wwPDB X-ray Structure Validation Summary Report (i)

Sep 2, 2023 – 07:19 PM EDT

PDB ID	:	3R3L
Title	:	Structure of NP protein from Lassa AV strain
Authors	:	Perbandt, M.; Brunotte, L.; Gunther, S.; Betzel, C.
Deposited on	:	2011-03-16
Resolution	:	2.45 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

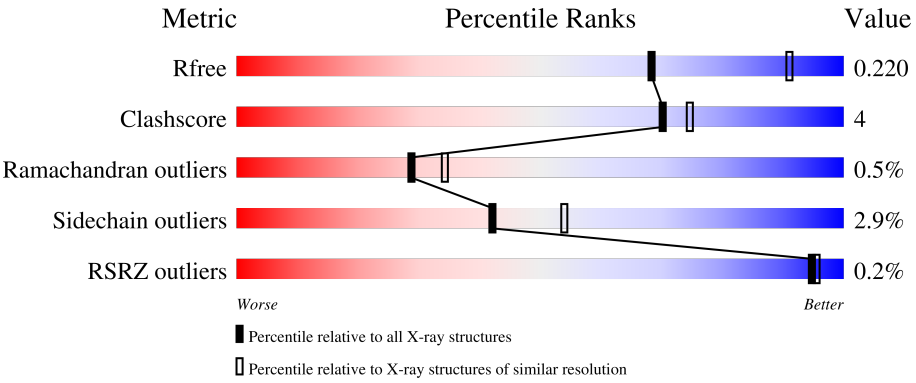
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.45 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)
RSRZ outliers	127900	1547 (2.46-2.42)

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 >=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% 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	582	81% 7% 12%
1	B	582	79% 9% 12%
1	C	582	71% 15% • 12%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12140 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	513	Total	C	N	O	S	0	0	0
			4009	2522	698	763	26			
1	B	514	Total	C	N	O	S	0	0	0
			4018	2528	700	764	26			
1	C	510	Total	C	N	O	S	0	0	0
			3986	2508	694	758	26			

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	570	ASP	-	expression tag	UNP Q9DQX7
A	571	TYR	-	expression tag	UNP Q9DQX7
A	572	LYS	-	expression tag	UNP Q9DQX7
A	573	ASP	-	expression tag	UNP Q9DQX7
A	574	HIS	-	expression tag	UNP Q9DQX7
A	575	ASP	-	expression tag	UNP Q9DQX7
A	576	GLY	-	expression tag	UNP Q9DQX7
A	577	HIS	-	expression tag	UNP Q9DQX7
A	578	HIS	-	expression tag	UNP Q9DQX7
A	579	HIS	-	expression tag	UNP Q9DQX7
A	580	HIS	-	expression tag	UNP Q9DQX7
A	581	HIS	-	expression tag	UNP Q9DQX7
A	582	HIS	-	expression tag	UNP Q9DQX7
B	570	ASP	-	expression tag	UNP Q9DQX7
B	571	TYR	-	expression tag	UNP Q9DQX7
B	572	LYS	-	expression tag	UNP Q9DQX7
B	573	ASP	-	expression tag	UNP Q9DQX7
B	574	HIS	-	expression tag	UNP Q9DQX7
B	575	ASP	-	expression tag	UNP Q9DQX7
B	576	GLY	-	expression tag	UNP Q9DQX7
B	577	HIS	-	expression tag	UNP Q9DQX7
B	578	HIS	-	expression tag	UNP Q9DQX7
B	579	HIS	-	expression tag	UNP Q9DQX7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	580	HIS	-	expression tag	UNP Q9DQX7
B	581	HIS	-	expression tag	UNP Q9DQX7
B	582	HIS	-	expression tag	UNP Q9DQX7
C	570	ASP	-	expression tag	UNP Q9DQX7
C	571	TYR	-	expression tag	UNP Q9DQX7
C	572	LYS	-	expression tag	UNP Q9DQX7
C	573	ASP	-	expression tag	UNP Q9DQX7
C	574	HIS	-	expression tag	UNP Q9DQX7
C	575	ASP	-	expression tag	UNP Q9DQX7
C	576	GLY	-	expression tag	UNP Q9DQX7
C	577	HIS	-	expression tag	UNP Q9DQX7
C	578	HIS	-	expression tag	UNP Q9DQX7
C	579	HIS	-	expression tag	UNP Q9DQX7
C	580	HIS	-	expression tag	UNP Q9DQX7
C	581	HIS	-	expression tag	UNP Q9DQX7
C	582	HIS	-	expression tag	UNP Q9DQX7

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	41	Total O 41 41	0	0

Continued on next page...

Continued from previous page...

