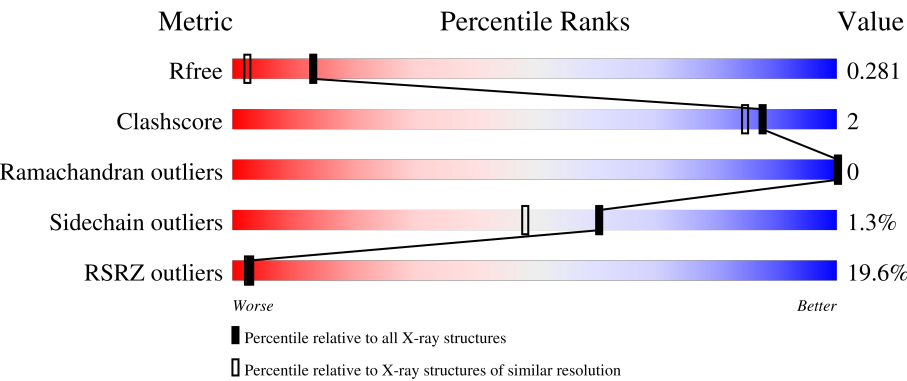
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 1.70 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))
$R_{free}$	180053	5551 (1.70-1.70)
Clashscore	190562	5924 (1.70-1.70)
Ramachandran outliers	187476	5846 (1.70-1.70)
Sidechain outliers	187428	5846 (1.70-1.70)
RSRZ outliers	180081	5554 (1.70-1.70)

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	136	<div><div>12%</div><div><div></div><div>85%</div><div>5%</div><div>10%</div></div></div>
1	B	136	<div><div>7%</div><div><div></div><div>84%</div><div>• •</div><div>10%</div></div></div>
1	C	136	<div><div>43%</div><div><div></div><div>75%</div><div>12%</div><div>•</div><div>12%</div></div></div>
1	D	136	<div><div>9%</div><div><div></div><div>86%</div><div>5%</div><div>•</div><div>8%</div></div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8009 atoms, of which 3716 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	122	Total	C	H	N	O	S	24	0	0
			1872	599	928	172	172	1			
1	B	123	Total	C	H	N	O	S	25	0	0
			1893	608	937	173	174	1			
1	C	119	Total	C	H	N	O	S	24	0	0
			1823	582	903	168	169	1			
1	D	125	Total	C	H	N	O	S	25	0	0
			1919	616	948	175	179	1			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	GLY	-	expression tag	UNP P0DTC9
A	40	SER	-	expression tag	UNP P0DTC9
A	63	GLY	ASP	variant	UNP P0DTC9
B	39	GLY	-	expression tag	UNP P0DTC9
B	40	SER	-	expression tag	UNP P0DTC9
B	63	GLY	ASP	variant	UNP P0DTC9
C	39	GLY	-	expression tag	UNP P0DTC9
C	40	SER	-	expression tag	UNP P0DTC9
C	63	GLY	ASP	variant	UNP P0DTC9
D	39	GLY	-	expression tag	UNP P0DTC9
D	40	SER	-	expression tag	UNP P0DTC9
D	63	GLY	ASP	variant	UNP P0DTC9

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	135	Total	O	0	0
			135	135		
2	B	141	Total	O	0	0
			141	141		

*Continued on next page...*

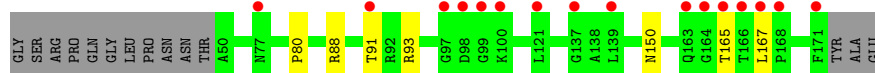
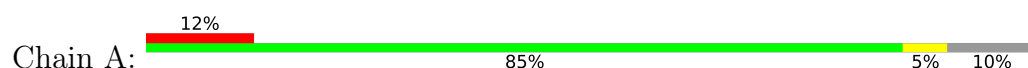
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	89	Total 89	O 89	0	0
2	D	137	Total 137	O 137	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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: Nucleoprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.01Å 55.66Å 84.89Å 90.00° 95.15° 90.00°	Depositor
Resolution (Å)	50.48 – 1.70 50.48 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.4 (50.48-1.70) 96.4 (50.48-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0425, BUSTER 2.10.4	Depositor
R, $R_{free}$	0.229 , 0.280 0.230 , 0.281	Depositor DCC
$R_{free}$ test set	2859 reflections (4.72%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.6	Xtriage
Anisotropy	0.411	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 48.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8009	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.66 % 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 1.4933e-03. 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.

