



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 03:42 am BST

PDB ID : 3T5N  
Title : 1.8A crystal structure of Lassa virus nucleoprotein in complex with ssRNA  
Authors : Hastie, K.M.; Liu, T.; King, L.B.; Ngo, N.; Zandonatti, M.A.; Woods, V.L.;  
de la Torre, J.C.; Saphire, E.O.  
Deposited on : 2011-07-27  
Resolution : 1.79 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

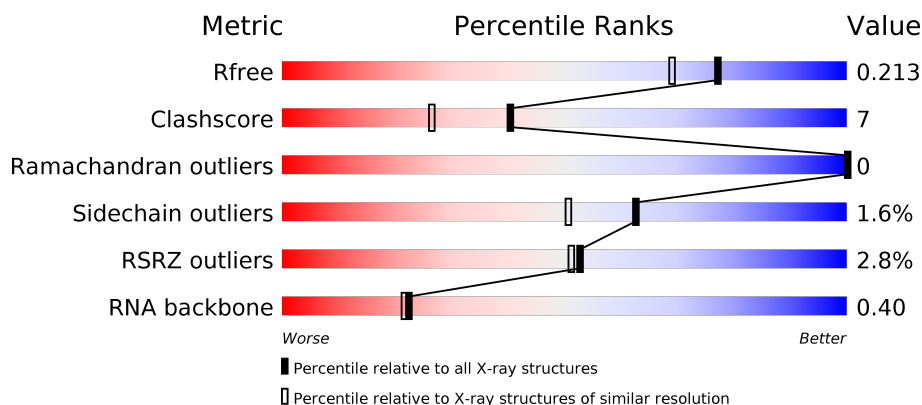
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.79 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)
RNA backbone	3102	1069 (2.40-1.16)

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 on 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	354	 2% 69% 9% 20%
2	C	6	 17% 33% 50% 17%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2492 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
1	A	282	Total	C	N	O	S	0	1	0
			2184	1371	380	422	11			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	EXPRESSION TAG	UNP Q5S585
A	-12	ALA	-	EXPRESSION TAG	UNP Q5S585
A	-11	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-10	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-9	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-8	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-7	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-6	HIS	-	EXPRESSION TAG	UNP Q5S585
A	-5	VAL	-	EXPRESSION TAG	UNP Q5S585
A	-4	ASP	-	EXPRESSION TAG	UNP Q5S585
A	-3	ASP	-	EXPRESSION TAG	UNP Q5S585
A	-2	ASP	-	EXPRESSION TAG	UNP Q5S585
A	-1	ASP	-	EXPRESSION TAG	UNP Q5S585
A	0	LYS	-	EXPRESSION TAG	UNP Q5S585

- Molecule 2 is a RNA chain called RNA (5'-R(P\*UP\*AP\*UP\*CP\*UP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	P	0	0	0
			122	55	17	44	6			

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Ni	0	0
			2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	180	Total 180	O 180	0	0
4	C	4	Total 4	O 4	0	0



- Molecule 1: Nucleoprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.16Å 71.74Å 56.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.00 – 1.79 45.14 – 1.79	Depositor EDS
% Data completeness (in resolution range)	99.7 (37.00-1.79) 99.7 (45.14-1.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 1.78Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, $R_{free}$	0.182 , 0.210 0.183 , 0.213	Depositor DCC
$R_{free}$ test set	1491 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.2	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 49.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2492	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.26% 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: NI

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	0.35	0/2214	0.53	1/2992 (0.0%)
2	C	2.06	9/134 (6.7%)	1.68	4/205 (2.0%)
All	All	0.60	9/2348 (0.4%)	0.66	5/3197 (0.2%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	4	U	P-OP1	-7.74	1.35	1.49
2	C	4	U	P-OP2	-7.39	1.36	1.49
2	C	4	U	N1-C2	-6.70	1.32	1.38
2	C	4	U	O5'-C5'	-5.90	1.33	1.42
2	C	5	C	P-OP1	-5.75	1.39	1.49
2	C	5	C	O5'-C5'	-5.54	1.33	1.42
2	C	5	C	N1-C6	-5.52	1.33	1.37
2	C	4	U	P-O5'	-5.39	1.54	1.59
2	C	5	C	P-OP2	-5.21	1.40	1.49

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	4	U	N1-C2-O2	-8.58	116.79	122.80
2	C	4	U	C2-N3-C4	-7.44	122.54	127.00
2	C	4	U	N1-C2-N3	5.89	118.44	114.90
1	A	264	MET	CG-SD-CE	-5.18	91.91	100.20
2	C	5	C	O4'-C1'-N1	5.02	112.22	108.20

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	2184	0	2227	31	0
2	C	122	0	64	4	0
3	A	2	0	0	0	0
4	A	180	0	0	8	0
4	C	4	0	0	0	0
All	All	2492	0	2291	33	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 7.