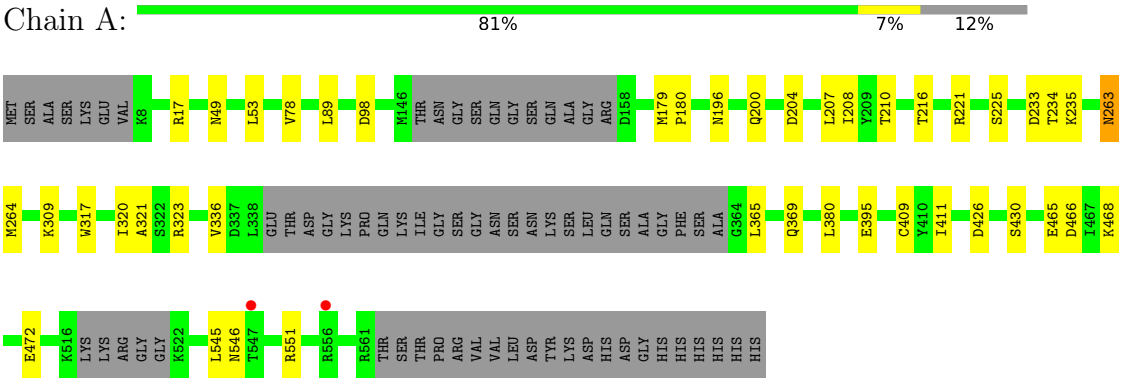
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	48	Total O 48 48	0	0
4	C	30	Total O 30 30	0	0

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 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.

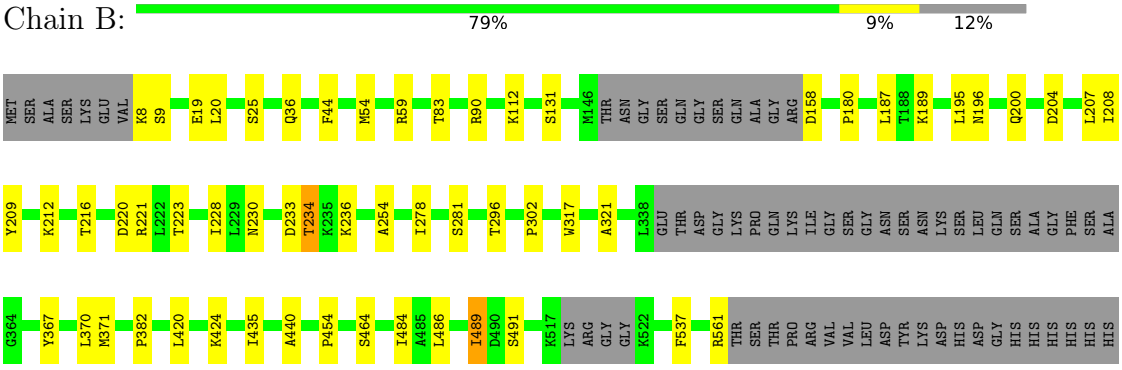
- Molecule 1: Nucleoprotein

Chain A:



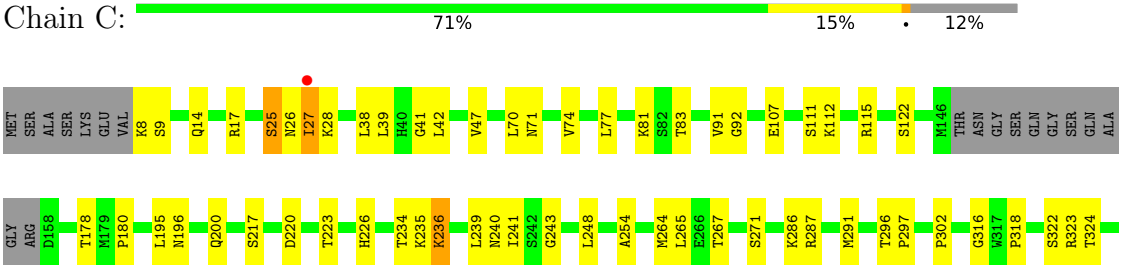
- Molecule 1: Nucleoprotein

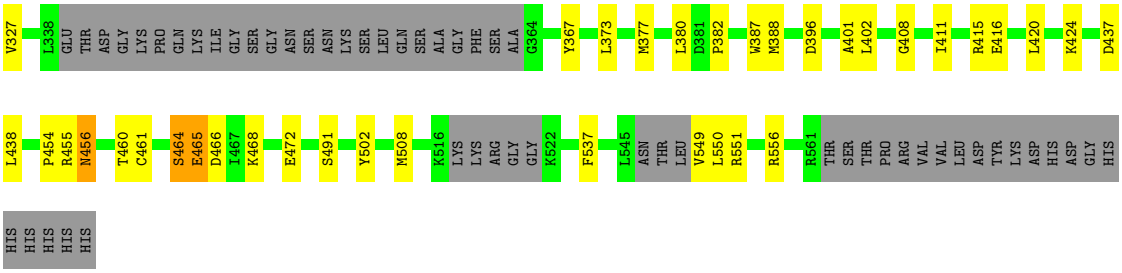
Chain B:



