



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 16, 2023 – 08:40 AM JST

PDB ID : 6LD1  
Title : Zika NS5 polymerase domain  
Authors : El Sahili, A.; Lescar, J.  
Deposited on : 2019-11-20  
Resolution : 1.40 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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

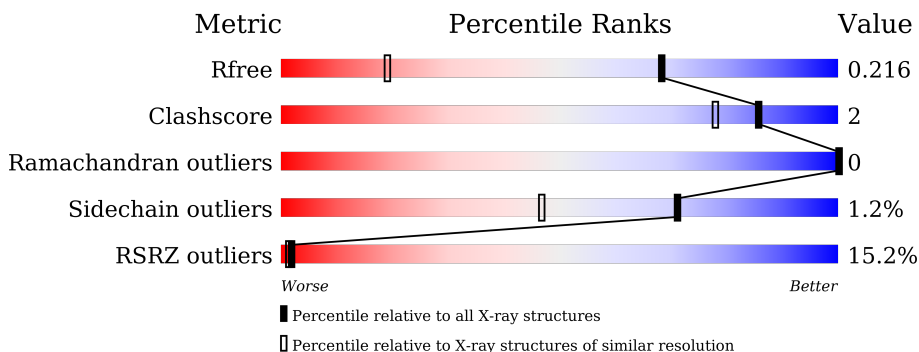
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	645	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 5309 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 RNA-directed RNA polymerase NS5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	574	4829	3048	881	867	33	0	15	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	247	MET	-	expression tag	UNP A0A024B7W1
A	248	HIS	-	expression tag	UNP A0A024B7W1
A	249	HIS	-	expression tag	UNP A0A024B7W1
A	250	HIS	-	expression tag	UNP A0A024B7W1
A	251	HIS	-	expression tag	UNP A0A024B7W1
A	252	HIS	-	expression tag	UNP A0A024B7W1
A	253	HIS	-	expression tag	UNP A0A024B7W1
A	254	SER	-	expression tag	UNP A0A024B7W1
A	255	SER	-	expression tag	UNP A0A024B7W1
A	256	GLY	-	expression tag	UNP A0A024B7W1
A	257	VAL	-	expression tag	UNP A0A024B7W1
A	258	ASP	-	expression tag	UNP A0A024B7W1
A	259	LEU	-	expression tag	UNP A0A024B7W1
A	260	GLY	-	expression tag	UNP A0A024B7W1
A	261	THR	-	expression tag	UNP A0A024B7W1
A	262	GLU	-	expression tag	UNP A0A024B7W1
A	263	ASN	-	expression tag	UNP A0A024B7W1
A	264	LEU	-	expression tag	UNP A0A024B7W1
A	265	TYR	-	expression tag	UNP A0A024B7W1
A	266	PHE	-	expression tag	UNP A0A024B7W1
A	267	GLN	-	expression tag	UNP A0A024B7W1
A	268	SER	-	expression tag	UNP A0A024B7W1
A	269	MET	-	expression tag	UNP A0A024B7W1

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0

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

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

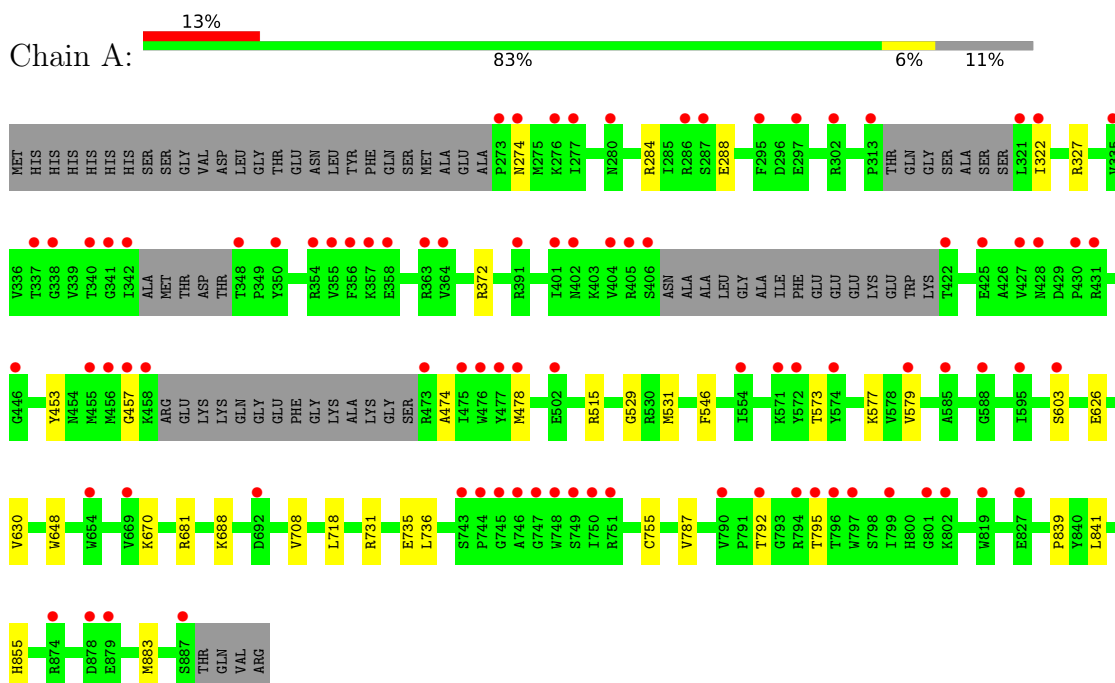
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	466	Total O 466 466	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: RNA-directed RNA polymerase NS5