## 5 Model quality

### 5.1 Standard geometry

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	0.96	0/971	1.23	1/1319 (0.1%)
1	B	0.96	0/984	1.32	9/1337 (0.7%)
1	C	0.83	0/946	1.27	3/1287 (0.2%)
1	D	1.01	1/999 (0.1%)	1.28	2/1356 (0.1%)
All	All	0.94	1/3900 (0.0%)	1.28	15/5299 (0.3%)

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	A	0	1
1	B	0	1
1	C	0	3
1	D	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	145	HIS	CG-CD2	-6.73	1.28	1.35

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	89	ARG	CD-NE-CZ	10.43	139.01	124.40
1	B	144	ASP	CA-CB-CG	9.24	121.84	112.60
1	C	144	ASP	CA-CB-CG	7.52	120.12	112.60
1	B	172	TYR	CA-C-O	-7.30	108.39	120.80
1	B	168	PRO	N-CA-CB	6.37	109.37	103.39
1	D	144	ASP	CA-CB-CG	6.25	118.85	112.60
1	A	150	ASN	OD1-CG-ND2	-6.05	116.55	122.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	89	ARG	NE-CZ-NH1	5.72	127.22	121.50
1	C	167	LEU	O-C-N	-5.66	117.06	121.14
1	B	169	LYS	CB-CA-C	5.32	118.75	109.65
1	B	89	ARG	NE-CZ-NH2	-5.17	114.55	119.20
1	C	113	LEU	CA-C-O	-5.13	115.36	120.96
1	B	145	HIS	CB-CG-CD2	5.13	137.86	131.20
1	B	145	HIS	CB-CG-ND1	-5.05	115.12	122.70
1	D	136	GLU	CB-CG-CD	5.03	121.16	112.60

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	88	ARG	Sidechain
1	B	89	ARG	Sidechain
1	C	163	GLN	Peptide
1	C	88	ARG	Sidechain
1	C	93	ARG	Sidechain
1	D	92	ARG	Sidechain

## 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	944	928	925	3	0
1	B	956	937	934	3	0
1	C	920	903	900	6	0
1	D	971	948	945	3	0
2	A	135	0	0	1	0
2	B	141	0	0	0	0
2	C	89	0	0	2	0
2	D	137	0	0	1	0
All	All	4293	3716	3704	15	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 2.

All (15) 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:58:GLN:NE2	1:C:62:GLU:O	2.36	0.58
1:C:97:GLY:N	2:C:202:HOH:O	2.36	0.54
1:D:160:GLN:HE21	1:D:160:GLN:HA	1.73	0.53
1:D:164:GLY:N	2:D:201:HOH:O	2.39	0.48
1:B:93:ARG:HH11	1:B:93:ARG:HG2	1.80	0.47
1:C:94:ILE:CG2	1:C:104:LEU:HD11	2.46	0.46
1:B:89:ARG:HD3	1:B:108:TRP:CZ2	2.51	0.45
1:A:165:THR:O	1:A:167:LEU:HD12	2.18	0.43
1:B:62:GLU:HG2	1:B:168:PRO:HB3	2.00	0.43
1:A:93:ARG:HH11	1:A:93:ARG:HG2	1.82	0.43
1:C:67:PRO:HG2	1:C:70:GLN:OE1	2.19	0.42
1:D:93:ARG:HG2	1:D:103:ASP:OD1	2.18	0.42
1:C:74:ILE:HD11	1:C:162:PRO:HA	2.03	0.41
1:A:91:THR:HG21	2:A:318:HOH:O	2.20	0.41
1:C:129:GLY:N	2:C:204:HOH:O	2.54	0.40

There are no symmetry-related clashes.