All (33) 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:51:GLN:O	1:A:55:ARG:HG3	1.83	0.79
1:A:61:ASP:OD1	1:A:166:VAL:HG13	1.85	0.76
1:A:269:LYS:NZ	4:A:484:HOH:O	2.19	0.76
1:A:70:LEU:HD13	1:A:244:TYR:CE1	2.25	0.71
2:C:5:C:C5	2:C:6:U:C4	2.83	0.65
1:A:40:HIS:HE1	4:A:465:HOH:O	1.81	0.64
1:A:70:LEU:HD13	1:A:244:TYR:CZ	2.33	0.63
2:C:5:C:C4	2:C:6:U:C4	2.89	0.60
1:A:18:ARG:NH1	4:A:491:HOH:O	2.35	0.60
1:A:310:ILE:HD13	1:A:318:PRO:HG2	1.84	0.58
1:A:216:THR:HG23	1:A:234:THR:HG21	1.86	0.57
1:A:223:THR:HG21	1:A:230:ASN:OD1	2.05	0.56
1:A:52:ARG:NE	4:A:512:HOH:O	2.15	0.55
1:A:52:ARG:NH2	4:A:512:HOH:O	2.39	0.51
1:A:8:LYS:N	4:A:461:HOH:O	2.43	0.51
1:A:55:ARG:NH2	1:A:236:LYS:O	2.43	0.51
1:A:269:LYS:HD3	1:A:317:TRP:CE2	2.46	0.50
1:A:164:TRP:HE1	2:C:2:U:H5	1.61	0.49
1:A:48:SER:O	4:A:490:HOH:O	2.19	0.49
1:A:256:ALA:O	1:A:264:MET:HG3	2.14	0.47
1:A:240:ASN:ND2	2:C:2:U:O2'	2.48	0.47
1:A:166:VAL:HG12	1:A:168:ASN:H	1.81	0.46

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:ASP:OD1	1:A:235:LYS:HB2	2.14	0.46
1:A:18:ARG:NH1	4:A:510:HOH:O	2.49	0.45
1:A:222:LEU:HA	1:A:222:LEU:HD23	1.84	0.45
1:A:269:LYS:HD2	1:A:316:GLY:O	2.16	0.45
1:A:15:SER:CB	1:A:289:LEU:HD11	2.47	0.44
1:A:216:THR:CG2	1:A:234:THR:HG21	2.48	0.42
1:A:204:ASP:HB3	1:A:222:LEU:HD21	2.02	0.42
1:A:70:LEU:HD23	1:A:70:LEU:HA	1.93	0.41
1:A:317:TRP:HB2	1:A:321:ALA:HB2	2.03	0.41
1:A:61:ASP:CG	1:A:166:VAL:HG13	2.41	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	279/354 (79%)	275 (99%)	4 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 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	244/306 (80%)	240 (98%)	4 (2%)	62	51

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70	LEU
1	A	72	GLN
1	A	179	MET
1	A	234	THR

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

Mol	Chain	Res	Type
1	A	40	HIS
1	A	240	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	5/6 (83%)	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](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates 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.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	282/354 (79%)	0.19	7 (2%) 57 56	14, 31, 61, 86	0
2	C	6/6 (100%)	0.91	1 (16%) 1 1	21, 41, 68, 80	0
All	All	288/360 (80%)	0.21	8 (2%) 53 51	14, 31, 65, 86	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	234	THR	2.9
2	C	7	C	2.8
1	A	295	ASP	2.5
1	A	339	GLU	2.4
1	A	167	LYS	2.2
1	A	163	VAL	2.1
1	A	192	GLN	2.1
1	A	237	SER	2.1

### 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 carbohydrates 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	NI	A	342	1/1	0.82	0.07	72,72,72,72	0
3	NI	A	341	1/1	0.90	0.11	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.