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.03Å 82.89Å 69.43Å 90.00° 114.52° 90.00°	Depositor
Resolution (Å)	19.69 – 1.40 20.41 – 1.40	Depositor EDS
% Data completeness (in resolution range)	97.6 (19.69-1.40) 97.6 (20.41-1.40)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 1.40Å)	Xtrriage
Refinement program	BUSTER 2.10.3 (3-OCT-2019)	Depositor
R, $R_{free}$	0.196 , 0.218 0.194 , 0.216	Depositor DCC
$R_{free}$ test set	6321 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.8	Xtrriage
Anisotropy	0.095	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 61.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.013 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	5309	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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: ZN, EDO

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	0.43	0/4949	0.57	0/6689

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 [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	4829	0	4724	21	0
2	A	12	0	18	0	0
3	A	2	0	0	0	0
4	A	466	0	0	3	0
All	All	5309	0	4742	21	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 (21) 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:787:VAL:HG11	1:A:883:MET:HB2	1.80	0.63
1:A:787:VAL:HG11	1:A:883:MET:CB	2.33	0.58
1:A:731:ARG:HD3	1:A:736[A]:LEU:HD21	1.86	0.57
1:A:731:ARG:NH2	1:A:735:GLU:HG2	2.19	0.57
1:A:322:ILE:HG21	1:A:327:ARG:HD2	1.89	0.55
1:A:531:MET:HE2	1:A:708:VAL:HG21	1.89	0.55
1:A:529:GLY:O	1:A:670:LYS:HE3	2.07	0.54
1:A:626:GLU:OE1	1:A:681[B]:ARG:HG2	2.11	0.51
1:A:457:GLY:HA2	1:A:474:ALA:HA	1.94	0.50
1:A:718[B]:LEU:HD21	1:A:839:PRO:HG2	1.94	0.49
1:A:453:TYR:HB2	1:A:579:VAL:HG22	1.94	0.48
1:A:718[B]:LEU:CD2	1:A:839:PRO:HG2	2.46	0.46
1:A:573:THR:O	1:A:577:LYS:HG2	2.15	0.46
1:A:515[B]:ARG:HG2	4:A:1348:HOH:O	2.16	0.45
1:A:718[A]:LEU:HD21	1:A:841:LEU:HD23	1.99	0.45
1:A:688:LYS:HE3	4:A:1005:HOH:O	2.18	0.44
1:A:372:ARG:HD3	1:A:546:PHE:CD1	2.53	0.43
1:A:755:CYS:HB2	1:A:792:THR:HG22	2.00	0.43
1:A:630:VAL:HG21	1:A:648:TRP:CD1	2.54	0.42
1:A:372:ARG:HD3	1:A:546:PHE:CE1	2.55	0.41
1:A:855:HIS:HE1	4:A:1425:HOH:O	2.05	0.40

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	579/645 (90%)	569 (98%)	10 (2%)	0	<a href="#">100</a> <a href="#">100</a>

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	518/560 (92%)	512 (99%)	6 (1%)	<span style="border: 1px solid blue; padding: 2px;">71</span> <span style="border: 1px solid gray; padding: 2px;">47</span>