## 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	120/136 (88%)	117 (98%)	3 (2%)	0	100	100
1	B	121/136 (89%)	118 (98%)	3 (2%)	0	100	100
1	C	117/136 (86%)	113 (97%)	4 (3%)	0	100	100
1	D	123/136 (90%)	119 (97%)	4 (3%)	0	100	100
All	All	481/544 (88%)	467 (97%)	14 (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 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	95/106 (90%)	94 (99%)	1 (1%)	65	54
1	B	96/106 (91%)	95 (99%)	1 (1%)	68	58
1	C	93/106 (88%)	91 (98%)	2 (2%)	45	29
1	D	97/106 (92%)	96 (99%)	1 (1%)	68	58
All	All	381/424 (90%)	376 (99%)	5 (1%)	61	48

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	80	PRO
1	B	144	ASP
1	C	64	LEU
1	C	163	GLN
1	D	160	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	59	HIS
1	C	83	GLN
1	C	145	HIS
1	D	59	HIS
1	D	145	HIS
1	D	150	ASN
1	D	153	ASN
1	D	160	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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	122/136 (89%)	0.77	16 (13%) 7 7	16, 32, 74, 93	0
1	B	123/136 (90%)	0.50	9 (7%) 21 23	16, 27, 65, 95	0
1	C	119/136 (87%)	2.03	59 (49%) 0 0	22, 45, 118, 164	0
1	D	125/136 (91%)	0.59	12 (9%) 13 14	15, 28, 84, 151	0
All	All	489/544 (89%)	0.96	96 (19%) 3 2	15, 32, 93, 164	0

All (96) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	64	LEU	6.8
1	C	108	TRP	6.4
1	B	172	TYR	5.8
1	C	168	PRO	5.6
1	A	166	THR	5.4
1	C	107	ARG	5.1
1	D	97	GLY	4.9
1	B	166	THR	4.7
1	C	94	ILE	4.5
1	C	104	LEU	4.4
1	A	99	GLY	4.4
1	A	168	PRO	4.3
1	C	56	LEU	4.3
1	C	123	TYR	4.2
1	C	59	HIS	4.2
1	A	164	GLY	4.2
1	C	58	GLN	4.1
1	C	167	LEU	4.1
1	C	129	GLY	4.1
1	C	137	GLY	4.0
1	B	97	GLY	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	131	ILE	3.9
1	C	93	ARG	3.8
1	C	92	ARG	3.8
1	C	106	PRO	3.7
1	C	99	GLY	3.7
1	C	103	ASP	3.7
1	A	139	LEU	3.7
1	C	132	TRP	3.7
1	A	97	GLY	3.7
1	C	165	THR	3.6
1	C	67	PRO	3.5
1	D	94	ILE	3.5
1	C	60	GLY	3.5
1	C	134	ALA	3.4
1	C	161	LEU	3.4
1	C	69	GLY	3.3
1	C	101	MET	3.1
1	C	66	PHE	3.0
1	C	125	ALA	3.0
1	B	164	GLY	3.0
1	D	99	GLY	3.0
1	C	57	THR	2.9
1	D	152	ALA	2.9
1	C	133	VAL	2.9
1	C	166	THR	2.9
1	C	62	GLU	2.8
1	C	124	GLY	2.8
1	D	137	GLY	2.8
1	C	63	GLY	2.7
1	C	55	ALA	2.7
1	A	167	LEU	2.7
1	A	165	THR	2.7
1	B	100	LYS	2.7
1	C	95	ARG	2.7
1	D	96	GLY	2.6
1	C	91	THR	2.6
1	D	68	ARG	2.6
1	C	136	GLU	2.6
1	B	99	GLY	2.6
1	D	136	GLU	2.5
1	C	141	THR	2.5
1	D	61	LYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	95	ARG	2.5
1	B	144	ASP	2.5
1	A	100	LYS	2.5
1	C	61	LYS	2.5
1	C	71	GLY	2.5
1	C	130	ILE	2.5
1	A	163	GLN	2.4
1	C	97	GLY	2.4
1	C	50	ALA	2.4
1	C	162	PRO	2.4
1	A	98	ASP	2.4
1	C	100	LYS	2.4
1	C	90	ALA	2.3
1	B	163	GLN	2.3
1	A	171	PHE	2.3
1	A	137	GLY	2.3
1	A	77	ASN	2.2
1	B	167	LEU	2.2
1	C	163	GLN	2.2
1	C	109	TYR	2.2
1	C	128	ASP	2.2
1	C	87	TYR	2.2
1	C	74	ILE	2.2
1	C	160	GLN	2.1
1	C	96	GLY	2.1
1	C	102	LYS	2.1
1	A	121	LEU	2.1
1	D	101	MET	2.1
1	C	98	ASP	2.1
1	D	93	ARG	2.1
1	C	152	ALA	2.1
1	C	68	ARG	2.0
1	A	91	THR	2.0

## 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 oligosaccharides in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.