- Molecule 1: Nucleoprotein

Chain C:




V327
L338
GLU
THR
ASP
GLY
LYS
PRO
GLN
LYS
ILE
GLY
SER
GLY
ASN
SER
ASN
LYS
SER
LEU
GLN
SER
ALA
GLY
PHE
SER
ALA
G364
Y367
L373
M377
L380
D381
P382
W387
M388
D396
A401
L402
G408
I411
R415
E416
L420
K424
D437
L438
P454
R455
N456
T460
C461
S464
E465
D466
T467
K468
E472
S491
Y502
M508
K516
LYS
LYS
ARG
GLY
GLY
K522
F537
L545
ASN
THR
LEU
V549
L550
R551
R556
R561
THR
SER
THR
PRO
ARG
VAL
VAL
LEU
ASP
TYR
LYS
ASP
HIS
ASP
GLY
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	176.32Å 176.32Å 56.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.45 88.16 – 2.44	Depositor EDS
% Data completeness (in resolution range)	51.4 (30.00-2.45) 51.4 (88.16-2.44)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.184 , 0.219 0.182 , 0.220	Depositor DCC
R_{free} test set	1878 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	33.2	Xtriage
Anisotropy	0.331	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , -16.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.37$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	0.315 for -h,-k,l 0.499 for h,-h-k,-l 0.354 for -k,-h,-l	Xtriage
Reported twinning fraction	0.332 for H, K, L 0.163 for -H, H+K, -L 0.163 for -h,-k,l 0.342 for -H-K, K, -L	Depositor
Outliers	0 of 37572 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12140	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.72% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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

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	0.32	0/4066	0.51	0/5489
1	B	0.31	0/4075	0.50	0/5500
1	C	0.34	0/4042	0.52	0/5454
All	All	0.32	0/12183	0.51	0/16443

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	4009	0	4089	20	0
1	B	4018	0	4102	29	0
1	C	3986	0	4064	43	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	41	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	48	0	0	0	0
4	C	30	0	0	0	0
All	All	12140	0	12255	88	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 4.

The worst 5 of 88 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:216:THR:HA	1:A:234:THR:HG21	1.73	0.70
1:C:455:ARG:O	1:C:456:ASN:HB2	1.94	0.67
1:B:196:ASN:O	1:B:200:GLN:HG2	1.99	0.63
1:B:420:LEU:HD13	1:B:435:ILE:HD12	1.85	0.58
1:C:81:LYS:HA	1:C:322:SER:H	1.68	0.58

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 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	505/582 (87%)	486 (96%)	19 (4%)	0	100	100
1	B	506/582 (87%)	490 (97%)	16 (3%)	0	100	100
1	C	500/582 (86%)	465 (93%)	28 (6%)	7 (1%)	11	10
All	All	1511/1746 (86%)	1441 (95%)	63 (4%)	7 (0%)	29	34

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	91	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	464	SER
1	C	465	GLU
1	C	27	ILE
1	C	240	ASN

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	450/505 (89%)	441 (98%)	9 (2%)	55	67
1	B	451/505 (89%)	439 (97%)	12 (3%)	44	57
1	C	447/505 (88%)	429 (96%)	18 (4%)	31	41
All	All	1348/1515 (89%)	1309 (97%)	39 (3%)	42	54

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	236	LYS
1	C	461	CYS
1	C	264	MET
1	C	323	ARG
1	C	549	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	501	GLN
1	B	36	GLN
1	B	263	ASN
1	C	263	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 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

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^{th} 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^2)	Q<0.9
1	A	513/582 (88%)	-0.03	2 (0%) 92 92	20, 20, 20, 26	1 (0%)
1	B	514/582 (88%)	0.01	0 100 100	20, 20, 20, 20	0
1	C	510/582 (87%)	-0.15	1 (0%) 95 95	20, 20, 20, 30	0
All	All	1537/1746 (88%)	-0.06	3 (0%) 95 95	20, 20, 20, 30	1 (0%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	556	ARG	2.4
1	C	27	ILE	2.0
1	A	547	THR	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

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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(Å^2)	Q<0.9
2	MN	B	584	1/1	0.95	0.07	30,30,30,30	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MN	B	583	1/1	0.97	0.10	30,30,30,30	0
2	MN	A	583	1/1	0.98	0.06	30,30,30,30	0
2	MN	C	583	1/1	0.98	0.06	30,30,30,30	0
3	ZN	B	585	1/1	0.98	0.17	20,20,20,20	0
3	ZN	C	584	1/1	0.98	0.09	20,20,20,20	0
2	MN	A	584	1/1	0.99	0.09	30,30,30,30	0
3	ZN	A	585	1/1	1.00	0.16	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.