



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 02:40 pm BST

PDB ID : 5000  
Title : Structure of the Rift Valley fever virus NSs protein core domain  
Authors : Barski, M.S.; Potter, J.A.; Schwarz-Linek, U.  
Deposited on : 2017-08-08  
Resolution : 2.20 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

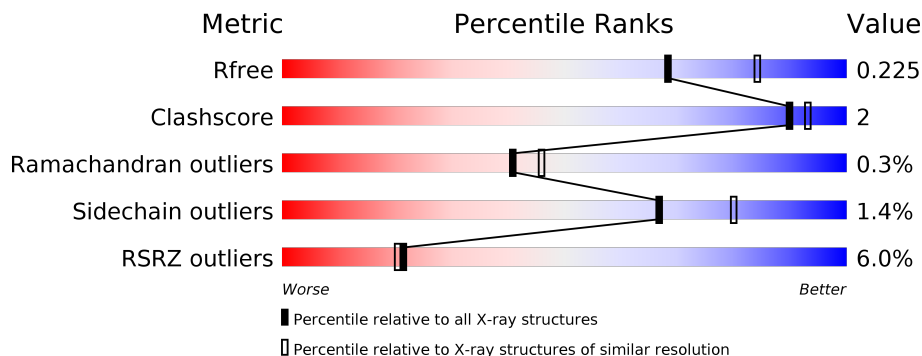
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	170	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1, green 92%, grey 5%);"></div> <div style="margin-left: 5px;"> <p>%</p> <p>92%</p> <p>5% ..</p> </div> </div>
1	B	170	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 10%, orange 1, yellow 5%, green 84%, grey 11%);"></div> <div style="margin-left: 5px;"> <p>10%</p> <p>84%</p> <p>5% • 11%</p> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2602 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 Non-structural protein NS-S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	166	1305	830	227	235	13	0	1	0
1	B	152	1194	760	211	212	11	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	GLY	-	expression tag	UNP A2SZX2
A	80	ALA	-	expression tag	UNP A2SZX2
A	81	MET	-	expression tag	UNP A2SZX2
A	82	VAL	-	expression tag	UNP A2SZX2
B	79	GLY	-	expression tag	UNP A2SZX2
B	80	ALA	-	expression tag	UNP A2SZX2
B	81	MET	-	expression tag	UNP A2SZX2
B	82	VAL	-	expression tag	UNP A2SZX2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	89	Total	O	0	0
			89	89		
2	B	14	Total	O	0	0
			14	14		

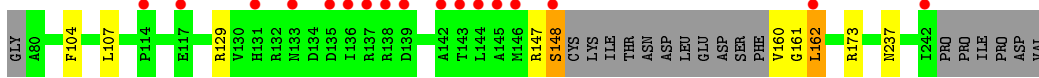
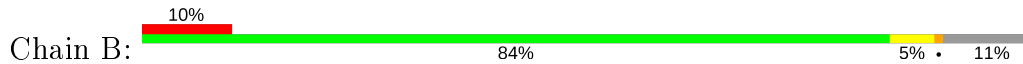
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Non-structural protein NS-S



- Molecule 1: Non-structural protein NS-S



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.50Å 124.50Å 174.26Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	107.82 – 2.20 58.62 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (107.82-2.20) 99.9 (58.62-2.20)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.97 (at 2.20Å)	Xtrriage
Refinement program	REFMAC 5.8.0107	Depositor
R, $R_{free}$	0.201 , 0.217 0.212 , 0.225	Depositor DCC
$R_{free}$ test set	2065 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	47.3	Xtrriage
Anisotropy	0.190	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2602	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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](#)

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	0.93	2/1333 (0.2%)	0.90	1/1806 (0.1%)
1	B	0.69	0/1215	0.90	0/1644
All	All	0.82	2/2548 (0.1%)	0.90	1/3450 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	126	SER	CB-OG	-6.11	1.34	1.42
1	A	228	SER	CB-OG	5.16	1.49	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	154	ASP	CB-CG-OD1	5.77	123.50	118.30

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	1305	0	1336	3	1
1	B	1194	0	1228	6	0
2	A	89	0	0	0	0
2	B	14	0	0	0	1
All	All	2602	0	2564	9	2

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 (9) 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:B:147:ARG:O	1:B:148:SER:OG	1.98	0.80
1:A:149:CYS:O	1:A:150:LYS:HB2	2.11	0.51
1:B:147:ARG:C	1:B:148:SER:OG	2.51	0.48
1:B:160:VAL:O	1:B:162:LEU:N	2.47	0.47
1:B:173:ARG:NH2	1:B:237:ASN:OD1	2.45	0.44
1:A:132:ARG:CD	1:A:137:ARG:HH11	2.32	0.42
1:B:129:ARG:HB3	1:B:129:ARG:HE	1.60	0.42
1:B:104:PHE:HB2	1:B:107:LEU:HB2	2.02	0.42
1:A:104:PHE:HB2	1:A:107:LEU:HB2	2.02	0.41

All (2) 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)
2:B:305:HOH:O	2:B:305:HOH:O[11_656]	2.03	0.17
1:A:150:LYS:NZ	1:A:150:LYS:NZ[10_556]	2.11	0.09

## 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	165/170 (97%)	161 (98%)	4 (2%)	0	100	100
1	B	148/170 (87%)	143 (97%)	4 (3%)	1 (1%)	22	22
All	All	313/340 (92%)	304 (97%)	8 (3%)	1 (0%)	41	46

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	161	GLY

### 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	147/150 (98%)	145 (99%)	2 (1%)	67 80
1	B	133/150 (89%)	131 (98%)	2 (2%)	65 78
All	All	280/300 (93%)	276 (99%)	4 (1%)	67 80

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

Mol	Chain	Res	Type
1	A	129	ARG
1	A	150	LYS
1	B	148	SER
1	B	162	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	HIS

### 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 carbohydrates 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 [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	166/170 (97%)	-0.02	2 (1%) 79 77	34, 44, 68, 91	0
1	B	152/170 (89%)	0.45	17 (11%) 5 4	52, 77, 122, 145	0
All	All	318/340 (93%)	0.21	19 (5%) 21 20	34, 59, 109, 145	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	145	ALA	7.5
1	B	146	MET	6.8
1	B	142	ALA	6.0
1	B	135	ASP	4.6
1	B	136	ILE	4.5
1	B	131	HIS	4.3
1	B	138	ARG	3.7
1	B	144	LEU	3.6
1	B	148	SER	3.5
1	B	242	ILE	3.4
1	A	244	PRO	3.2
1	B	133	ASN	3.0
1	B	143	THR	2.9
1	B	139	ASP	2.5
1	B	117	GLU	2.3
1	B	137	ARG	2.3
1	A	131	HIS	2.2
1	B	114	PRO	2.1
1	B	162	LEU	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

There are no carbohydrates in this entry.

### 6.4 Ligands

There are no ligands in this entry.

### 6.5 Other polymers

There are no such residues in this entry.