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	274	ASN
1	A	284	ARG
1	A	288	GLU
1	A	478	MET
1	A	603	SER
1	A	795	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	274	ASN
1	A	825	HIS
1	A	870	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](#)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

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	$\# Z  > 2$
2	EDO	A	902	-	3,3,3	0.82	0	2,2,2	0.30	0
2	EDO	A	903	-	3,3,3	0.57	0	2,2,2	0.07	0
2	EDO	A	901	-	3,3,3	0.54	0	2,2,2	0.30	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	A	902	-	-	0/1/1/1	-
2	EDO	A	903	-	-	0/1/1/1	-
2	EDO	A	901	-	-	0/1/1/1	-

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 [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	574/645 (88%)	0.94	87 (15%) <b>2</b> <b>1</b>	15, 25, 51, 64	6 (1%)

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	746	ALA	15.2
1	A	797	TRP	14.2
1	A	745	GLY	11.8
1	A	477	TYR	11.4
1	A	748	TRP	11.1
1	A	350	TYR	10.9
1	A	747	GLY	9.9
1	A	342	ILE	8.9
1	A	422	THR	8.0
1	A	473	ARG	6.8
1	A	794	ARG	6.5
1	A	795	THR	6.4
1	A	796	THR	6.3
1	A	321	LEU	6.0
1	A	431	ARG	6.0
1	A	456	MET	5.8
1	A	406	SER	5.7
1	A	457	GLY	5.5
1	A	887	SER	5.5
1	A	276	LYS	5.3
1	A	744	PRO	5.2
1	A	364	VAL	5.0
1	A	337	THR	4.9
1	A	273	PRO	4.8
1	A	692	ASP	4.6
1	A	405	ARG	4.2
1	A	801	GLY	4.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	313	PRO	4.0
1	A	572	TYR	3.9
1	A	879	GLU	3.8
1	A	357	LYS	3.8
1	A	302	ARG	3.8
1	A	750	ILE	3.8
1	A	654	TRP	3.8
1	A	341	GLY	3.7
1	A	455[A]	MET	3.7
1	A	475	ILE	3.7
1	A	348	THR	3.6
1	A	354	ARG	3.6
1	A	338	GLY	3.6
1	A	588	GLY	3.6
1	A	355	VAL	3.6
1	A	476	TRP	3.6
1	A	799	ILE	3.5
1	A	274	ASN	3.4
1	A	603	SER	3.2
1	A	425	GLU	3.1
1	A	340	THR	3.0
1	A	579	VAL	3.0
1	A	402	ASN	2.8
1	A	363	ARG	2.8
1	A	428	ASN	2.7
1	A	802	LYS	2.7
1	A	878	ASP	2.6
1	A	743	SER	2.6
1	A	792	THR	2.5
1	A	478	MET	2.5
1	A	356	PHE	2.5
1	A	749	SER	2.5
1	A	458	LYS	2.4
1	A	595	ILE	2.4
1	A	404	VAL	2.4
1	A	277	ILE	2.4
1	A	391	ARG	2.4
1	A	295	PHE	2.3
1	A	430	PRO	2.3
1	A	427	VAL	2.3
1	A	669	VAL	2.3
1	A	322	ILE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	819	TRP	2.2
1	A	827	GLU	2.2
1	A	446	GLY	2.2
1	A	335	VAL	2.2
1	A	287	SER	2.2
1	A	358	GLU	2.2
1	A	502	GLU	2.2
1	A	401	ILE	2.2
1	A	790	VAL	2.2
1	A	554	ILE	2.1
1	A	874	ARG	2.1
1	A	574	TYR	2.1
1	A	280	ASN	2.1
1	A	571	LYS	2.1
1	A	751	ARG	2.0
1	A	297	GLU	2.0
1	A	286	ARG	2.0
1	A	585	ALA	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 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(Å <sup>2</sup> )	Q<0.9
2	EDO	A	902	4/4	0.76	0.18	37,38,38,38	0
2	EDO	A	901	4/4	0.81	0.12	52,52,52,52	0
2	EDO	A	903	4/4	0.88	0.13	29,31,33,35	0
3	ZN	A	904	1/1	0.98	0.07	26,26,26,26	0
3	ZN	A	905	1/1	1.00	0.06	17,17,17,17	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.